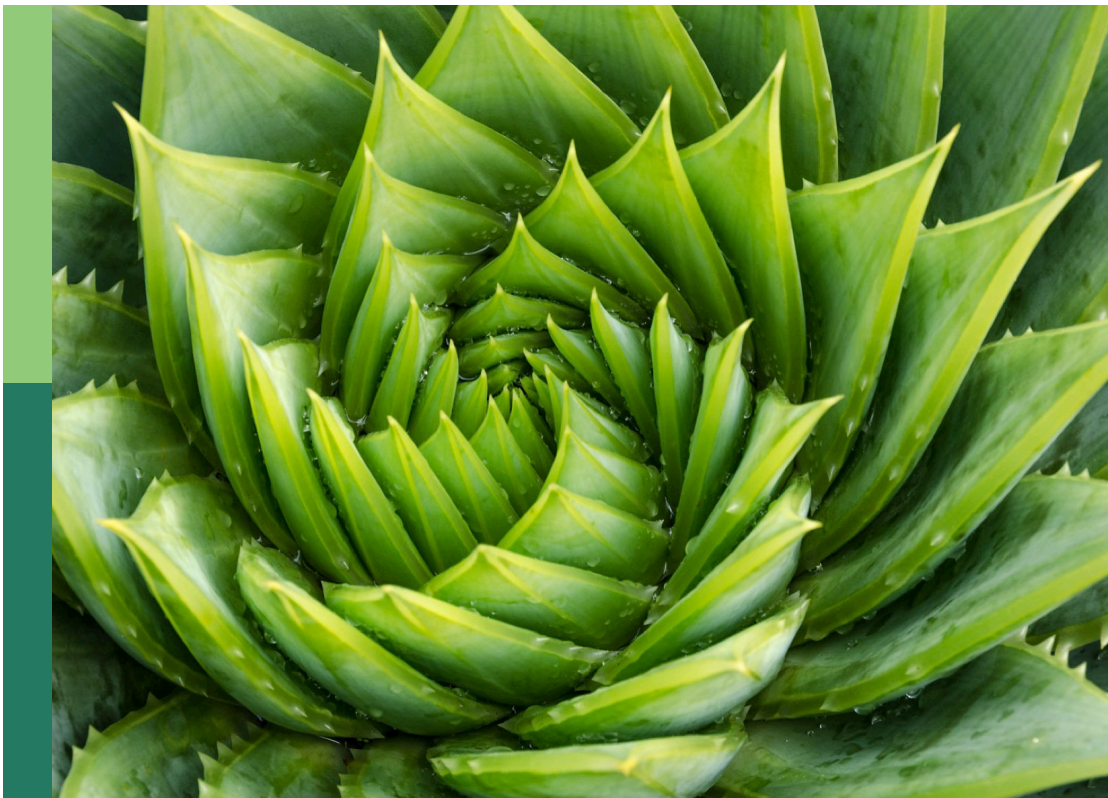
frontiers | Research Topics

Abiotic stress: Molecular genetics and genomics, volume II

Edited by
Mukesh Jain, Rohini Garg, Rajeev K. Varshney and Prasanta Kumar Subudhi

Published in
Frontiers in Plant Science



ISSN 1664-8714
ISBN 978-2-83251-541-9
DOI 10.3389/978-2-83251-541-9

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Abiotic stress: Molecular genetics and genomics, volume II

Topic editors

Mukesh Jain — Jawaharlal Nehru University, India
Rohini Garg — Shiv Nadar University, India
Rajeev K. Varshney — Murdoch University, Australia
Prasanta Kumar Subudhi — Louisiana State University, United States

Citation

Jain, M., Garg, R., Varshney, R. K., Subudhi, P. K., eds. (2023). *Abiotic stress: Molecular genetics and genomics, volume II*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83251-541-9

Table of contents

frontiers | Frontiers in Plant Science

TYPE Editorial
PUBLISHED 18 January 2023
DOI 10.3389/fpls.2022.1101139

OPEN ACCESS

EDITED AND REVIEWED BY
Luisa M. Sandalio,
Spanish National Research Council (CSIC), Spain

*CORRESPONDENCE
Rohini Garg
rohini.garg@snu.edu.in

SPECIALTY SECTION
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

RECEIVED 17 November 2022
ACCEPTED 09 December 2022
PUBLISHED 18 January 2023

CITATION
Garg R, Subudhi PK, Varshney RK and Jain M (2023) Editorial: Abiotic stress: Molecular genetics and genomics, volume II.
Front. Plant Sci. 13:1101139.
doi: 10.3389/fpls.2022.1101139

Editorial: Abiotic stress: Molecular genetics and genomics, volume II

Rohini Garg[1]*, Prasanta K. Subudhi[2], Rajeev K. Varshney[3] and Mukesh Jain[4]

[1]Department of Life Sciences, School of Natural Sciences, Shiv Nadar University, Gautam Buddh Nagar, India, [2]School of Plant, Environmental, and Soil Sciences, Louisiana State University Agricultural Center, Baton Rouge, LA, United States, [3]Centre for Crop & Food Innovation, State Agricultural Biotechnology Centre, Food Futures Institute, Murdoch University, Perth, WA, Australia, [4]School of Computational & Integrative Sciences, Jawaharlal Nehru University, New Delhi, India

KEYWORDS

abiotic stress, functional genomics, gene regulatory networks, climate change, comparative transcriptomics, genome-wide association studies, phylogenetic analysis

Editorial on the Research Topic
Abiotic stress: Molecular genetics and genomics, volume II

Abiotic stresses constitute major threat to farming system worldwide. The ongoing climate change is further exacerbating the global farming landscape due to increased frequency and intensity of abiotic stresses leading to reduced productivity and stability in crop plants (Ashikari and Ma, 2015; Hussain et al., 2019). In recent years, there has been a remarkable improvement in crop productivity due to development and implementation of innovative breeding and genetic tools and technologies (Varshney et al., 2021). However, increasing human population and rising living standards are expected to increase the global food demand in coming years. Therefore, further agricultural innovations are required to meet this challenge. Since most cultivars have been developed to perform well under optimal environments with minimal perturbations (Kukal and Irmak, 2018; Bharadwaj et al., 2021), there is a need for more stress-tolerant crop varieties to sustain crop productivity under adverse environments (Raza et al., 2021).

Climate variations threaten both global food security (Godfray et al., 2010) and sustainability of the farming system (Porfirio et al., 2018). The transformation of agriculture towards sustainability inspires research to mitigate the impact of climate change induced abiotic stresses (Wheeler and Von Braun, 2013). Among others, use of cultivars with enhanced adaptation to abiotic stresses is the most logical and economical approach to have significant impact on both sustainability and food security at a global scale (Chaturvedi et al., 2017). Therefore, understanding the plants' response to various abiotic stresses at the whole genome level, discovery and characterization of important natural variants, and elucidation of abiotic stress adaptation mechanisms using modern

genomic tools is crucial for designing next-generation climate-resilient crop cultivars (Garg et al., 2016; Palit et al., 2020; Rajkumar et al., 2022).

The current Research Topic "Abiotic Stress: Molecular Genetics and Genomics, Volume II" encompasses a collection of 30 original research articles and one review on wide range of topics, such as genome-wide survey of key abiotic stress tolerance genes and their characterization, genome-wide association studies (GWAS), comparative genome-scale transcriptomic, ionic, degradome, microRNA profiling, selection of key salt tolerance genes during breeding process, and abiotic stress tolerance mechanisms. These research articles provide novel insights into plants' responses and adaptation to several abiotic stresses, such as salinity, alkalinity, drought, temperature extremes, nutrient and metal toxicity, in a variety of plant species.

Gene family analysis for discovery of candidate gene(s) implicated in abiotic stress tolerance

The analysis of a set of evolutionary related homologous genes (gene families) can provide important insights into their distinct/overlapping functions and identify candidate gene(s) involved in important biological processes including abiotic stress response (Jain et al., 2010; Singh and Jain; Singh et al., 2017). The article by Wang et al. reported identification of 24 metal tolerance protein (MTP) encoding genes in peanut (*AhMTP*) in a genome-wide survey followed by detailed phylogenetic relationship, gene structure, protein structure, and gene expression analyses. The authors reported that the differential response of *AhMTP* genes to Fe, Cd, and Zn exposure in two peanut cultivars with contrasting response to metal toxicity may be due to differential metal translocation from roots to shoots. In another study, Yuan et al. identified 39 members of cysteine synthase (*CSase*) gene family in alfalfa (*Medicago sativa* L.). The authors performed a systematic phylogeny, gene structure, conserved domain and synteny analysis of this gene family. The overexpression of a *CSase* gene improved alkali tolerance by increasing the antioxidant and osmolyte production in alfalfa. A genome-wide analysis of late embryogenesis abundant (LEA) proteins in mung bean, adzuki bean and cowpea by Singh et al. provided insights into their structural and functional diversity in the three *Vigna* species. One interesting finding of this study was that LEA-6 group was missing in the mung genome and all seven groups were preset in the cowpea genome. The gene expression studies involving seven mung bean genes demonstrated their role in heat stress response. In a comprehensive survey of soybean genome, Xu et al. identified 22 TLP (tubby-like protein) genes which were analyzed for their phylogenetic relationship, gene structure and motif analyses. The functional characterization of a candidate gene, *GmTLP8*, demonstrated its role in drought and salinity stress responses by triggering the downstream stress-responsive genes. Likewise, another genome-wide study by Liu et al. identified 38 *GmTIFY* transcription factor genes. The overexpression of two of these genes (*GmTIFY10e* and *GmTIFY10g*) in transgenic *Arabidopsis* and soybean plants showed improved salt tolerance compared with wild-type plants, whereas the RNAi lines exhibited enhanced sensitivity to salt stress. Further, evidence of the involvement of these genes in the ABA signaling pathway was also provided.

The study by Huang et al. reported identification of 26 members of the plant-specific Teosinte Branched1/Cycloidea/Proliferating Cell Factor (TCP) transcription factor family in Pak-choi [*Brassica campestris* (syn. *Brassica rapa*) ssp. *chinensis* var. communis]. The gene expression analysis revealed the differential expression of several members in response to different types of abiotic stresses. In another article, Li et al. provided a comprehensive analysis of cysteine-rich polycomb-like protein (CPP) family transcription factors in *Brassica napus* and its two diploid progenitors. The analysis suggested that whole genome duplication and transposed duplication might be responsible for the expansion of CPP gene family in *B. napus* during allopolyploidization and few of *BnCPP* genes undergo neo- or sub-functionalization. The expression analysis under salinity stress revealed the adaptive advantage of allopolyploid *B. napus* compared with the diploid progenitors.

The study by Yu et al. reported 37 Calcineurin B-like protein-interacting protein kinase (CIPK) encoding genes in the ornamental woody plant *Lagerstroemia indica* and performed various analyses thereof. The collinearity and synonymous substitution rate analyses revealed that most of duplicated *LiCIPK*s were retained by the two whole genome duplication events. Several *LiCIPK*s exhibited differential expression under different abiotic stress conditions. Further, role of *LiCIPK30* in improving salt and osmotic tolerance was demonstrated *via* its overexpression in *Arabidopsis*. The above studies provide an important resource for the prioritization of candidate genes for further investigations into their function and mechanism of action in abiotic stress responses.

Candidate gene based studies for improvement of abiotic stress tolerance

Chang et al. successfully cloned and characterized a *WUSCHEL-RELATED HOMEOBOX GENE 11* (*WOX11*) gene from a hybrid walnut (*JrWOX11*), which was found induced by abscisic acid (ABA), salt, and polyethylene glycol. Based on gene

expression and overexpression studies, role of *JrWOX11* in improving salinity and osmotic tolerance through enhanced root system was demonstrated. This study provided the molecular basis of differentiating the trees that are difficult-to-root and easy-to-root. Li et al. characterized a novel NAC transcription factor (*NtNAC053*) in tobacco, which was induced by salt and drought stresses, localized in the nucleus, and acts as a transcriptional activator. This study demonstrated that enhanced tolerance to drought and salt stresses in *NtNAC053* overexpressing transgenic tobacco plants could be due to enhanced antioxidant system through activation of downstream stress-responsive genes. Lu et al. provided evidence for the existence of two isoforms, *ZmPP2C26L* and *ZmPP2C26S*, of *ZmPP2C26*, a clade B member of maize PP2C family. The *zmpp2c26* mutant exhibited enhanced drought tolerance, whereas overexpression of *ZmPP2C26L* and *ZmPP2C26S* significantly decreased drought tolerance in *Arabidopsis* and rice. The authors suggested dephosphorylation of ZmMAPK3 and ZmMAPK7 by the ZmPP2C26 protein as a possible mechanism to reduce drought tolerance and photosynthesis activity. The functional characterization of a valine-glutamine motif-containing gene in wheat, *TaVQ14*, was performed by Cheng et al. to reveal its role in salt and drought tolerance. The *TaVQ14* overexpressing lines in *Arabidopsis* showed improved salt and drought tolerance *via* scavenging reactive oxygen species and calcium signaling. Qin et al. focused on functional characterization of a transcription factor gene *HbMYB44* of rubber tree (*Hevea brasiliensis* Müll. Arg), which enhanced tolerance to salinity, drought, and osmotic stresses in overexpressing *Arabidopsis* plants. This gene also helped recovery of root damage in the overexpression plants by application of phytohormones such as ABA, methyl jasmonic acid, gibberellic acid, and salicylic acid which suggested *HbMYB44*'s versatile role in regulating multiple phytohormone signaling and stress tolerance pathways. Wang et al. demonstrated the role of Mitogen-activated Protein Kinase 1 in improving shading tolerance in rapeseed (*Brassica napus*) *via* increased photosynthetic capacity in *BnaMAPK1*-overexpressing plants. RNA sequencing revealed that *BnaMAPK1* positively regulated photosynthesis capability possibly by controlling antenna protein complex in photosystem II to respond to shading stress. Further, *BnaLHCB3* was identified as an interacting partner of BnaMAPK1 *via* yeast two-hybrid and split-luciferase complementation assays.

Jin et al. demonstrated that the ectopic expression of a soybean NHX gene, *GmNHX6* (encoding a Golgi-localized sodium/hydrogen exchanger), enhanced alkaline tolerance in *Arabidopsis* and soybean by maintaining high K^+ content and low Na^+/K^+ ratio. A natural sequence variation in the promoter region of *GmNHX6* was associated with the alkaline tolerance in soybean germplasm and the promoter of *GmNHX6* isolated from an alkaline tolerant soybean variety exhibited stronger activity in response to alkali stress. In another study, Sun et al. utilized a T-DNA insertion mutant of *OsCIPK18* (*cipk18*) encoding a CBL-interacting protein kinase and defined an *OsCIPK18*-dependent transcriptomic network involved in ammonium toxicity response. In addition, the role of *OsCIPK18* as a key node in auxin and ABA signaling pathways under ammonium stress was proposed. Zhang et al. showed improvement in tolerance to drought and ABA in transgenic *Arabidopsis* plants overexpressing a glutathione S-transferase (*CsGSTU8*) from tea plant (*Camellia sinensis*). Further molecular insights were provided by demonstrating the binding of a trancription factor, CsWRKY48, to the promoter of *CsGSTU8* to regulate its induction under drought stress and ABA treatment.

Genome-wide studies for discovery of candidate genes involved in cold/freezing stress

Cold and/or freezing stress is one of the major environmental factors, that limits the productivity of several plants. In the study by Wang et al., transcriptome profiling of leaves of two alfalfa genotypes with contrasting responses under freezing stress (-10°C) followed by co-expression network analysis revealed the importance of ATP-binding cassette (ABC) C subfamily genes, *ABCC8* and *ABCC3*, in freezing tolerance. Further, this study also demonstrated the contribution of Ca^{2+} signal transduction and CBF/DREB1 related genes towards tolerance to freezing stress. In another study, Islam et al. compared the transcriptional landscape in pseudostem and leaf blade tissues of endophyte-positive (E+) and endophyte-free (E−) tall fescue (*Festuca arundinacea*), a cool-season perennial grass, at three diurnal temperature conditions. The differential gene expression profiling revealed eight candidate genes, including orthologs of rice phytochrome A, phytochrome C, and ethylene receptor genes, which might be the possible route underlying freezing tolerance in tall fescue. A comparative transcriptomic analysis involving maize genotypes with contrasting response to low-temperature stress by Meng et al. revealed that both photosynthesis and antioxidant metabolism pathways played important role in conferring cold tolerance during seed germination stage. This was supported by data on increased antioxidant capacity in resistant line compared with the susceptible line. Further, Tang et al. demonstrated that cold treatment induced global DNA demethylation in *Hevea brasiliensis* and demethylation in the upstream regions of the genes was associated with higher gene expression.

Studies addressing heavy metal stress response/tolerance

Heavy metal contamination not only reduces crop yield significantly, but also poses risks to human health. Identifying the molecular mechanisms heavy metal uptake can help in developing plants for phytoremediation as well as crops with reduced accumulation of such toxic metals. To investigate the molecular mechanism of cadmium (Cd) stress tolerance, Wang et al. performed RNA-seq analysis in *Tamarix hispida* treated with Cd stress for different time points. The functional annotation of differentially expressed genes identified genes involved in ion binding, signal transduction, stress sensing, hormone responses and ROS metabolism contributing toward Cd stress tolerance. Further, *ThUGT* from the ABA-signaling pathway was identified as a candidate gene to improve Cd stress tolerance by reducing Cd uptake and regulation of ROS. Another study by Paape et al. used GWAS approach with the seedlings of *Medicago truncatula* HapMap collection exposed to Cd and mercury (Hg) stress revealed significant genetic diversity for these phenotypic traits. Some important candidate genes in the QTL regions included, membrane associated ATP-binding cassette transporters, P-type ATPase transporters, oxidative stress response genes, and stress related UDP-glycosyltransferases, which can be the useful targets to design plants with reduced heavy metal accumulation. The study also suggested to exploit wild accessions of *Medicago* for genetic improvement due to macroevolutionary conservation of heavy metal and stress response genes in this model plant.

Lee et al. analyzed the interactions between arsenic (As) and eight essential ions in a rice core collection under non-stress and stress conditions to elucidate the impact of environmental and genotypic differences, and identified the genetic factors regulating As accumulation. This GWAS study provides evidence that *indica* populations are superior in reducing As accumulation compared with *japonica* populations. A potential candidate gene, *AIR2* (arsenic-induced RING finger protein), whose expression was lower in *indica* compared with *japonica* subspecies, was suggested for marker-assisted selection in developing rice varieties with improved grain quality.

Studies providing insights into salinity and/or osmotic stress response/tolerance

A comparative analysis of transcriptomes of two contrasting clones (R7, salt-tolerant and S4, salt-sensitive) of *Fraxinus velutina* reported by Ma et al. revealed the upregulation of several stress-responsive genes in the salt-tolerant clone. Salt stress induced the expression of genes involved in proline biosynthesis, starch and sucrose metabolism, and those encoding antioxidant enzymes, which might contribute towards enhanced salt tolerance. Further, leaves and roots of both of these clones were subjected to miRNA and degradome analysis by Liu et al. to understand the role of miRNAs in defense response of plants to salt stress. This study revealed multiple and somewhat distinct miRNA/target modules regulating different biological processes in leaves (antioxidant system and auxin signaling) and roots (ROS scavenging, cell proliferation, and ion homeostasis) under salt stress. Based on GWAS of various growth and agronomic traits, genetic basis of salt-alkali tolerance was investigated by Zhang et al. and at least nine significant QTLs and 20 candidate genes related to salt-alkali stress tolerance were identified. By coupling the sequence variation, annotation and differential expression, few important candidate genes, such as *BnABA4, BnBBX14, BnVTI12, BnPYL8*, and *BnCRR1* were identified for breeding salt-alkali-tolerant *B. napus* varieties. Guan et al. identified five haplotypes of *GmSALT3* by using genome resequencing data from 279 Chinese soybean landraces. Using five PCR-based haplotype-specific markers developed in this study, the authors demonstrated their efficiency in distinguishing salt-tolerant and salt-sensitive soybean lines and tracing the salt-tolerant haplotype in soybean pedigree.

In a study by Galić et al., a link between biochemistry and genetics of osmotic stress tolerance in maize plants was established *via* investigating the variability in responses of a panel of elite maize inbred lines for the stress-related traits at the seedling stage. The overall analysis revealed genomic regions linked to stress responsive traits that harbor the genes associated with osmotic-stress signaling, osmolyte accumulation and regulation of peroxisomes gene ontology terms. The integrated transcriptomics and antioxidant profiling of two contrasting Chinese chestnut (*Castanea mollissima* BL.) varieties in response to gall wasp *Dryocosmus kuriphilus* (GWDK) infestation at different time points by Zhu et al. revealed new insights into the chestnut-GWKD interactions and identified candidate genes for further functional validation and molecular-aided breeding of gall wasp-resistant chestnut varieties.

The review article by Zhang et al. included in this Research Topic summarized the mechanisms associated with adaptative response to salinity, drought, and cold stresses as well as crosstalk among them in the model legume, *M. truncatula*. The genetic and molecular resources provided in this review should be useful for investigating and improving abiotic stress tolerance in legume crops. Future investigation on the impact of combination of abiotic stresses and use of wild species (*M. ruthenica*) is suggested for the retention of abiotic stress tolerance.

Due to increased awareness for improving sustainability of farming systems with minimal carbon footprints, it has become imperative to design climate smart crop varieties. However, accomplishing this goal is challenging due to involvement of multiple genes and pathways, and interactions among them. Despite tremendous advances made during last few decades,

there are still gaps in comprehensive understanding of the plant's response to abiotic stresses at molecular level. The collection of articles highlighting the relevance of genome-wide discovery of key genes associated with abiotic stress adaptation and their characterization in a wide range of plants using genomic approaches clearly contribute toward that goal. Since a tremendous number of natural variations for tolerance to abiotic stresses exists in the available cultivated and wild germplasm resources, future research should focus on mining of key abiotic stress related genes and their superior alleles, and gene pyramiding to develop crop varieties adapted to multiple stresses. The articles presented in this special issue not only enrich our understanding of the molecular basis of plants' adaptative responses to abiotic stresses, but also should help in successful breeding of crop varieties adapted to climate-change using marker-assisted selection and genome editing tools.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work, and approved it for publication.

Funding

This work was supported by the Department of Biotechnology, Government of India (RG and MJ), Science and Engineering Research Board, Government of India (MJ), United States Department of Agriculture-National Institute of Food and Agriculture (Grant No. 2018-67013-27618 to PS), Food Futures Institute of Murdoch University (RV).

Acknowledgments

We would like to express our sincere gratitude and appreciation to all authors for their excellent contributions and reviewers for their evaluations and constructive comments to make this Research Topic successful.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Ashikari, M., and Ma, J. F. (2015). Exploring the power of plants to overcome environmental stresses. *Rice* 8, 10. doi: 10.1186/s12284-014-0037-y

Bharadwaj, C., Tripathi, S., Soren, K. R., Thudi, M., Singh, R. K., Sheoran, S., et al. (2021). Introgression of "QTL-hotspot" region enhances drought tolerance and grain yield in three elite chickpea cultivars. *Plant Genome* 14, e20076. doi: 10.1002/tpg2.20076

Chaturvedi, A. K., Bahuguna, R. N., Shah, D., Pal, M., and Jagadish, S. V. K. (2017). High temperature stress during flowering and grain filling offsets beneficial impact of elevated CO2 on assimilate partitioning and sink-strength in rice. *Sci. Rep.* 7, 8227. doi: 10.1038/s41598-017-07464-6

Garg, R., Shankar, R., Thakkar, B., Kudapa, H., Ksrishnamurthy, L., Mantri, N., et al. (2016). Transcriptome analyses reveal genotype- and developmental stage-specific molecular responses to drought and salinity stresses in chickpea. *Sci. Rep.* 6, 19228. doi: 10.1038/srep19228

Godfray, H. C. J., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., Nisbett, N., et al. (2010). The future of the global food system. *Philos. Trans. R. Soc. B: Biol. Sci.* 365 (1554), 2769–2777. doi: 10.1098/rstb.2010.0180

Hussain, H. A., Men, S., Hussain, S., Chen, Y., Ali, S., Zhang, S., et al. (2019). Interactive effects of drought and heat stresses on morpho-physiological attributes, yield, nutrient uptake, and oxidative status in maize hybrids. *Sci. Rep.* 9, 3890. doi: 10.1038/s41598-019-40362-7

Jain, M., Ghanashyam, C., and Bhattacharjee, A. (2010). Comprehensive expression analysis suggests overlapping and specific roles of rice glutathione s-transferase genes during development and stress responses. *BMC Genomics* 11, 73. doi: 10.1186/1471-2164-11-73

Kukal, M. S., and Irmak, S. (2018). Climate-driven crop yield and yield variability and climate change impacts on the US great plains agricultural production. *Sci. Rep.* 8, 1–18. doi: 10.1038/s41598-018-21848-2

Palit, P., Kudapa, H., Zougmore, R., Kholova, J., Whitbread, A., Sharma, M., et al. (2020). An integrated research framework combining genomics, systems biology, physiology, modelling and breeding for legume improvement in response to elevated CO_2 under climate change scenario. *Curr. Plant Biol.* 22, 100149. doi: 10.1016/j.cpb.2020.100149

Porfirio, L. L., Newth, D., Finnigan, J. J., and Cai, Y. (2018). Economic shifts in agricultural production and trade due to climate change. *Palgrave Commun.* 4 (1), 1–9. doi: 10.1057/s41599-018-0164-y

Rajkumar, M. S., Garg, R., and Jain, M. (2022). Genome-wide discovery of DNA polymorphisms *via* resequencing of chickpea cultivars with contrasting response to drought stress. *Physiol. plant.* 174 (1), e13611. doi: 10.1111/ppl.13611

Raza, A., Tabassum, J., Kudapa, H., and Varshney, R. K. (2021). Can omics deliver temperature resilient ready-to-grow crops? *Crit. Rev. Biotechnol.* 41, 1209–1232. doi: 10.1080/07388551.2021.1898332

Singh, V. K., Rajkumar, M. S., Garg, R., and Jain, M. (2017). Genome-wide identification and co-expression network analysis provide insights into the roles of auxin response factor gene family in chickpea. *Sci. Rep.* 7, 10895. doi: 10.1038/s41598-017-11327-5

Varshney, R. K., Bohra, A., Roorkiwal, M., Barmukh, R., Cowling, W. A., Chitikineni, A., et al. (2021). Fast-forward breeding for a food-secure world. *Trends Genet.* 37, 1124–1136. doi: 10.1016/j.tig.2021.08.002

Wheeler, T., and Von Braun, J. (2013). Climate change impacts on global food security. *Science* 341, 508–513. doi: 10.1126/science.1239402

frontiers in Plant Science

ORIGINAL RESEARCH
published: 16 November 2021
doi: 10.3389/fpls.2021.794241

Selection of the Salt Tolerance Gene *GmSALT3* During Six Decades of Soybean Breeding in China

Rongxia Guan[1†], Lili Yu[1†], Xiexiang Liu[1†], Mingqiang Li[1], Ruzhen Chang[1], Matthew Gilliham[2] and Lijuan Qiu[1*]*

[1] The National Key Facility for Crop Gene Resources and Genetic Improvement, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing, China, [2] Australian Research Council Centre of Excellence in Plant Energy Biology, School of Agriculture, Food and Wine, Waite Research Institute, The University of Adelaide, Glen Osmond, SA, Australia

OPEN ACCESS

Edited by:
Rajeev K. Varshney, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), India

Reviewed by:
Dawei Xin, Northeast Agricultural University, China
Qingshan Chen, Northeast Agricultural University, China

***Correspondence:**
Rongxia Guan
guanrongxia@caas.cn
Lijuan Qiu
qiulijuan@caas.cn

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

Received: 13 October 2021
Accepted: 25 October 2021
Published: 16 November 2021

Citation:
Guan R, Yu L, Liu X, Li M, Chang R, Gilliham M and Qiu L (2021) Selection of the Salt Tolerance Gene GmSALT3 During Six Decades of Soybean Breeding in China. Front. Plant Sci. 12:794241. doi: 10.3389/fpls.2021.794241

Salt tolerance is an important trait that affects the growth and yield of plants growing in saline environments. The salt tolerance gene *GmSALT3* was cloned from the Chinese soybean cultivar Tiefeng 8, and its variation evaluated in Chinese wild soybeans and landraces. However, the potential role of *GmSALT3* in cultivation, and its genetic variation throughout the history of Chinese soybean breeding, remains unknown. Here we identified five haplotypes of *GmSALT3* in 279 Chinese soybean landraces using a whole genome resequencing dataset. Additionally, we developed five PCR-based functional markers: three indels and two cleaved amplified polymorphic sequences (CAPS) markers. A total of 706 Chinese soybean cultivars (released 1956–2012), and 536 modern Chinese breeding lines, were genotyped with these markers. The Chinese landraces exhibited relatively high frequencies of the haplotypes H1, H4, and H5. H1 was the predominant haplotype in both the northern region (NR) and Huanghuai region (HHR), and H5 and H4 were the major haplotypes present within the southern region (SR). In the 706 cultivars, H1, H2, and H5 were the common haplotypes, while H3 and H4 were poorly represented. Historically, H1 gradually decreased in frequency in the NR but increased in the HHR; while the salt-sensitive haplotype, H2, increased in frequency in the NR during six decades of soybean breeding. In the 536 modern breeding lines, H2 has become the most common haplotype in the NR, while H1 has remained the highest frequency haplotype in the HHR, and H5 and H1 were highest in the SR. Frequency changes resulting in geographically favored haplotypes indicates that strong selection has occurred over six decades of soybean breeding. Our molecular markers could precisely identify salt tolerant (98.9%) and sensitive (100%) accessions and could accurately trace the salt tolerance gene in soybean pedigrees. Our study, therefore, not only identified effective molecular markers for use in soybean, but also demonstrated how these markers can distinguish *GmSALT3* alleles in targeted breeding strategies for specific ecoregions.

Keywords: soybean, salt tolerance, *GmSALT3*, haplotype, markers

INTRODUCTION

Salt stress is a major environmental factor affecting agricultural plant productivity, which in turn, can threaten food security (Zhu, 2016). The ability of plants to complete their life cycles while growing in saline environments is the ultimate measure of salt tolerance (Parida and Das, 2005). Salt tolerance variation in soybean, a moderately salt-tolerant crop, has been studied throughout the last half century (Abel and MacKenzie, 1964; Fehr et al., 1971; Läuchli and Wieneke, 1979; Durand and Lacan, 1994; El-Samad and Shaddad, 1997; An et al., 2002). Salt tolerance in soybean seedlings is reportedly associated with a dominant gene proposed as *Ncl* (Abel, 1969). Additionally, the major salt tolerance quantitative trait locus (QTL) was mapped repeatedly on chromosome 3 (Chr. 3) in different salt tolerant germplasms (Lee et al., 2004; Hamwieh and Xu, 2008; Hamwieh et al., 2011; Ha et al., 2013; Guan et al., 2014a). The candidate gene *Glyma03g32900.1* has been proposed to underpin the conserved locus on Chr. 3. Researchers have isolated *Glyma03g32900.1* from different soybean germplasm, defined independently as *GmCHX1*, *GmSALT3,* and *Ncl* (Guan et al., 2014b; Qi et al., 2014; Do et al., 2016). GmSALT3 has been shown to limit Na^+ and Cl^- accumulation in soybean shoots, thereby increasing soybean salt tolerance (Liu et al., 2016; Qu et al., 2021). Additionally, novel putative loci related to chloride and leaf chlorophyll concentration have been mapped on soybean chromosomes 2, 13, 14, 16, and 20, through QTL mapping and genome-wide association studies (Zeng et al., 2017; Do et al., 2018). Furthermore, Zhang et al. (2019) identified a cation diffusion facilitator, *GmCDF1* (*Glyma08g10200*), related with salt tolerance at the germination stage, but no interaction between this candidate gene and *Glyma03g32900.1* was observed. While these studies have bolstered genetic resources for breeding salt tolerant soybeans (Zeng et al., 2017; Do et al., 2018), much remains to be learned.

Genetic diversity in *GmCHX1* has been evaluated in 23 genetically distinct soybeans; a conserved coding sequence was observed in the salt-tolerant germplasm with various genotypes detected within the 12 salt-sensitive soybeans (Qi et al., 2014). Our previous research revealed nine haplotypes (H1–H9) in *GmSALT3*, with H1 observed in salt tolerant plants and H2–H5 observed in salt-sensitive soybean landraces (Guan et al., 2014b). However, little is known about the distribution of these *GmSALT3* alleles more widely across Chinese soybean cultivars.

Intriguingly, whole-genome resequencing of 106 soybeans revealed three structural variants of *GmCHX1*, of which SV-1 and SV-2, respectively, correspond to H2 and H1 of the nine Chinese soybean germplasm haplotypes, and SV-3 was a novel allele (Patil et al., 2016). Using coding regions of 216 soybean accessions from South Korea, China, and Japan, 40 haplotypes were observed, nine of which were observed in *Glycine max* accessions from China (Lee et al., 2018). Such results have encouraged us to evaluate the genomic variation of *GmSALT3* widely across representative Chinese soybean landraces, to identify novel variants, and to gain insights into how different alleles have been selected during Chinese soybean breeding over the years. Specifically, the objectives of this study were to (1) analyze genomic variation for *GmSALT3* using whole genome resequencing data of 279 representative Chinese soybean landraces, (2) develop a set of PCR-based markers and assess marker accuracy for the selection of the *GmSALT3* gene, and (3) explore how the haplotypes of *GmSALT3* have been selected and promoted in soybeans during historical breeding process.

MATERIALS AND METHODS

Plant Materials

We obtained a diverse set of 706 soybean cultivars released during 60 years of breeding (1956–2012) from the Chinese Academy of Agricultural Sciences (CAAS), as well as 536 soybean breeding lines from the National Soybean Regional Trials in China (2013–2017). These soybeans were used to investigate the historical allelic variation in soybean salt tolerance gene *GmSALT3* and breeding lines were used to evaluate the selection efficiency of molecular markers developed from *GmSALT3*.

Salt Tolerance Evaluation

We tested a diverse set of 536 soybean breeding lines for salt tolerance in a rain shelter that excluded rainfall but allowed plants to grow under ambient light and temperature (Institute of Crop Sciences, CAAS). Experiments were performed according to previous reports (Jiang et al., 2013; Liu et al., 2016). Ten seeds of each line were sowed in a pot filled with vermiculite and were thinned to six plants after five days. Twenty-four pots were placed in one tray and 11 days after sowing (DAS), when the unifoliate leaves were fully expanded, 2 L of salt solution (200 mM NaCl) was added to each tray. The same volume of NaCl solution was added to each tray at 13 and 15 DAS, respectively. Ten days after the last addition of the salt solution, leaf chlorosis was observed and scored accordingly: 1 = healthy green leaves, no damage observed; 2 = slight chlorosis, light yellowish color observed in true leaves; 3 = moderate chlorosis, chlorosis observed in trifoliate leaves; 4 = severe chlorosis, more than 75% of the leaf area showed chlorosis; 5 = dead, plants were completely withered. The experiment was performed with 3 pots of each genotype.

Analysis of Genomic Sequence Variation

We used the genomic dataset of 279 Chinese soybean landraces (**Supplementary Table 1**; Li et al., 2020) to determine genomic sequence variation in *GmSALT3*. SNPs with a minor allele frequency $\leq$0.01, or with missing data $>$0.1, and indels with maximum length $>$10 bp were discarded. Annotation was carried out based on the soybean reference genome Wm82.a2.v1[1] and transcript sequence information (Guan et al., 2014b; Qi et al., 2014; Lee et al., 2018). SNPs and indels within the 5′UTR and the genomic region of *GmSALT3* were used for haplotype investigation.

Development of Functional Markers

We developed functional markers to distinguish the haplotypes of *GmSALT3*. We created three indel markers based on the ∼150 bp

[1]https://phytozome.jgi.doe.gov/

insertion in the promoter region, the 4 bp deletion and the 3.78-kb insertion in the coding sequence of *GmSALT3*. Additionally, we developed cleaved amplified polymorphic sequences (CAPS) markers, H3-*Mbo*II and H4-*Nla*III, for haplotypes H3 and H4 based on a GC > TG substitution at exon 4 and a splice site AG > AT substitution at the end of intron 2. The primers used to amplify and distinguish haplotypes of *GmSALT3* are shown in **Table 1**.

Genomic DNA Isolation and Genotyping of Soybeans

Genomic DNA was isolated from leaves of each accession by using a Genomic DNA Purification Kit (Thermo Fisher Scientific, Lithuania), and 100 ng DNA was used for PCR amplification with a T100 thermal cycler (Bio-Rad). We performed the PCR experiment using 20 μl reaction mixtures containing 100 ng genomic DNA, 2 μl 10× EasyTaq Buffer (with Mg^{2+}), 1.5 μl 2.5 mM dNTPs, 2 μl each of 2 μM primer stock, and 1 U EasyTaq DNA Polymerase (TransGen Biotech, Beijing, China) under the following thermal cycler conditions: 95°C for 5 min, then 35 cycles at 95°C for 30 s, 55°C for 30 s, and 72°C for 50 s, and followed by a final extension of 5 min at 72°C. Products of PCR of H3-*Mbo*II and H4-*Nla*III were digested with, respectively, 1 U of restriction enzymes *Mbo*II and *Nla*III in 1× reaction buffer for 1 h at 37°C. PCR products and enzyme-digested products were separated on a 1.5% agarose gel or 6% polyacrylamide gel.

RESULTS

Haplotypes of *GmSALT3* and Their Geographical Distribution

The multi-allelic salt tolerance gene *GmSALT3*, was previously isolated from Chinese soybean cultivar Tiefeng 8, and five distinct haplotypes (H1–H5) have previously been described in Chinese soybean landraces (Guan et al., 2014b). H1 was found in the salt tolerant soybean Tiefeng 8. H2 was found in salt sensitive soybean cultivar 85-140, of which a 3.78-kb retro transposon insertion resulted in a truncated GmSALT3 protein (reference Williams 82: W82). Compared with the H1 haplotype, H3–H5 had a ∼150 bp insertion at position-147 of the promoter region and a TCGA insertion at position-103 in common with each other (Guan et al., 2014b). We analyzed the haplotypes of *GmSALT3* from a panel of 279 Chinese landraces (**Supplementary Table 1**) and identified 57 polymorphic sites, including 53 SNPs and four indels (**Supplementary Table 2**). These SNPs and indels formed six haplotypes (H1–H4, H5-1, and H5-2) (**Figure 1** and **Supplementary Table 2**). H1–H4 were the same as reported previously (Guan et al., 2014b). Both H5-1 and H5-2 had a 4 bp deletion in exon 2 which resulted in frameshift as earlier reported in H5, while H5-1 differed from H5-2 by 10 SNPs and one indel in intron 1 (**Supplementary Table 2**). Given that H5-1 and H5-2 had identical 714-nt cDNA sequence, we ascribed them as haplotype H5.

We found that H1, H4, and H5 were the highest frequency haplotypes in Chinese soybean landraces (37.3, 21.8, and 29.7%, respectively). When examining the three main eco-regions, the northern region (NR) which includes the area above 40°N, the Huanghuai region (HHR) in middle China (30°N to 40°N), and the southern region (SR) which ranges from Hainan Island (19°N) to Shanghai (31°N) (Li et al., 2008), H1 was the most abundant haplotype in both the NR and the HHR, while H4 and H5 together, were the most abundant haplotypes in the SR of China (**Figure 1**, **Supplementary Figure 1**, and **Supplementary Table 1**).

Development of Haplotype-Specific Markers

To identify different *GmSALT3* alleles in a simple and effective way, we developed three indel markers and two CAPS markers. Indel marker Pro-Ins was designed to distinguish H1 and H2 haplotypes from H3 to H5; the H1 and H2 haplotypes amplified a 623 bp fragment, while H3–H5 amplified an approximate 775 bp fragment due to ∼150 bp insertion in the promoter region (**Figure 2**). Indel marker H2-Ins was designed to distinguish H2 from the other haplotypes using a common forward primer (H2-Ins-F) and two haplotype-specific reverse primers that were designed according to a 3.78-kb insertion (H2-Ins-R/H2) or exon 3 (H2-Ins-R/H1), and amplified a 565 bp fragment in H2, and 364 bp fragment in the H1 and H3–H5 haplotypes (**Figure 2B**). The marker H5-Del amplified a 210 bp sequence from H5 and a 214 bp sequence from the other four haplotypes. The CAPS marker H3-*Mbo*II amplified a 296 bp fragment from all haplotypes, and the fragment from the H3 haplotype could be digested into three fragments (125, 119, and 52 bp) by *Mbo*II, whereas the products of the other haplotypes were cleaved into two fragments (244 and 52 bp). To distinguish H4 from the other four haplotypes, CAPS marker H4-*Nla*III was developed, and an amplicon of H4 was digested into four fragments (111, 87, 42, and 22 bp) by *Nla*III, the products of H1 and H2 were digested into two fragments (133 and 129 bp) and those of H3 and H5 into three fragments (133, 87, and 42 bp) (**Figure 2B**).

Changes of *GmSALT3* Haplotype Frequencies During Six Decades of Chinese Soybean Breeding

We used the markers developed in this study to explore the genetic diversity of *GmSALT3* in a total of 706 modern Chinese soybean cultivars (released between 1956 and 2012). All five haplotypes were found in these cultivars (n = 323 H1, 259 H2, 5 H3, 6 H4, and 111 H5). Different haplotypes were not equally distributed throughout the three eco-regions in China. The frequencies of cultivar haplotypes in the NR were as follows: 39.6% H1, 49.1% H2, 0.2% H3, 1.0% H4, and 8.7% H5. In the HHR, 63.1% of the genotypes were H1, 18.7% were H2, 0.9% were H3, 0.4% were H4, and 16.8% were H5. Intriguingly, H5 was the most frequent haplotype (68.8%) in the SR, followed by H1 (25.0%), H3 (2.1%), and H4 (4.2%); H2 was the least frequent haplotype, found in only one cultivar (**Figure 3A**).

Additionally, we analyzed frequency changes of the haplotypes in the NR and HHR since the 1950s, at 10-year release intervals, over a 60-year time span [SR was not included due to a smaller

TABLE 1 | Haplotype-specific markers for *GmSALT3*.

Marker name		Primer sequence	Marker type	Size (bp)
Pro-Ins	F	GGGTTGTGCCTAAATAGCA	Indel	623/775/777
	R	AAGGAAGAGCGTGGTTCA		
H2-Ins	F	GCGGGAGTAATGTTATCGG	Indel	
	R-H1	CGATTAGCTCCACCAACCCT		364
	R-H2	GTCGTATCTTGGGAGAGGAG		565
H3-*Mbo*II	F	TATGGTGGCTAAGCAGGTG	CAPS	(125 + 119 + 52)/(244 + 52)
	R	CAGTGAGTTCGGTAAGTTGC		
H4-*Nla*III	F	AAAGCGCATAAGTTATAACACAAAAT	CAPS	(133 + 129)/(111 + 87 + 42 + 22)/(133 + 87 + 42)
	R	GAATGTAACCCTATCATGTCTGTCA		
H5-Del	F	CTGTCCATCACGGCTTTCC	Indel	210/214
	R	CTATAGTAGGTCCACCTGAGAA		

Position from start codon		-107	-100	-20	39	63	683–686	695	774	1060	1234	1284	1285	1349	1548	1772	1816	1903	2124	2170	2355	2368	No. of accessions		
Haplotype																							NR	HHR	SR
H1		C	A	C	A	C	GTGCT	G	A	A	C	G	C	A	G	T	G	A	T	T	T	G	32	55	17
H2		C	C	G	A	C	GTGCT	G	A	A	C	G	C	A	G	T	G	T	T	T	T	G	12	1	1
H3		CTCGA	C	G	T	T	GTGCT	G	T	T	C	T	G	A	C	G	A	A	A	C	G	G	2	6	9
H4		CTCGA	C	G	T	T	GTGCT	T	T	T	C	G	C	A	C	G	A	A	A	C	G	A	10	9	42
H5	-1	CTCGA	C	G	T	T	G	G	T	T	C	G	C	G	G	T	G	A	T	T	G	G	1	15	53
	-2	CTCGA	C	G	T	T	G	G	T	T	G	G	C	G	G	T	G	A	T	T	G	G	0	3	11
Amino acid position in H1					13	21	228-229	232	258	354	412	428	429	450	516	591	606	635	708	724	785	790			
Amino acid in H1					E	P			Q	S	L	L	Q	E	M	L	V	I	T	L	F	E			
Amino acid change					D	P	Framshift	splicing	H	C		F	E		I	W	I		T	L	L	K			

FIGURE 1 | Haplotypes of *GmSALT3*. Red color represents the same nucleotides as that of H1 type, blue color represents variation in the 5′UTR and coding regions different from that of H1 type. See **Supplementary Table 2** for the whole genomic variation of *GmSALT3*.

sample of cultivars (48)]. Eight cultivars released in 2011 and 2012 were merged with those of the 2000s. In the NR, H1 and H2 frequencies exhibited fluctuating patterns from the 1950s to 1970s (**Figure 3B**). H1 decreased from 60.5% to 39.3% and H2 increased from 23.7 to 48.8% during the 1970s to 2000s (**Figure 3B**). H3 and H4 occurred at low frequencies, where they were, respectively, present in only 1 and 6 cultivars.

Five haplotypes were observed in the HHR; where H1 increased in frequency from 50.0 to 77.8% (1960s to 1970s), then gradually decreased to 62.5% in the 21st century. H2 increased in frequency to 25.8% and H5 decreased to 12.1% (**Figure 3C**). Such variation in allelic frequency over the past 60 years indicates that H1 and H2 were selected for in the NR, and H1 was favored in the HHR (**Figure 4**).

Validation of Functional Markers in Modern Soybean Breeding Lines

To evaluate the genotype-phenotype relationship among haplotypes in modern soybean lines, a validation panel of 536 breeding lines (National Soybean Regional Trials, 2013–2017) was tested for salt tolerance and genotyped using the five PCR-based functional markers determined in this study. Four haplotypes were observed in the breeding lines (**Figure 4**), with 35.1% of soybean lines possessing H1, 41.0% possessing H2, and 23.7% possessing H5. H3 was only observed in one line from the SR. In the NR, H2 was the predominant haplotype (73.3%); H5 was only observed in 7.6% of lines. Whereas in the HHR, H1 and H5 were expressed in 67.0% and 19.3% of lines, respectively. In the SR, H5 and H1 were the two highest frequency haplotypes (56.7% and 41.3%). Modern breeding line frequencies of H2 in the NR, and H1 in the SR, were much higher compared with soybean cultivars released from 1956 to 2012 (**Figure 4**). These results suggest positive selection for favored haplotypes in different eco-regions during modern cultivation.

Of the 536 modern breeding lines, 186 lines were salt tolerant (with chlorosis scores of 1 or 2) and 350 lines were salt sensitive (with chlorosis scores 3–5) (**Figure 5**). The correlation of genotype-phenotype showed that H2 (220 lines), H5 (127 lines) and one breeding line which possessed H3, were all salt sensitive, while 98.9% of the lines possessing H1 were salt tolerant (**Figure 5B**). Two breeding lines with the H1 haplotype showed moderate leaf scorch (chlorosis score 3), suggesting that different loci other than *GmSALT3* might affect salt tolerance in these accessions.

Pedigree Tracing of *GmSALT3* in Soybean Varieties

To examine potential use of the *GmSALT3* PCR-based markers in a soybean breeding program, we traced the pedigree of four soybean cultivars. Zhonghuang 13 is a salt tolerant cultivar that exemplifies widely adaptable soybean cultivar in China, it was registered in nine provinces ranged from 29°N to 42°N, and was also the first Chinese soybean cultivar used for *de novo genome* assembly (Shen et al., 2018). Wenfeng 7 is a historical soybean cultivar that has previously been used for salt tolerance analysis (Ren et al., 2012; Pi et al., 2016), and shares Jüxuan 23 as the common ancestral parent with Zhonghuang 13 (**Figure 6**). Zhonghuang 30 is an abiotic stress resistant cultivar that is

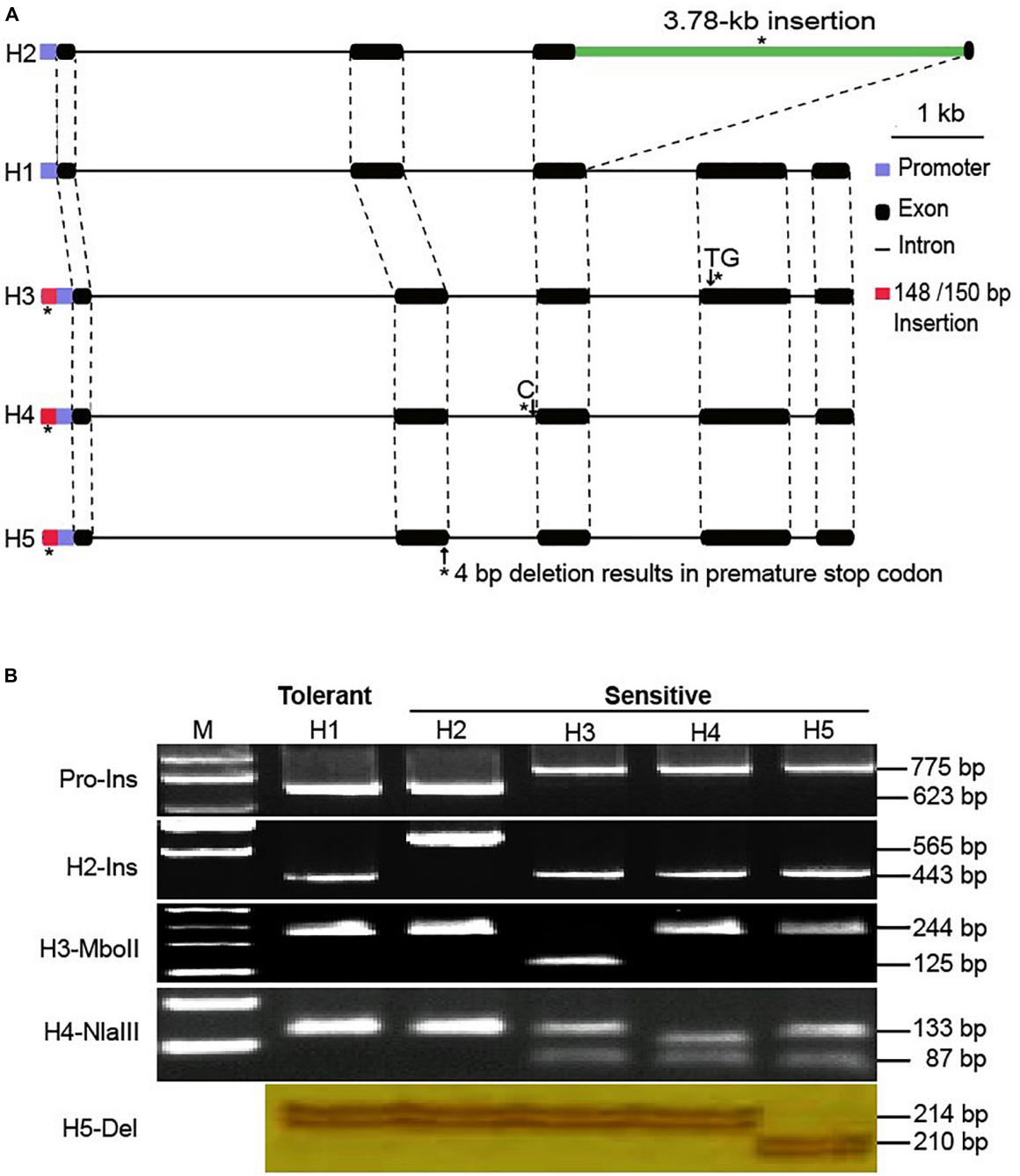


FIGURE 2 | Variation of *GmSALT3* and haplotypes (H1–H5) revealed by indel and CAPS markers. Schematic diagrams show the structure of five *GmSALT3* haplotypes. Asterisks indicate variations that were used for PCR-based marker development **(A)**; Markers that distinguish five haplotypes in soybean varieties **(B)**.

widely used in the NR (He et al., 2016), and Zhonghuang 39 is a salt sensitive sister line of Zhonghuang 30. These cultivars and 19 related ancestral lines in the pedigree were assayed with the five molecular markers. The phenotypic salt response and corresponding *GmSALT3* haplotypes (H1–H5) of the pedigree lines are marked in **Figure 6**. Zhonghuang 13, Zhonghuang 30, Wenfeng 7 were H1 type salt tolerant cultivars. The salt tolerance gene in Wenfeng 7 could be traced back to Jüxuan 23. The tolerance gene in Zhonghuang 13 could be traced back to Zhengzhou 135 or 58-161. The *GmSALT3* allele (haplotype H2) in Zhonghuang 39, the sensitive sister line of Zhonghuang 30, was traced back to W82, while the tolerant allele in Zhonghuang 30 was inherited from Zhengzhou 135 (**Figure 6**). These results indicate that the H1 haplotype is present in all the salt tolerant

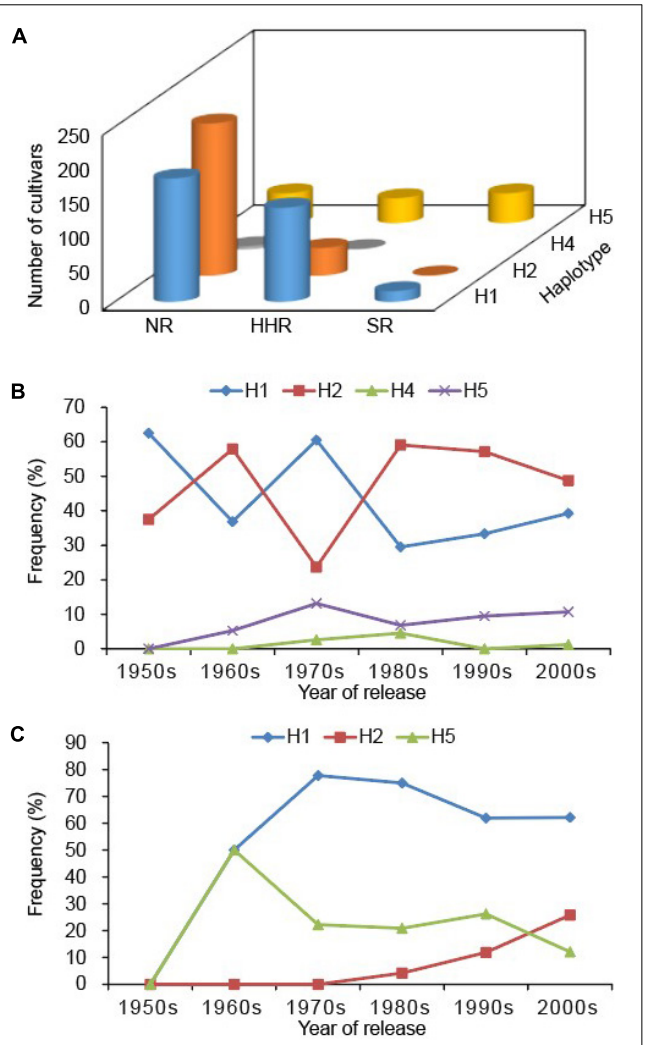


FIGURE 3 | *GmSALT3* allele characterization in Chinese soybean varieties. Distribution of haplotypes among varieties from different ecoregions **(A)**; Frequency changes of haplotypes in the northern region of China (NR) **(B)**; and Huanghuai region of China (HHR) **(C)** over 60 years of soybean breeding. The distribution of H3 is not shown because only five cultivars expressed this haplotype.

ancestors in the pedigree, and that sensitive ancestors expressed haplotypes H2 or H5.

DISCUSSION

Genomic Diversity of *GmSALT3* in Chinese Soybean Landraces

Gene cloning related to agronomic traits has facilitated the development allele-specific functional DNA markers, which in turn, have been used in soybean research to validate genotype-phenotype relationships and facilitate breeding efforts (Shin and Lee, 2012; Tsubokura et al., 2013, 2014; Jia et al., 2014; Zhai et al., 2014; Liu et al., 2015). *GmSALT3* was identified as gene that controls Na^+ and Cl^- uptake in soybean shoots and improves yield in saline conditions without negative consequences on soybean yield under non-saline field conditions (Liu et al., 2016). We previously found at least nine haplotypes of *GmSALT3* in 172 Chinese soybean landraces and 57 wild soybeans by analyzing promoter and coding sequences of the gene (Guan et al., 2014b). Additionally, 3 and 40 different haplotypes of *GmSALT3* were observed in 106 and 216 diverse soybean accessions, respectively (Patil et al., 2016; Lee et al., 2018). In the present study, genetic diversity in *GmSALT3* was analyzed using whole genome resequencing data from 279 representative Chinese soybean landraces (Li et al., 2020). Genomic variants were classified into six haplotypes (H1–H4, H5-1, and H5-2). Haplotypes H1–H4 were the same as previously found. The intron variations between Haplotype H5-2 and H5-1 (**Supplementary Table 2**), indicate that the common frameshift mutation (TGCT deletion) in exon 2 of H5-1 and H5-2 arose independently in south of China from different lineages (**Figure 1** and **Supplementary Figure 2**). H5-1 and H5-2 were assigned as H5 haplotype due to the same frameshift mutation and their identical amino acid sequences (**Figure 1**).

In the 40 *GmSALT3* haplotypes observed by Lee et al. (2018), Ha, HTn, Hd-2, and Hd-3, which correspond to H1, H2, H5-1, and H5-2 in present study, were the four majority haplotypes in cultivated soybean (Lee et al., 2018). Haplotypes Hst-7, Hss-11, Hss-12, and Hd-1 were shown in only one Chinese soybean accession in their report, indicating these are minor alleles in Chinese soybeans. The H3 and H4 haplotypes identified in the present study were not found in any of the three structural variations listed in Patil's report (Patil et al., 2016), or in the 40 haplotypes of Lee's study (Lee et al., 2018). This lack of concordance is not surprising considering the greater number of Chinese soybean accessions used in the present experiment.

Functional Markers of *GmSALT3* and Their Potential Use in Soybean Breeding

Using the functional markers we developed in the present study (**Figure 2** and **Table 1**), we successfully validated the predictive accuracy of the markers across 536 modern breeding lines. Prediction accuracy was 98.9% for salt-tolerant lines and 100% for salt-sensitive lines. H1 and H2 shared similar promoter sequences, while H3–H5 had ∼150 bp and 4 bp insertion in the promoter region (**Figure 2**). However, previous research found no insertion in the *GmSALT3* promoter region (Patil et al., 2016). This difference might due to (1) the ∼150 bp insertion sequence having low similarity with the W82 reference genome, thus it could not be perfectly assembled to the reference genome, or (2) the limitations of next-generation sequencing data in finding long fragment insertions/deletions using paired-end reads of 45 bp or 75 bp (Lam et al., 2010). The 4 bp insertion in the promoter of *GmSALT3* was observed in our 279 resequencing dataset, while the ∼150 bp insertion was not (**Figure 1**). Reduced gene expression of the salt sensitive haplotypes has been reported by several researchers (Guan et al., 2014b; Do et al., 2016; Lee et al., 2018). This common variation in the promoter region of salt

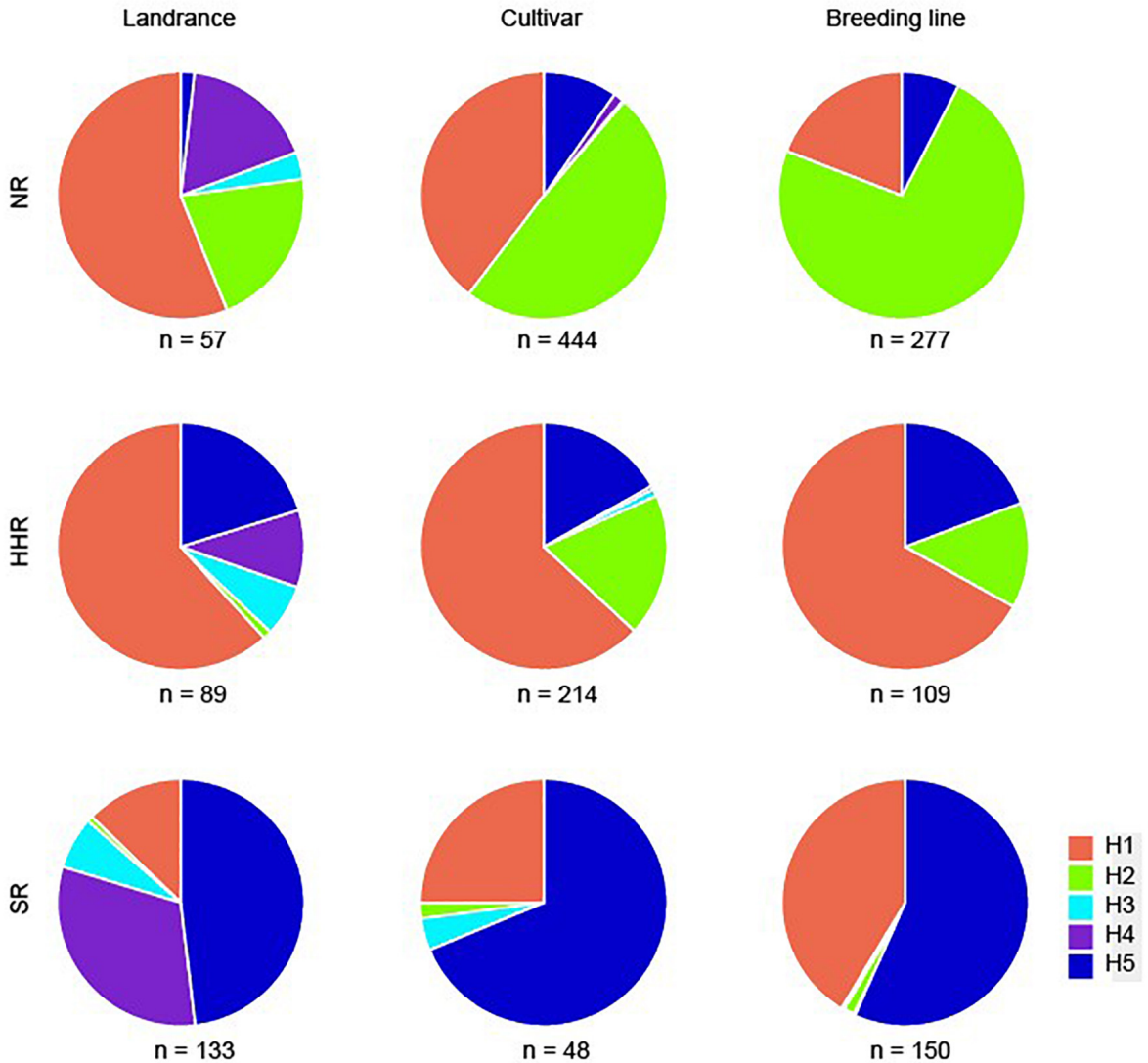


FIGURE 4 | Distribution of *GmSALT3* haplotypes in Chinese soybean landraces, cultivars, and breeding lines from the three ecoregions. NR, northern region of China; HHR, Huanghuai region of China; SR, southern region of China.

sensitive haplotypes raises the possibility that these insertions might affect the expression of *GmSALT3*.

Patil et al. (2016) conducted six KASPar assays with three structural variants in *GmSALT3* and achieved precise identification of tolerant and sensitive genotypes with over 91% accuracy. Additionally, Lee et al. (2018) developed five molecular markers for the seven haplotype groups in 216 soybean accessions, including three codominant and two dominant markers, and achieved 98.8% predictive accuracy for 173 soybean accessions. In this study, we successfully predicted salt tolerant vs. sensitive accessions using only two of the five codominant markers we developed, i.e., using Pro-Ins to distinguish H1 and H2 from H3 to H5, and using H2-Ins to distinguished H1 from H2. Prediction efficacy was 98.9% for salt tolerant and 100% for salt sensitive accessions.

Interestingly, two accessions (SW10710 and CW10948) with salt-sensitive genotypes showed high or moderate salt tolerance (Lee et al., 2018). We also found that two of our H1 type soybean lines showed salt sensitivity, indicating that other genes may impact salt tolerance in these accessions, and these genotypes are useful material to identify new sources of salt tolerance. Candidate loci that might influence this are hinted at through past literature. For instance, in the population derived from Kefeng No. 1 and Nannong1138-2, a major QTL on linkage group G (Chr. 18) was identified in both green house and field conditions which correlates to altered salt sensitivity (Chen et al., 2008). Furthermore, a novel locus on Chr. 13 related to leaf sodium content was observed in soybean germplasm Fiskeby III (Do et al., 2018). These results suggest that other loci, in addition to the major gene *GmSALT3* affect salt tolerance in soybean.

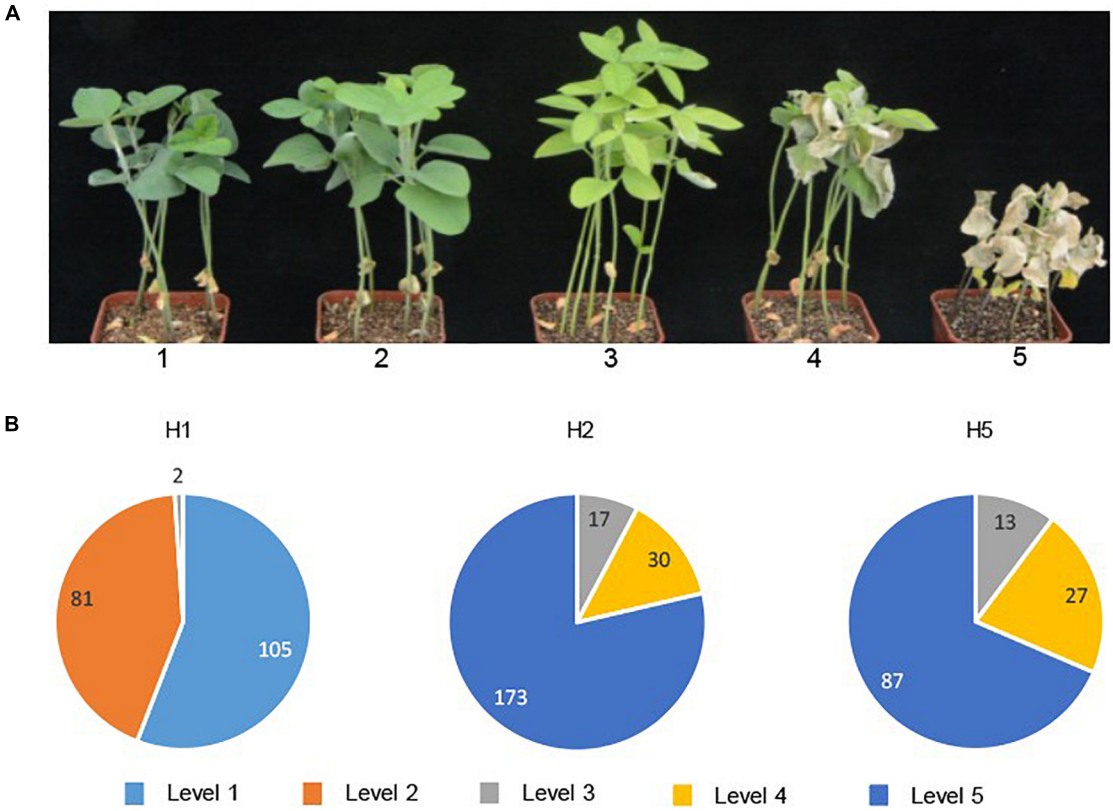


FIGURE 5 | Salt tolerance evaluation of 536 modern soybean breeding lines. Chlorosis scores were used for salt tolerance rating after 15 days of NaCl treatment (plants with score ratings 1 or 2 were considered salt tolerant; 3–5 were salt sensitive) **(A)**; In each haplotype, the number of lines belonging to each chlorosis score for salt tolerance ratings (level 1 to level 5) are shown **(B)**.

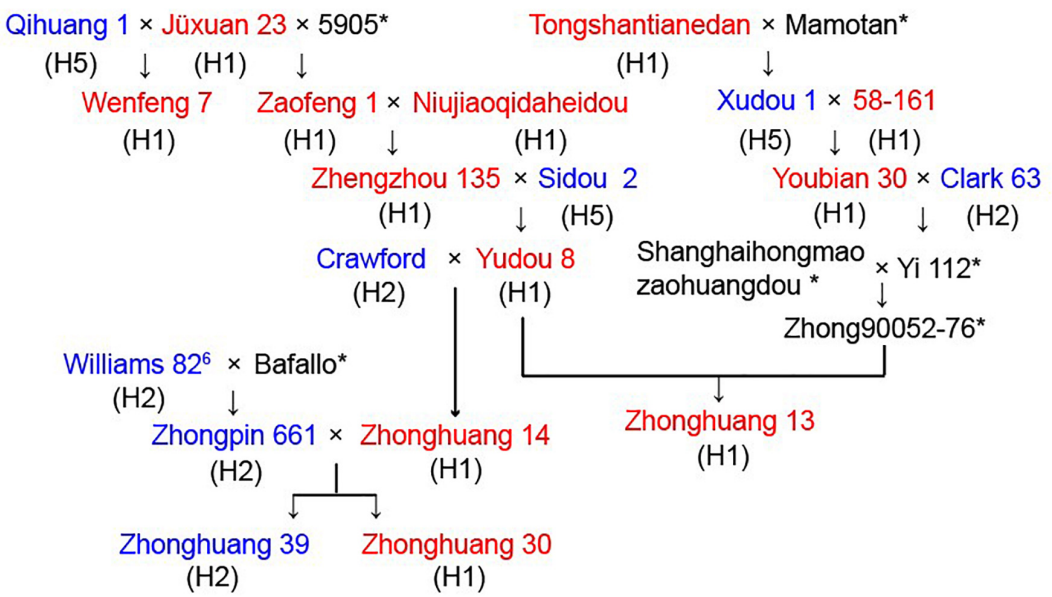


FIGURE 6 | Pedigree of four soybean cultivars for the salt tolerance gene *GmSALT3*. Salt tolerance of each cultivar is indicated in red (tolerant), blue (sensitive), and haplotypes (H1–H5) are shown in parentheses under the name of cultivar. * indicate seeds of these cultivars were unavailable.

Future work should focus on defining the genetic relationship of these loci and *GmSALT3*; how might alleles be interacting to impact salt tolerance in particular soybean germplasms?

Soybean [*Glycine max* (L.) Merr.] is an essential crop worldwide due to its high protein and oil content; it provides valuable source of nutrition for human consumption and main protein source in for livestock feed (Bellaloui et al., 2014). As one of the major crops grown in China, soybean is widely used for edible oil and soy-based foods such as tofu, soymilk, sauces, and many other products. However, from 2004 to 2016, the land area in Chinese soy production decreased by 25%, much of that land being converted to cropland for maize production (Ren et al., 2021). Reclamation of slightly saline soil, therefore, could be a potential way to increase productivity of soybean yield.

In China, modern soybean breeding started as early as 1923 and regional trials of soybean production began in the 1950s. We found that in soybean landraces, H1 (salt tolerant haplotype) was at highest frequencies in both the NR and HHR, whereas H4 and H5 (salt-sensitive) were highest in the SR. In soybean cultivars, the highest frequencies of H1 were observed in the HHR, and salt-sensitive haplotypes H2 and H5 most frequently occurred in the NR and SR, respectively (**Figure 3**). In modern breeding lines, the frequency of H2 in the NR increased to 73.3%, and H1 in the SR increased to 41.3%, while H3 and H4 were nearly eliminated from modern breeding lines. The high frequency of the H1 haplotype observed in the HHR is consistent with the distribution of saline soils along the eastern coast of China and saline conditions in the Yellow River Delta (Meng et al., 2016). No significant difference of yield-related traits of tolerant (T) and sensitive (S) near isogenic lines NIL-T and NIL-S has been observed under non-saline field conditions, indicating that the H1 haplotype has no yield penalty (Liu et al., 2016). Additionally, selection of the salt sensitive haplotype H2 in the NR during modern times might be attributed to climate change, water reduction and subsequent decreasing soil salinization in the region (Guo et al., 2016). Likewise, the slight frequency increase of H1 haplotype in the SR (**Figure 4**) may be related to increasing soil salinization along the Yangtze River Basin due to vegetable production practices (Zhang et al., 2021). That being said, soybean breeding efforts are geared toward many factors including high seed yield, high oil content, and high protein content. Therefore, we cannot rule out the possibility that pleiotropy of *GmSALT3* or linkage drag affects the selection of haplotypes in different ecoregions.

Previously, simple sequence repeat (SSR) markers were used to trace the origin of the salt tolerance gene; however, the multi-allelic nature of SSR markers decreased prediction accuracy of salt-tolerant genotypes (Lee et al., 2004). We were able, however, to clearly distinguish the possible sources of alleles in a soybean pedigree by using functional markers (**Figure 6**). In addition, previous field studies suggested that the parent strains of Wenfeng 7, Jüxuan 23, and Qihuang 1 exhibited high or moderate salt tolerance (Shao, 1988). In contrast, our research indicated that the parent strain Qihuang 1 was an H5 type salt-sensitive cultivar, thus indicating the predictive value of functional markers and their potential use in soybean breeding.

The ability to select salt-tolerant soybean germplasm via accurate and economic genotyping methods is essential for modern soybean production. Here, we developed PCR-based functional markers of *GmSALT3*, and demonstrated their high efficiency in the selection of salt tolerant or sensitive breeding lines. By using only Pro-Ins and H2-Ins, for which amplicons can be separated on standard agarose gels, salt-tolerant genotypes (H1) are readily distinguished from salt-sensitive genotypes (H2–H5). Most soybean breeding laboratories should be able to employ this method of identification and selection. Our exploration of variation in *GmSALT3* throughout known Chinese soybean breeds, in addition, provided valuable insight into the long-term effect of phenotypic selection on the distribution of *GmSALT3* haplotypes. Further, it provided a basis for the utilization of *GmSALT3* allele ecoregion-specific breeding; which, under changing global conditions, remains key to economic and food security.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

RG, RC, and LQ designed the experiments. RG and LQ directed the project. RG performed the design of molecular markers. LY and XL performed the phenotype screen and data analysis. XL and ML performed the genotype test of the soybeans. RG, LY, MG, and LQ wrote the manuscript. All authors commented on the manuscript.

FUNDING

This research was funded by the National Natural Science Foundation of China (31830066), Central Public-Interest Scientific Institution Basal Research Fund (S2021ZD02), and Fundamental Research Funds for Central Non-Profit of Institute of Crop Sciences, Chinese Academy of Agricultural Sciences (S2018QY03).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.794241/full#supplementary-material

Supplementary Figure 1 | Geographical distribution of five haplotypes (H1–H5) of *GmSALT3* in Chinese soybean landraces. NR, the northern region of China; HHR, the Huanghuai region of China; SR the southern region of China.

Supplementary Figure 2 | Distribution of Chinese soybean landraces possessing haplotype H5-1 and H5-2. Red dots, H5-1; Black dots, H5-2.

REFERENCES

Abel, G. (1969). Inheritance of the capacity for chloride inclusion and chloride exclusion by soybeans. *Crop Sci.* 9, 697–698. doi: 10.2135/cropsci1969.0011183x000900060006x

Abel, G., and MacKenzie, A. J. (1964). Salt tolerance of soybean varieties (*Glycine max* L. Merrill) during germination and later growth. *Crop Sci.* 4, 157–161. doi: 10.2135/cropsci1964.0011183x000400020010x

An, P., Inanaga, S., Cohen, Y., Kafkafi, U., and Sugimoto, Y. (2002). Salt tolerance in two soybean cultivars. *J. Plant Nutr.* 25, 407–423.

Bellaloui, N., Mengistu, A., Zobiole, L., Abbas, H. K., and Kassem, M. A. (2014). Soybean seed phenolics, sugars, and minerals are altered by charcoal rot infection in mg iii soybean cultivars. *Food Nutr. Sci.* 5, 1843–1859. doi: 10.4236/fns.2014.519198

Chen, H., Cui, S., Fu, S., Gai, J., and Yu, D. (2008). Identification of quantitative trait loci associated with salt tolerance during seedling growth in soybean (*Glycine max* L.). *Aust. J. Agr. Res.* 59, 1086–1091. doi: 10.1071/ar08104

Do, T. D., Chen, H., Hien, V. T., Hamwieh, A., Yamada, T., Sato, T., et al. (2016). Ncl synchronously regulates Na+, K+, and Cl- in soybean and greatly increases the grain yield in saline field conditions. *Sci. Rep.* 6:19147. doi: 10.1038/srep19147

Do, T. D., Vuong, T. D., Dunn, D., Smothers, S., Patil, G., Yungbluth, D. C., et al. (2018). Mapping and confirmation of loci for salt tolerance in a novel soybean germplasm, Fiskeby III. *Theor. Appl. Genet.* 131, 513–524. doi: 10.1007/s00122-017-3015-0

Durand, M., and Lacan, D. (1994). Sodium partitioning with the shoot of soybean. *Physiol. Plant.* 91, 65–71. doi: 10.1034/j.1399-3054.1994.910110.x

El-Samad, H. A., and Shaddad, M. (1997). Salt tolerance of soybean cultivars. *Bio Plantarum* 39, 263–269.

Fehr, W., Caviness, C., Burmood, D., and Pennington, J. (1971). Stage of development descriptions for soybeans, *Glycine max* (L.) Merrill. *Crop sci.* 11, 929–931. doi: 10.2135/cropsci1971.0011183x001100060051x

Guan, R., Chen, J., Jiang, J., Liu, G., Liu, Y., Tian, L., et al. (2014a). Mapping and validation of a dominant salt tolerance gene in the cultivated soybean (*Glycine max*) variety Tiefeng 8. *Crop J.* 2, 358–365. doi: 10.1016/j.cj.2014.09.001

Guan, R., Qu, Y., Guo, Y., Yu, L., Liu, Y., Jiang, J., et al. (2014b). Salinity tolerance in soybean is modulated by natural variation in GmSALT3. *Plant J.* 80, 937–950. doi: 10.1111/tpj.12695

Guo, S., Ruaan, B., Guan, X., Wang, S., and Li, Y. (2016). Analysis on spatial-temporal evolution of soil salinity and its driving factors in Hetao irrigation district during recent 30 years. *China Rural Water Hydropower* 9, 937–950.

Ha, B. K., Vuong, T. D., Velusamy, V., Nguyen, H. T., Shannon, J. G., and Lee, J. D. (2013). Genetic mapping of quantitative trait loci conditioning salt tolerance in wild soybean (*Glycine soja*) PI 483463. *Euphytica* 193, 79–88. doi: 10.1007/s10681-013-0944-9

Hamwieh, A., and Xu, D. (2008). Conserved salt tolerance quantitative trait locus (QTL) in wild and cultivated soybeans. *Breed. Sci.* 58, 355–359. doi: 10.1270/jsbbs.58.355

Hamwieh, A., Tuyen, D., Cong, H., Benitez, E., Takahashi, R., and Xu, D. (2011). Identification and validation of a major QTL for salt tolerance in soybean. *Euphytica* 179, 451–459. doi: 10.1007/s10681-011-0347-8

He, J., Du, Y. L., Wang, T., Turner, N. C., Xi, Y., and Li, F. M. (2016). Old and new cultivars of soya bean (*Glycine max* L.) subjected to soil drying differ in abscisic acid accumulation, water relations characteristics and yield. *J. Agron. Crop Sci.* 202, 372–383.

Jia, H., Jiang, B., Wu, C., Lu, W., Hou, W., Sun, S., et al. (2014). Maturity group classification and maturity locus genotyping of early-maturing soybean varieties from high-latitude cold regions. *PLoS One* 9:e94139. doi: 10.1371/journal.pone.0094139

Jiang, J., Guan, R., Guo, Y., Chang, R., and Qiu, L. (2013). Simple evaluation method of tolerance to salt at seedling stage in soybean. *Acta Agron. Sinica* 39, 1248–1256. doi: 10.3724/sp.j.1006.2013.01248

Lam, H.-M., Xu, X., Liu, X., Chen, W., Yang, G., Wong, F.-L., et al. (2010). Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. *Nat. Genet.* 42, 1053–1059. doi: 10.1038/ng.715

Läuchli, A., and Wieneke, J. (1979). Studies on growth and distribution of Na+, K+ and Cl− in soybean varieties differing in salt tolerance. *Z. Pflanzenemaehr. Bodenkd* 142, 3–13. doi: 10.1002/jpln.19791420103

Lee, G., Boerma, H., Villagarcia, M., Zhou, X., Carter, T., Li, Z., et al. (2004). A major QTL conditioning salt tolerance in S-100 soybean and descendent cultivars. *Theor. Appl. Genet.* 109, 1610–1619. doi: 10.1007/s00122-004-1783-9

Lee, S., Kim, J.-H., Sundaramoorthy, J., Park, G. T., Lee, J.-D., Kim, J. H., et al. (2018). Identification of GmSALT3 haplotypes and development of molecular markers based on their diversity associated with salt tolerance in soybean. *Mol. Breed.* 38:86.

Li, C., Li, Y. H., Li, Y., Lu, H., Hong, H., Tian, Y., et al. (2020). A domestication-associated gene GmPRR3b regulates the circadian clock and flowering time in soybean. *Mol. Plant* 13, 745–759. doi: 10.1016/j.molp.2020.01.014

Li, Y., Guan, R., Liu, Z., Ma, Y., Wang, L., Li, L., et al. (2008). Genetic structure and diversity of cultivated soybean (*Glycine max* (L.) Merr.) landraces in China. *Theor. Appl. Genet.* 117, 857–871. doi: 10.1007/s00122-008-0825-0

Liu, G., Zhao, L., Averitt, B. J., Liu, Y., Zhang, B., Chang, R., et al. (2015). Geographical distribution of GmTfl1 alleles in Chinese soybean varieties. *Crop J.* 3, 371–378. doi: 10.1016/j.cj.2015.05.004

Liu, Y., Yu, L., Qu, Y., Chen, J., Liu, X., Hong, H., et al. (2016). GmSALT3, which confers improved soybean salt tolerance in the field, increases leaf Cl-exclusion prior to Na+ exclusion but does not improve early vigor under salinity. *Front. Plant Sci.* 7:1485. doi: 10.3389/fpls.2016.01485

Meng, L., Zhou, S., Zhang, H., and Bi, X. (2016). Estimating soil salinity in different landscapes of the yellow rive delta through lLandsat OLI/TIRS and ETM+ data. *J. Coast. Conserv.* 20, 271–279. doi: 10.1007/s11852-016-0437-9

Parida, A. K., and Das, A. B. (2005). Salt tolerance and salinity effects on plants: a review. *Ecotoxicol. Environ. Saf.* 60, 324–349. doi: 10.1016/j.ecoenv.2004.06.010

Patil, G., Do, T., Vuong, T. D., Valliyodan, B., Lee, J. D., Chaudhary, J., et al. (2016). Genomic-assisted haplotype analysis and the development of high-throughput SNP markers for salinity tolerance in soybean. *Sci. Rep.* 6:19199. doi: 10.1038/srep19199

Pi, E., Qu, L., Hu, J., Huang, Y., and Du, L. (2016). Mechanisms of soybean roots' tolerances to salinity revealed by proteomic and phosphoproteomic comparisons between two cultivars. *Mol. Cell Proteom.* 15, 266–288. doi: 10.1074/mcp.M115.051961

Qi, X., Li, M. W., Xie, M., Liu, X., Ni, M., Shao, G., et al. (2014). Identification of a novel salt tolerance gene in wild soybean by whole-genome sequencing. *Nat. Commun.* 5:4340. doi: 10.1038/ncomms5340

Qu, Y., Guan, R., Bose, J., Henderson, S. W., Wege, S., Qiu, L., et al. (2021). Soybean CHX-type ion transport protein GmSALT3 confers leaf Na+ exclusion via a root derived mechanism, and Cl- exclusion via a shoot derived process. *Plant Cell Environ.* 44, 856–869. doi: 10.1111/pce.13947

Ren, D., Yang, H., Zhou, L., Yang, Y., Liu, W., Hao, X., et al. (2021). The land-water-food-environment nexus in the context of China's soybean import. *Adv. Water Resour.* 151:103892.

Ren, S., Weeda, S., Li, H., Whitehead, B., Guo, Y., Atalay, A., et al. (2012). Salt tolerance in soybean WF-7 is partially regulated by ABA and ROS signaling and involves withholding toxic Cl– ions from aerial tissues. *Plant Cell Rep.* 31, 1527–1533. doi: 10.1007/s00299-012-1268-2

Shao, G. (1988). Relationships of distribution of salt tolerant soybean cultivars and saline field. *Crops* 2, 34–36.

Shen, Y., Liu, J., Geng, H., Zhang, J., Liu, Y., Zhang, H., et al. (2018). De novo assembly of a Chinese soybean genome. *Sci. China Life Sci.* 61, 871–884.

Shin, J. H., and Lee, S. H. (2012). Molecular markers for the E2 and E3 genes controlling flowering and maturity in soybean. *Mol. Breed.* 30, 1793–1798. doi: 10.3389/fpls.2019.01303

Tsubokura, Y., Matsumura, H., Xu, M., Liu, B., Nakashima, H., Anai, T., et al. (2013). Genetic variation in soybean at the maturity locus E4 is involved in adaptation to long days at high latitudes. *Agron* 3, 117–134. doi: 10.3390/agronomy3010117

Tsubokura, Y., Watanabe, S., Xia, Z., Kanamori, H., Yamagata, H., Kaga, A., et al. (2014). Natural variation in the genes responsible for maturity loci E1. E2, E3 and E4 in soybean. *Ann. Bot.* 113, 429–441. doi: 10.1093/aob/mct269

Zeng, A., Chen, P., Korth, K., Hancock, F., Pereira, A., Brye, K., et al. (2017). Genome-wide association study (GWAS) of salt tolerance in worldwide soybean germplasm lines. *Mol. Breed.* 37:30.

Zhai, H., Lu, S., Wang, Y., Chen, X., Ren, H., Yang, J., et al. (2014). Allelic variations at four major maturity E genes and transcriptional abundance of the E1 gene are associated with flowering time and maturity of soybean cultivars. *PLoS One* 9:e97636. doi: 10.1371/journal.pone.0097636

Zhang, W., Liao, X., Cui, Y., Ma, W., Zhang, X., Du, H., et al. (2019). A cation diffusion facilitator, GmCDF1, negatively regulates salt tolerance in soybean. *PLoS Genet.* 15:e1007798. doi: 10.1371/journal.pgen.1007798

Zhang, Z., Sun, D., Tang, Y., Zhu, R., Li, X., Gruda, N., et al. (2021). Plastic shed soil salinity in China: current status and next steps. *J. Clean. Prod.* 296:126453.

Zhu, J. (2016). Abiotic stress signaling and responses in plants. *Cell* 167, 313–324.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 09 December 2021
doi: 10.3389/fpls.2021.795919

CsGSTU8, a Glutathione S-Transferase From *Camellia sinensis*, Is Regulated by CsWRKY48 and Plays a Positive Role in Drought Tolerance

*Yongheng Zhang, Jingyuan He, Yezi Xiao, Yingao Zhang, Yingqin Liu, Siqing Wan, Lu Liu, Yuan Dong, Huan Liu and Youben Yu**

College of Horticulture, Northwest A&F University, Xianyang, China

OPEN ACCESS

Edited by:
Rajeev K. Varshney, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), India

Reviewed by:
Xinchao Wang, Tea Research Institute, Chinese Academy of Agricultural Sciences (CAAS), China
Huiying Li, Wuhan Botanical Garden, Innovative Academy of Seed Design (INASEED), Chinese Academy of Sciences (CAS), China

***Correspondence:**
Youben Yu
yyben@163.com

Specialty section:
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

Received: *15 October 2021*
Accepted: *17 November 2021*
Published: *09 December 2021*

Citation:
Zhang Y, He J, Xiao Y, Zhang Y, Liu Y, Wan S, Liu L, Dong Y, Liu H and Yu Y (2021) CsGSTU8, a Glutathione S-Transferase From Camellia sinensis, Is Regulated by CsWRKY48 and Plays a Positive Role in Drought Tolerance. Front. Plant Sci. 12:795919. doi: 10.3389/fpls.2021.795919

Glutathione S-transferases (GSTs) constitute a large family of enzymes with a wide range of cellular functions. Recently, plant GSTs have gained a great deal of attention due to their involvement in the detoxification of electrophilic xenobiotics and peroxides under adverse environmental conditions, such as salt, cold, UV-B and drought stress. A previous study reported that a GST gene (*CsGSTU8*) in tea plant was distinctly induced in response to drought, suggesting this gene plays a critical role in the drought stress response. In this study, by using quantitative real-time PCR (qRT-PCR) and β-glucuronidase (GUS) reporter lines, we further demonstrated that *CsGSTU8* was upregulated in response to drought stress and exogenous abscisic acid (ABA) treatments. Overexpression of *CsGSTU8* in *Arabidopsis* resulted in enhanced drought tolerance as indicated by the improved scavenging of excess amounts of reactive oxygen species (ROS) under drought conditions. Furthermore, we found that CsWRKY48 acts as a transcriptional activator and that its expression is induced in response to drought stress and ABA treatment. Electrophoretic mobility shift assays (EMSAs), dual-luciferase (LUC) assays and transient expression assays in tea plant leaves revealed that CsWRKY48 directly binds to the W-box elements in the promoter of *CsGSTU8* and activates its expression. Taken together, our results provide additional knowledge of drought stress responses in tea plant.

Keywords: *Camellia sinensis*, glutathione S-transferases (GSTs), ROS, WRKY TF, drought stress

INTRODUCTION

Adverse environmental conditions, especially drought conditions, greatly limit plant growth and development (Shao et al., 2008). Plants have evolved intricate defensive systems to survive under drought-stress conditions, specifically, regulating the balance of reactive oxygen species (ROS) (including hydrogen peroxide (H_2O_2), superoxide (O^{2-}), $OH^{\bullet}$ and $^1O_2$) in cells is an indispensable strategy (Carvalho, 2008; Singh et al., 2019). On the one hand, ROS are necessary for plant growth, as plants actively produce ROS that serve as signal transduction molecules for growth under suitable conditions. On the other hand, when plant experience drought stress, excessive amounts of ROS

accumulate in cells and cause irreversible damage to membranes, proteins, and RNA and DNA molecules, even resulting in oxidative destruction of the cells (Moller et al., 2007; Choudhury et al., 2017). Thus, accelerating the expression or activities of antioxidant enzymes and antioxidants, which are involved in ROS scavenging, is critical for the drought tolerance of plants (Smirnoff, 1993; Gill and Tuteja, 2010; Qi et al., 2018).

Glutathione S-transferases (GSTs) constitute a large family of enzymes with a wide range of cellular functions in plants, including protecting organisms against oxidative stress under stress conditions by participating in ROS scavenging (Jha et al., 2011; Rong et al., 2014; Qi et al., 2018). Plant GSTs can be divided into eight distinct subclasses according to their protein sequence and function: Phi (GSTF), Tau (GSTU), Theta (GSTT), Zeta (GSTZ), Lambda (GSTL), Elongation factor 1 gamma (EF1G), DHAR, and TCHQD proteins (Sasan et al., 2011). Recently, several classes of GSTs, including Tau GSTs, have received a great deal of attention in plants due to the assistance of these GSTs in regulating the metabolism of oxidized molecules under drought conditions (Liu et al., 2013b; Xu et al., 2016; Choudhury et al., 2017). The expression level of *GSTs* in plants was reported to be positively correlated with the rate of oxidized molecule scavenging and to contribute to drought stress tolerance in previous studies (Kumar et al., 2013; Stavridou et al., 2021). For example, transgenic *Arabidopsis* plants overexpressing tomato *LeGSTU2* show enhanced resistance to drought stress via increased antioxidative enzyme activities for scavenging excess ROS (Xu et al., 2015). Liu et al. (2013a) reported that overexpression of a zeta GST gene from *Pyrus pyrifolia* in tobacco improved O^{2-} scavenging under drought conditions, thus resulting in increased tolerance to drought. Moreover, overexpression of a GST gene (*ThGSTZ1*) from *Tamarix hispida* improves drought tolerance by enhancing the ability to scavenge ROS (Yang et al., 2014).

Although *GSTs* have been observed to protect cells from oxidative stress under drought-stress conditions, knowledge about the intricate regulation of *GSTs* under drought stress remains scarce. To date, several *GSTs* have been reported to be directly activated by specific transcription factors to regulate drought stress tolerance. For example, in wheat, the BES/BZR family transcription factor TaBZR2 functions positively in the drought response by activating *TaGST1* to scavenge drought-induced $O2^{-}$ accumulation (Cui et al., 2019). A wheat ethylene-response factor, TaERF3, was also reported to directly activate the expression of *TaGST6* by binding to the *TaGST6* promoter to improve drought tolerance (Rong et al., 2014). Furthermore, Ren et al. (2020) demonstrated that a NAC transcription factor, ZmNST3, enhances maize drought stress tolerance by directly binding to the promoters of *GST* and *GlnRS* and activating their expression.

Camellia sinensis is a commercially important perennial evergreen woody crop species that is widely cultivated worldwide and is highly susceptible to drought stress (Parmar et al., 2019). Although several tea plant *GSTs* have been reported to be involved in cold and drought stress responses based on their expression levels (Parmar et al., 2019; Samarina et al., 2020; Sun et al., 2020), the roles of tea plant *GSTs* in drought stress are still poorly understood. *CsGSTU8* was found to be strongly induced in response to drought stress in previous studies (Xia et al., 2019) and according to our transcriptome data (unpublished), thus prompting us to investigate its role in drought stress. In this study, we used quantitative real-time PCR (qRT-PCR) and transgenic β-glucuronidase (GUS) reporter lines to confirm the upregulation of *CsGSTU8* under drought and abscisic acid (ABA) treatments. Furthermore, transgenic *Arabidopsis* plants expressing *CsGSTU8* showed improved ROS scavenging, thus contributing to drought resistance. In addition, CsWRKY48, a tea plant transcriptional activator, whose expression is induced in response to drought and ABA treatment, can directly bind to the promoter and activate the expression of *CsGSTU8*. Taken together, our findings provide evidence for the positive role of the GSTU8 protein in drought stress and are helpful for understanding the mechanism of the drought response in tea plant.

MATERIALS AND METHODS

Plant Materials and Treatment

Hydroponically grown Longjing-changye annual tea seedlings were preincubated in a greenhouse at Northwest A&F University Yangling (34°20′N, 108°24′E), Shaanxi Province, China, under natural light, and the temperature was maintained at 20/25°C in the dark/light. Seedlings displaying consistent growth were divided into three treatment groups: normal, 15% polyethylene glycol (PEG) 6,000 and ABA (100 μM) treatments. The first to third leaves were collected at 0, 4, 8, 12, 24, and 48 h after treatment for RNA extraction. For all the treatments, three biological replicates were included.

RNA Extraction and qRT-PCR Analysis

Total RNA extraction and qRT-PCR were carried out as previously described by Wan et al. (2020). The primers used in the expression analysis are listed in **Supplementary Table 1**.

Transformation of *Arabidopsis* With *CsGSTU8* and Subcellular Localization of CsGSTU8

The CDS of *CsGSTU8* was amplified from the cDNA of Longjing-changye tea plant according to the sequence information reported in the Tea Plant Information Archive (TPIA) (Xia et al., 2019). Then, the Kpn I and BamH I sites of pCAMBIA2300-GFP vector were selected to generate 35S::CsGSTU8-GFP construct. To obtain transgenic *Arabidopsis*, *Agrobacterium tumefaciens* strain GV3101 harboring the 35S::CsGSTU8-GFP construct was transformed into wild-type (WT) *Arabidopsis* ecotype Columbia (Col-0) by using the floral-dip method (Zhang et al., 2006). T3-generation homozygous lines (OE3 and OE7) and Col-0 (WT) plants were used for further analysis.

To determine the subcellular localization of the CsGSTU8 protein, GV3101 cells harboring *35S*::CsGSTU8-GFP and *35S*::GFP were cultured and adjusted to an OD600 of 0.6 via an infiltration buffer (10 mM MES, 10 mM $MgCl_2$ and 100

μM acetosyringone; pH 5.7). After 2 h incubation at 25°C, the cultures were transiently expressed in 4-week-old tobacco (*Nicotiana benthamiana*) leaves, and the GFP signal was observed with a BX63 Automatic Fluorescence Microscope (Olympus, Japan) at 48 h after infiltration. The primers used are listed in **Supplementary Table 1**.

Stress Tolerance Assays

Transgenic and WT *Arabidopsis* seeds were surface sterilized with 20% NaClO (including 0.1% Triton X-100) and sown on Murashige and Skoog (MS) agar media that included mannitol (0, 100, 125, and 150 mM). Media containing seeds were placed in a controlled growth chamber under 16/8 h (25/20°C) day/night conditions for growth after incubation at 4 °C for 3 days. The survival rate was calculated after 6-days (normal conditions) or 10-days (mannitol conditions) of growth. Seedlings with green leaves under mannitol conditions were considered alive.

For drought treatment in soil, WT and transgenic seeds were germinated in soil (vermiculite: humus = 1:1) and grown under normal conditions for 7 days. Then watering was withheld until drought stress symptoms occurred. Relative electrolyte leakage was measured to compare the drought tolerance of WT and transgenic seedings following previously report (Jiang et al., 2019).

Measurement of Reactive Oxygen Species Contents, Malondialdehyde Contents and Glutathione S-Transferase Activity

3,3′-Diaminobenzidine (DAB) staining and nitro blue tetrazolium (NBT) staining were performed to detect the accumulation of H_2O_2 and O^{2-} in the leaves according to our previously described methods (Zhang et al., 2020). The contents of H_2O_2 and O^{2-} were measured following the previous report (Shi et al., 2010). To determine the Malondialdehyde (MDA) content, a thiobarbituric acid (TBA) reaction was performed according to a previously described method (Draper et al., 1992). The GST activity was measured according to the protocol provided with a glutathione-S-transferase Assay Kit (BCO350, Solarbio, China).

β-Glucuronidase Reporter Construction and β-Glucuronidase Activity Assays

The 1,500 bp promoter of *CsGSTU8* was inserted into a pCAMBIA1300-GUS vector (Hind III and BamH I) to activate the GUS gene. Then, the construct was transformed into WT *Arabidopsis*; 7-day-old T2-generation seedings were treated with ABA (1 μM) and dehydrated (20% PEG 6,000). After 12 h of treatment, GUS staining and activity were measured as described previously (Jefferson et al., 1987). The primers used are listed in **Supplementary Table 1**.

Electrophoretic Mobility Shift Assays

Full-length CsWRKY48 (TEA008513) was inserted into a pGEX4T-1 vector to construct a GST-CsWRKY48 expression vector, and the recombinant plasmid was transformed into *E. coli* strain Rosetta2 (EC1014, Weidi Biotechnology, Shanghai, China). Expression of the GST and GST-CsWRKY48 proteins was induced by 1 mM isopropyl-b-D-thiogalactoside (IPTG) at 16°C for 20 h, and then the fusion proteins were purified using glutathione Sepharose (GSTrap^{TM} HP, GE Healthcare, United States). A biotin-labeled probe containing "W-box" elements in the promoter of *CsGSTU8* was synthesized by Tsingke. EMSAs were carried out according to the protocol provided with a chemiluminescent EMSA kit (GS009, Beyotime Biotechnology, China).

Transient Transcriptional Activation Assays

The 1,500 bp promoter fragment of *CsGSTU8* was inserted into the luciferase (LUC) reporter plasmid pGreen II 0800, which contains a Renilla luciferase (REN) gene under the control of the *35S* promoter used as an internal control. The *35S*::CsWRKY48-GFP vector as effector. The effector plasmids and the reporter plasmids were transformed into GV3101 (pSoup), and the effector and corresponding reporter in GV3101 (pSoup) were mixed together at a proportion of 2:1 and subsequently injected into tobacco leaves. Forty-eight hours later, the LUC signal was visualized with a CCD system (Lumazone Pylon 2048B, Princeton, United States). The activities of LUC and REN were determined according to the protocol of a dual-Luciferase reporter assay kit (FR201, TransGen Biotech, China) by using a full-wavelength multifunctional enzyme labeling instrument (Tecan Infinite M200PRO, Tecan, Switzerland).

Two-month-old Fudingdaba tea plants grown from seeds in a growth chamber were selected for transient expression in accordance with the methods of Shui et al. (2021) with slight modifications. *A. tumefaciens* strain GV3101 harboring *35S*::CsWRKY48-GFP and *35S*::GFP was injected into the different sides of the same leaf. At 3 days after injection, leaves were collected for RNA isolation, and RT-PCR and qRT-PCR were then carried out to assess the transformation results and the measure expression of *CsGSTU8*, respectively. The primers used are listed in **Supplementary Table 1**.

Statistical Analysis

SPSS 19.0 software was used for statistical data analysis. Data for tissue-specific expression, survival rate and relative electrolyte leakage were subjected to one-way analysis of variance, and the differences between means were assessed by Duncan's multiple range tests. Other data were examined by Student's *t*-tests. The values are represented as means ± standard deviations.

RESULTS

Identification, Expression in Different Tissues and Subcellular Localization of CsGSTU8

Our transcriptome data (**Supplementary Figure 1**) and those from the TPIA[1] revealed that TEA019065 was strongly responsive

[1]http://tpia.teaplant.org/

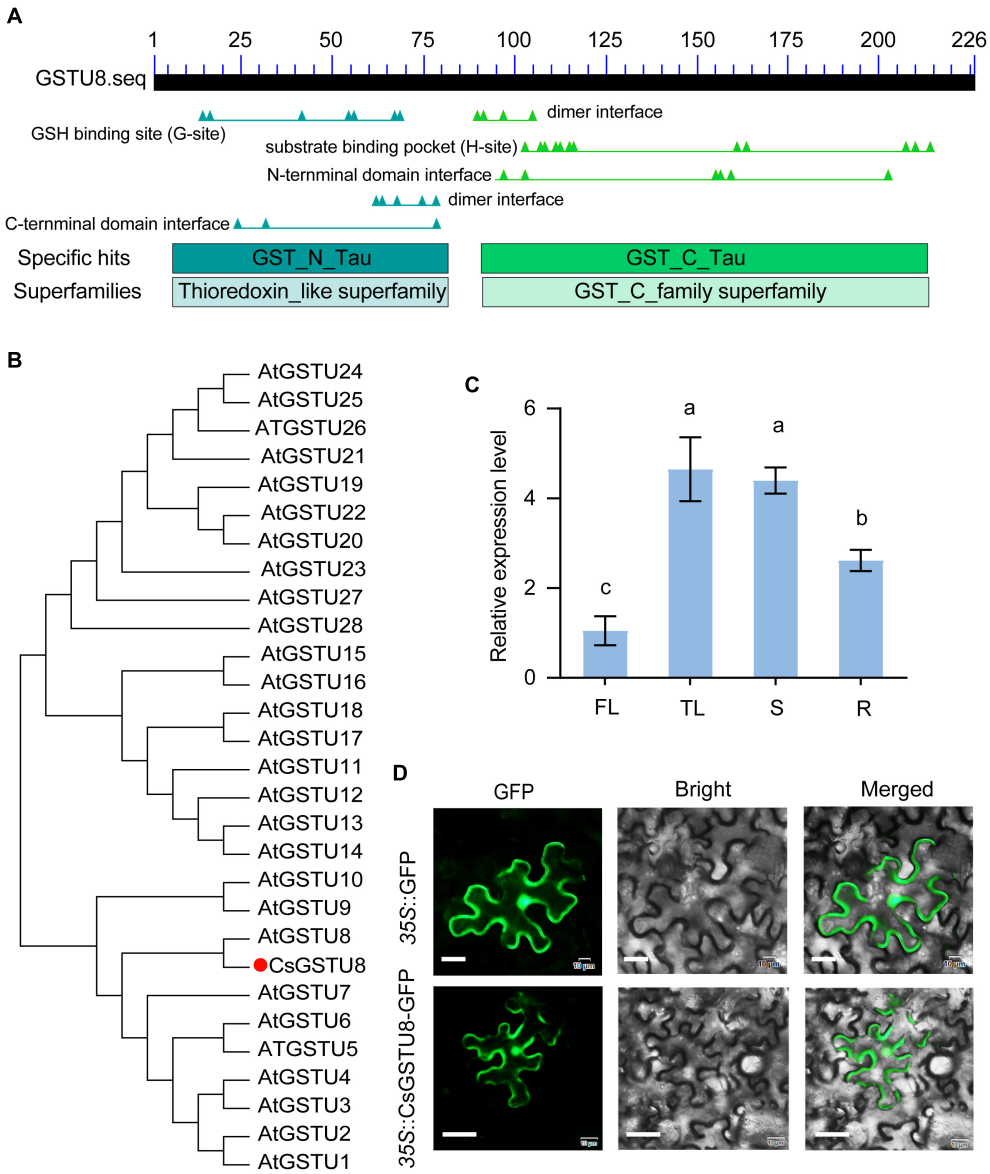


FIGURE 1 | Characterization of CsGSTU8 from *Camellia sinensis*. **(A)** Conserved domain analysis of the CsGSTU8 protein sequence retrieved from the NCBI database (https://www.ncbi.nlm.nih.gov/Structure/cdd/). **(B)** Phylogenetic analysis of CsGSTU8 and GSTU proteins from *Arabidopsis*. **(C)** Expression of CsGSTU8 in different tissues. FL (first leaf), TL (third leaf), S (stems), R (root tips). The data are the means ± SDs of three independent experiments. The values not followed by the same letter are significantly different according to Duncan's multiple range test ($P < 0.05$). **(D)** Subcellular localization of 35S::CsGSTU8-GFP and 35S::GFP in tobacco cells,. Bar = 20 μm.

to drought stress (Xia et al., 2019). In addition, it was also reported to be distinctly induced in response to drought stress in both drought-tolerant and drought-sensitive tea plants (Parmar et al., 2019). These findings prompted us to investigate the role of this gene in drought stress. Therefore, we isolated the complete CDS of TEA019065 from Longjing-changye seedings and found that it encodes a GST belonging to the Tau subfamily and harbors a GSH-binding site in the N-terminal region (**Figure 1A**). The protein comprises 226 amino acids and has a molecular weight of 26.36 kDa. Phylogenetic analysis including all Tau GST proteins from *Arabidopsis* showed that TEA019065 is closely related to AtGSTU8 (**Figure 1B**); hence, it was named CsGSTU8 in this study.

To explore the expression levels of *CsGSTU8* in different tissues, we investigated its expression in the first leaves, third leaves, stems and root tips via qRT-PCR. The results showed that *CsGSTU8* tended to expressed greater in the mature organization (the third leaves and stem) than young organs (the first leaves and root tips) (**Figure 1C**).

A. tumefaciens strain GV3101 harboring *35S*::CsGSTU8-GFP and *35S*::GFP was introduced into tobacco leaves to investigate the subcellular localization of the CsGSTU8 protein. The

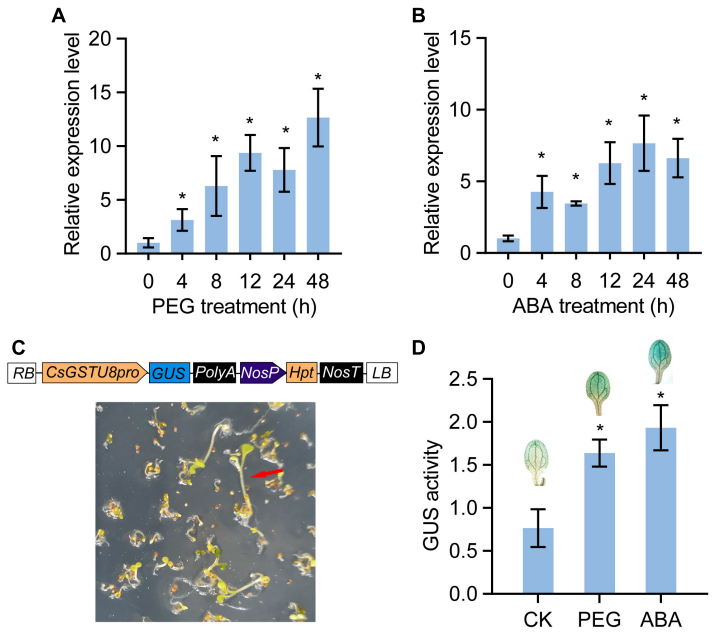


FIGURE 2 | Expression patterns of *CsGSTU8* under PEG and ABA treatments. **(A)** Expression of *CsGSTU8* after PEG treatment. **(B)** Expression of *CsGSTU8* under ABA treatment. **(C)** Construction of *CsGSTU8pro*::GUS transgenic lines. Hygromycin B was used to screen transgenic *Arabidopsis* plants. **(D)** Measurement of the GUS activities in *CsGSTU8pro*::GUS transgenic *Arabidopsis* leaves under PEG and ABA treatments. The Data are presented as the means ± SDs of three independent experiments. Significant differences were determined using Student's *t*-test (*$P < 0.05$).

fluorescence of GFP fused to CsGSTU8 showed no difference from that of the GFP control, which means that the CsGSTU8 protein was distributed throughout the cell (**Figure 1D**).

Expression Analysis of *CsGSTU8* Under Drought and Abscisic Acid Treatments

To validate the drought response of *CsGSTU8*, tea plants were treated with 15% PEG 6,000 to mimic drought stress, and the transcript levels of *CsGSTU8* were measured by qRT-PCR. The results showed that the expression of *CsGSTU8* gradually increased in response to drought stress (**Figure 2A**). Notably, ABREs involved in ABA responsiveness were found in the promoter of *CsGSTU8* (**Supplementary Figure 2**), implying that the *CsGSTU8* expression is induced in response to ABA treatment. qRT-PCR was used to measure the expression of *CsGSTU8* under ABA treatment, and as expected, ABA induced the expression of *CsGSTU8* (**Figure 2B**). To better understand *CsGSTU8* expression patterns, a *CsGSTU8pro*::GUS transgenic line was generated and subjected to drought and ABA; the results showed that GUS activity significantly increased under drought stress and ABA treatment (**Figures 2C,D**). Taken together, these results further reveal that the expression of *CsGSTU8* was induced in response to drought stress and ABA treatment.

CsGSTU8 Confers Drought Tolerance to Transgenic *Arabidopsis* Plants

Transgenic *Arabidopsis* plants were generated to explore the function of *CsGSTU8* in tolerance to drought stresses. Two T3-generation homozygous lines (OE3 and OE7) with high expression of *CsGSTU8* gene were selected for further analysis (**Supplementary Figure 3**). Various concentrations of mannitol were employed to mimic drought stress. As expected, in the absence of mannitol condition, the survival rates of the two transgenic lines were indistinguishable from those of the WT (**Figures 3A,B**). The application of mannitol inhibited the growth of both the WT and transgenic lines in a dose-dependent manner; however, the survival rates of the OE3 and OE7 lines significantly increased compared to those of the WT plants under 100, 125, and 150 mM mannitol conditions (**Figures 3A,B**).

Moreover, under soil water shortage conditions, the WT plants exhibited a more severely wilted phenotype than did the OE3 and OE7 transgenic lines (**Figure 3C**). We further measured the relative electrolyte leakage to evaluate membrane damage in *Arabidopsis* leaves, and the results showed that there was no difference between the WT and transgenic plants under normal conditions, while under drought stress, the relative electrolyte leakage of the two transgenic lines was lower than that of the WT (**Figure 3D**). These results indicate that overexpression of *CsGSTU8* enhances drought tolerance in *Arabidopsis*.

Overexpression of *CsGSTU8* Alleviates Reactive Oxygen Species and Malondialdehyde Accumulation in *Arabidopsis*

As it is well established that GSTs contribute to alleviating excess amounts of ROS when plants experience stress (Sharma et al., 2014; Yang et al., 2014), we evaluated the GST activity and

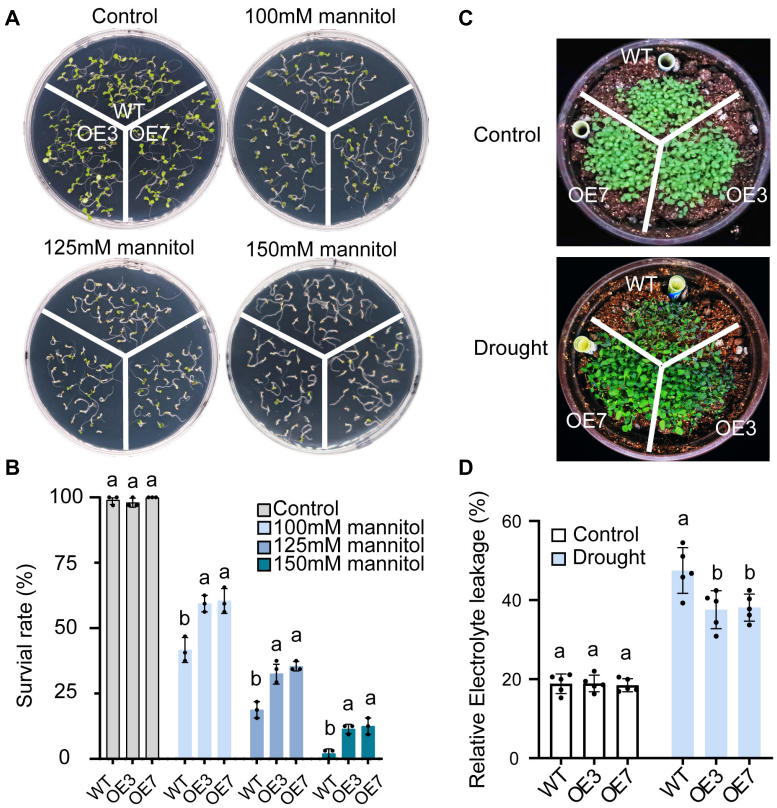


FIGURE 3 | Overexpression of *CsGSTU8* improves drought stress tolerance in transgenic *Arabidopsis*. **(A)** Growth phenotypes under control and mannitol conditions in MS media. **(B)** Measurement of survival rate in MS media. The Data are presented as the means ± SDs of three independent experiments. **(C)** Growth phenotypes under control and drought stress conditions in the soil. **(D)** Measurement of relative electrolyte leakage. The Data are the means ± SDs of five independent experiments. The Values not followed by the same letter are significantly different according to Duncan's multiple range test ($P < 0.05$).

ROS content to explore whether enhanced drought tolerance is associated with a decrease of ROS levels in transgenic lines. The results showed that GST activity was higher in the transgenic plants than in the WT plants under both control and drought conditions (**Figure 4B**). Interestingly the contents of H_2O_2 and O^{2-} in the WT and transgenic lines were not significantly different under normal conditions, while under drought stress, the OE3 and OE7 transgenic lines accumulated less H_2O_2 and O^{2-} than the WT plants did (**Figures 4A,C,D**). MDA is considered a marker of membrane lipid peroxidation caused by excessive ROS (Jambunathan, 2010), so we also measured the MDA content of WT and transgenic lines under drought stress. The transgenic lines consistently accumulated lower MDA than did the WT under drought conditions (**Figure 4E**). These results implying that *CsGSTU8* overexpression in *Arabidopsis* alleviates ROS and MDA accumulation under drought stress.

CsWRKY48 Is a Transcriptional Activator Involved in Responses to Drought and Abscisic Acid in Tea Plant

Our team previously reported that *CsWRKYIIc3* (annotated as *CsWRKY48* in the TPIA) is involved in responses to drought and ABA in tea plant Shaancha No. 1 tea plant (Xiao et al., 2020). Our unpublished transcriptome data also demonstrated that *CsWRKY48* expression was induced in response to drought (**Supplementary Figure 1**). In this study, qRT-PCR was used to explore whether *CsWRKY48* responds to drought and ABA treatment in Longjing-changye seedings, and the results showed that *CsWRKY48* was indeed induced in responses to drought and ABA treatment (**Figures 5C,D**). Amino acid sequence analysis and subcellular localization results revealed that CsWRKY48 contains a highly conserved WRKY DNA-binding domain (**Figure 5A**) and localizes to the nucleus (**Figure 5E**), so we fused the CDS of *CsWRKY48* with the Gal4-binding domain in the pGBKT7 vector, after which the construct was transferred to yeast strain Y2H Gold to investigate its transactivation capacity. The results showed that the fusion of CsWRKY48 with Gal4 was able to activate *HIS3*, *ADE2* and *MEL1* and grow on selective media (**Figure 5B**), indicating that CsWRKY48 has potential transcriptional activation.

CsWRKY48 Bind to the *CsGSTU8* Promoter and Activates the Expression of *CsGSTU8*

W-boxes were found in the promoter of *CsGSTU8* (**Supplementary Figure 2**), implying that the expression of

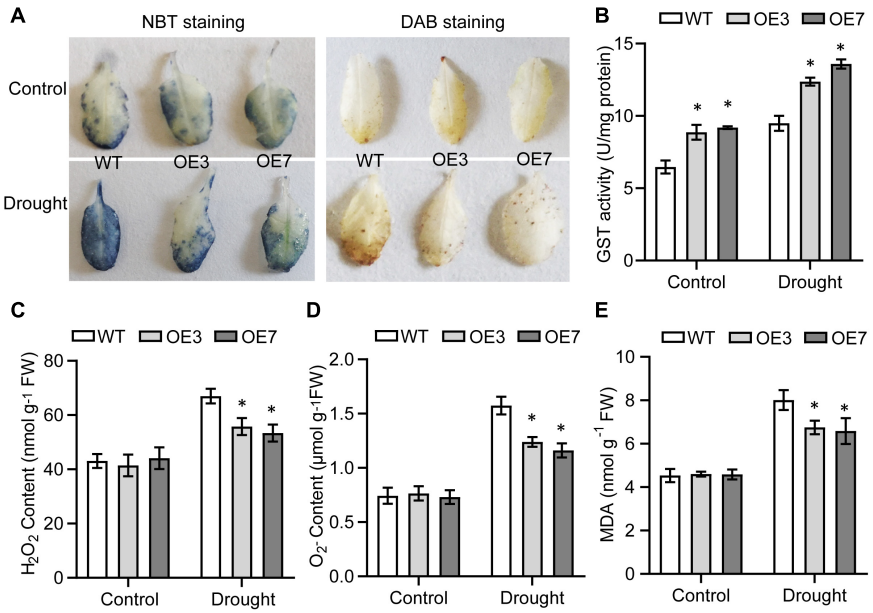


FIGURE 4 | Overexpression of *CsGSTU8* alleviated ROS and MDA accumulation under drought stress. **(A)** Histochemical detection of O^{2-} and H_2O_2 via NBT staining and DAB staining, respectively. **(B)** Measurement of GST activity under control and drought conditions. **(C)** Measurement of H_2O_2 content under control and drought conditions. **(D)** Measurement of O^{2-} content under control and drought conditions. **(E)** Measurement of MDA content under control and drought conditions. The Data are presented as the means ± SDs of three independent experiments. Significant differences were determined using Student's *t*-test (*$P < 0.05$).

CsGSTU8 might be regulated by WRKY transcription factors. Remarkably, our results revealed that the expression of *CsGSTU8* similar to that of *CsWRKY48*, was induced in response to both drought and ABA treatment, and we inferred that CsWRKY48 may be involved in the regulation of *CsGSTU8* expression. Thus, EMSAs and dual-LUC transient expression assays were carried out to verify this hypothesis. First, GST and GST-CsWRKY48 were purified for use in EMSAs and the band shifts were observed when the GST-CsWRKY48 protein was incubated together with a biotin-labeled probe (containing the core sequences (TGAC) of W-boxes), while band shifts were not observed in the GST protein control (**Figure 6A**), which indicated that CsWRKY48 can directly bind to the *CsGSTU8* promoter. Dual-LUC transient expression assays showed that the transient expression of *CsWRKY48* significantly increased the LUC/REN ratio of the reporter relative to that of the corresponding empty control in tobacco leaves (**Figures 6B–D**). In addition, we transiently overexpressed *CsWRKY48* in tea plant leaves (**Figures 6E,F**), and the qRT-PCR results showed that *CsWRKY48* overexpression increased the expression of *CsGSTU8* in tea plant (**Figure 6G**). Taken together, our results indicate that CsWRKY48 can directly bind to the promoter of *CsGSTU8* and activate its expression.

DISCUSSION

GSTs with a wide range of functions in plants, including protecting cells from oxidative damage during stresses (Sheehan et al., 2001). A series of studies have indicated that the expression of GST genes is induced in response to drought stress in various plants, including barley (Rezaei et al., 2013), tomato (Xu et al., 2015), wheat (Wang et al., 2019), potato (Islam et al., 2018), pepper (Islam et al., 2019), and rice (Jain et al., 2010). Previous studies in tea plant have reported that drought stress also induces the expression of members of *CsGSTs* (Parmar et al., 2019; Xia et al., 2019; Sun et al., 2020), including *CsGSTU8*, implying that *GSTs* from tea plant may have conserved functions in response to drought stress, similar to cases in other plants. Based on this information, we investigated the role and regulation of *CsGSTU8* in response to drought stress in this study.

As it well established that abiotic stress, such as drought, salinity and cold alter the ABA levels in plants. ABA plays an important role as an essential mediator in triggering plant responses to various abiotic stresses (Hartung et al., 1988). Several *GSTs* have been reported to be associated with ABA signaling and to be involved in the stress response. For example, Sharma et al. (2014) reported that *OsGSTU4* was induced in response to ABA and was involved in ABA-dependent processes that provide stress tolerance to transgenic plants. In wheat, the TaERF3-actived *TaGST6* expression was enhanced by ABA to respond to salt and drought stress (Rong et al., 2014). Our results indicated that the expression of *CsGSTU8* was increased in response to ABA and drought treatments, meanwhile, the ABA- and drought-responsive transcription factor CsWRKY48 could directly activate the expression of *CsGSTU8*. Thus, we inferred that ABA signaling might be involved in the drought-induced expression of *CsGSTU8*.

One of effects of drought stresses on plants is the excess accumulation of ROS. A series of studies have shown that transgenic plants expressing *GSTs* contribute to drought stress

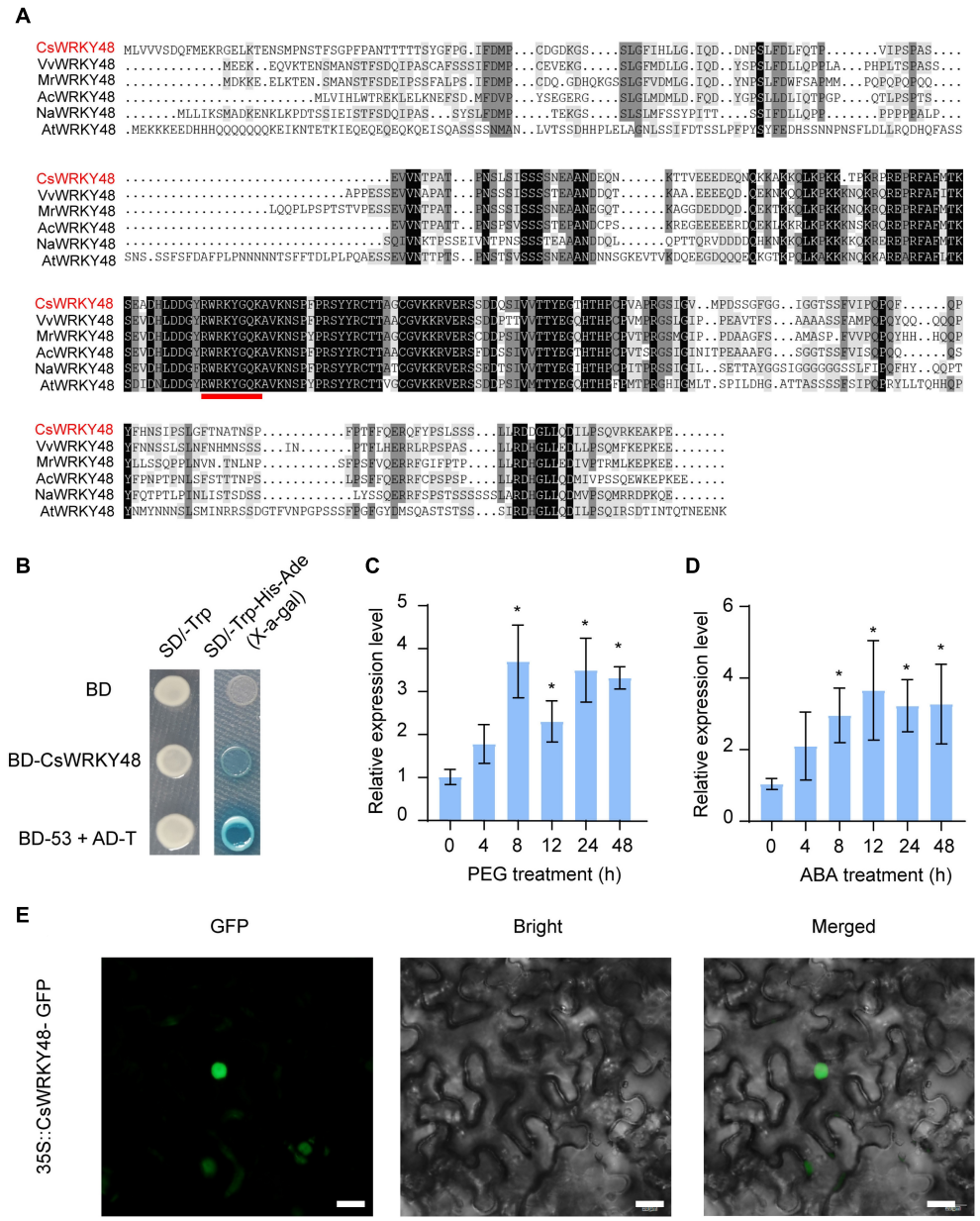


FIGURE 5 | CsWRKY48 acts as a transcriptional activator and is expressed in response to drought and ABA treatment. **(A)** Multiple sequence alignment of CsWRKY48 and its homologs in other plant species. The red line indicates the WRKY DNA-binding domain. **(B)** Transactivation analysis of CsWRKY48 in yeast. **(C)** Transcript levels of *CsWRKY48* under PEG treatment. **(D)** Transcript levels of *CsWRKY48* under ABA treatment. **(E)** Subcellular localization of 35S::CsWRKY48-GFP in tobacco cells. Bar = 20 μm. The data are presented as the means ± SDs of three independent experiments. Significant differences were determined using Student's *t*-test (*$P < 0.05$).

tolerance by scavenging excess ROS accumulation (Liu et al., 2013a; Srivastava et al., 2019; Yang et al., 2019). Similarly, in the current study, transgenic *Arabidopsis* plants expressing *CsGSTU8* exhibited a stress-tolerant phenotype when subjected to drought stress accompanied by less accumulation of ROS level, suggesting the positive role of CsGSTU8 in ROS scavenging. GST proteins catalyze the conjugation of GSH to an array of hydrophobic and electrophilic substrates, including ROS, thus, protecting the cell from oxidative burst (Kumar and Trivedi, 2018). During catalysis, the binding and correct orientation of GSH are governed by the conserved GSH binding site (G-site), and the substrate binding pocket (H-site) assists in the binding of substrates by providing a hydrophobic environment (Basantani and Srivastava, 2007; Kumar and Trivedi, 2018). CsGSTU8 harboring the conserved GSH bind site and the substrate binding pocket, suggesting the capacity of CsGSTU8 to catalyze the conjugation of GSH to an array of substrates, which was further confirmed by the higher GST activity of transgenic plants compared to that in WT plants.

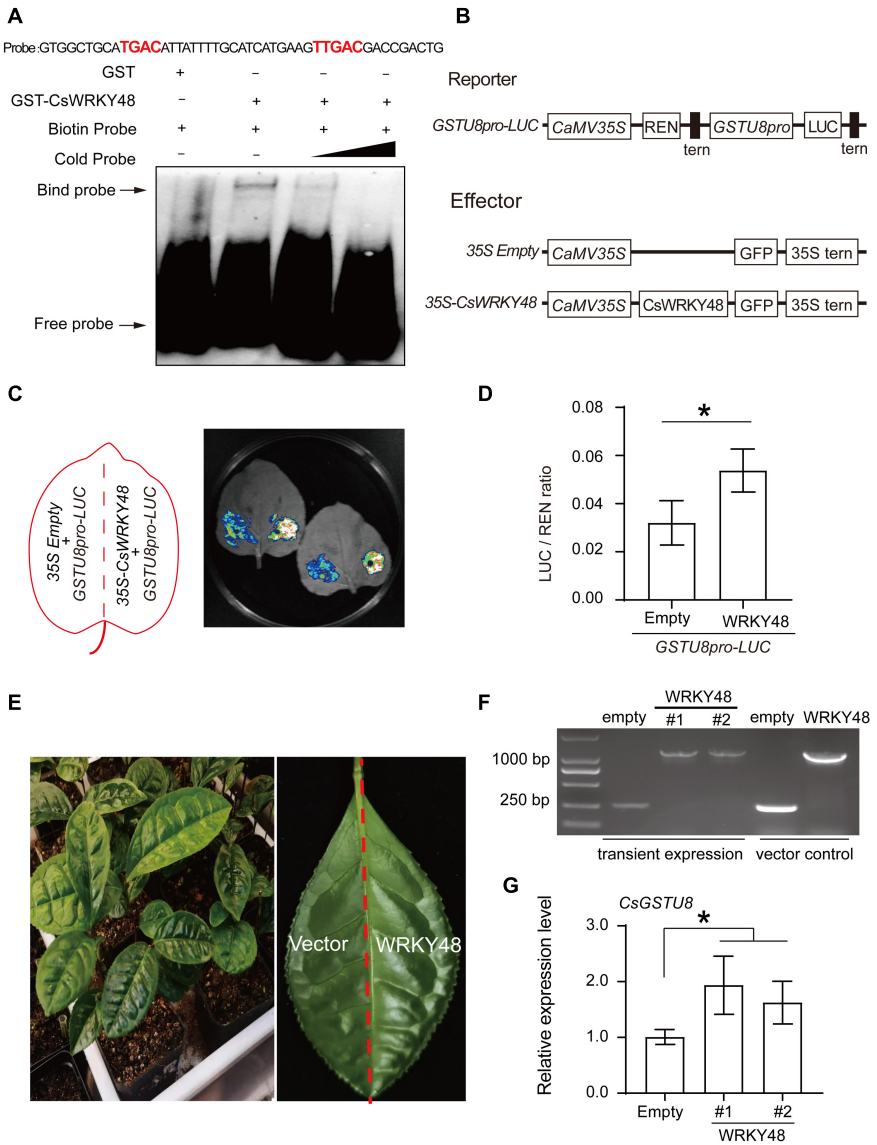


FIGURE 6 | CsWRKY48 binds to the promoter of *CsGSTU8* and activates its expression. **(A)** EMSAs using purified GST and GST-CsWRKY48 proteins incubated together with biotin-labeled probes of W-box core elements (TGAC sequences) presented in *CsGSTU8* promoter. -: absence; +: presence. **(B)** Schematics of the reporter and effector constructs used in the dual-LUC assays. **(C)** LUC signal detected in tobacco leaves. **(D)** LUC/REN ratio, as measured by the dual-LUC assays. **(E)** Transient expression of 35S::CsWRKY48-GFP and 35S::GFP in tea plant leaves. **(F)** RT-PCR was used to identify the transformation of 35S::CsWRKY48-GFP and 35S::GFP by using specific primers. **(G)** Expression of *CsGSTU8* in transformed tea plant leaves. The data are presented as the means ± SDs of three independent experiments. Significant differences were determined using Student's *t*-test (*$P < 0.05$).

Thus, the lower ROS level of transgenic plants under drought stress can be explained by the fact that transgenic plants have an improved capacity for GSH conjugation to ROS and result in an enhanced ability to scavenge ROS, which constitutes the foundation for their stress-tolerant phenotype.

During signal transduction and the adaptive response, transcription factors including NAC, MYB, bZIP, DREB, and WRKY ones usually bind to the promoters of effector genes that encode enzymes, chaperones and ion/water channels and directly regulate their expression to induce an adaptive response under drought conditions (Joshi et al., 2016; Wei et al., 2019). As effector genes encoding enzymes involved in ROS scavenging under stress, *GSTs* have been found to be directly regulated by several type transcription factors in response to drought stress in plants, such as ERF type (Rong et al., 2014), bHLH type (Cui et al., 2019), and NAC type (Ren et al., 2020). WRKY family members regulate the expression of target genes by binding to the W-box cis-elements in their promoters (Deng et al., 2020; Zhao et al., 2020). For example, AtWRKY57 can directly bind to the W-box in *AtRD29A* and *AtNCED3* promoters and activate their expression (Jiang et al., 2012). Our EMSA and Dual-LUC transient expression assays demonstrated that a drought- and

ABA-responsive transcription factor form tea plant, CsWRKY48, can directly bind to the W-box in promoter of *CsGSTU8* and activate its expression. These results suggest that drought stress may first induces the expression of *CsWRKY48* and then activates the *CsGSTU8* transcription, thus causing an adaptive response.

CONCLUSION

In conclusion, our study revealed that *CsGSTU8*, which is positively activated by CsWRKY48, enhances the drought tolerance of transgenic *Arabidopsis* by increasing ROS scavenging under drought stress. This study provides valuable knowledge for understanding the function of CsGSTU8 and the underlying molecular mechanism of drought tolerance in tea plant.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

YY and YoZ designed the experiments. YoZ and JH performed the experiments and data analysis. YoZ and YX wrote the manuscript. SW, YiZ, and YL provided valuable advice on the manuscript. LL, YD, and HL revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the China Agriculture Research System of MOF and MARA (CARS-19), the Agricultural Special Fund Project of Shaanxi Province (NYKJ-2020-YL-13), the special fund for University-Supported Extension Model (XTG2021-04), and Key R&D Program of Shaanxi Province (2020LSFP3-16).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.795919/full#supplementary-material

REFERENCES

Basantani, M., and Srivastava, A. (2007). Plant glutathione transferases — a decade falls short. *Canadian J. Botany* 85, 443–456.

Carvalho, M. H. C. D. (2008). Drought stress and reactive oxygen species: production, scavenging and signaling. *Plant Signal. Behav.* 7, 599–601. doi: 10.4161/psb.3.3.5536

Choudhury, F. K., Rivero, R. M., Blumwald, E., and Mittler, R. (2017). Reactive oxygen species, abiotic stress and stress combination. *Plant J.* 90, 856–867. doi: 10.1111/tpj.13299

Cui, X. Y., Gao, Y., Guo, J., Yu, T. F., Zheng, W. J., Liu, Y. W., et al. (2019). BES/BZR transcription factor TaBZR2 positively regulates drought responses by activation of TaGST1. *Plant Physiol.* 180, 605–620. doi: 10.1104/pp.19.00100

Deng, B., Wang, W., Ruan, C., Deng, L., Yao, S., and Zeng, K. (2020). Involvement of CsWRKY70 in salicylic acid-induced citrus fruit resistance against Penicillium digitatum. *Hortic. Res.* 7:157.

Draper, H. H., Squires, E. J., Mahmood, H., Wu, J., Agarwa, S., and Hadley, M. (1992). A comparative evaluation of thiobarbituric acid methods for the determination of malondialdehyde in biological materials. *Free Radic. Biol. Med.* 15, 353–363. doi: 10.1016/0891-5849(93)90035-s

Gill, S. S., and Tuteja, N. (2010). Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. *Plant Physiol. Biochem.* 48, 909–930. doi: 10.1016/j.plaphy.2010.08.016

Hartung, W., Radin, J. W., and Hendrix, D. L. (1988). Abscisic acid movement into the apoplastic solution of water-stresses cotton leaves. *Plant Physiol.* 86, 908–913.

Islam, M. S., Choudhury, M., Majlish, A. K., Islam, T., and Ghosh, A. (2018). Comprehensive genome-wide analysis of Glutathione S-transferase gene family in potato (*Solanum tuberosum* L.) and their expression profiling in various anatomical tissues and perturbation conditions. *Gene* 639, 149–162. doi: 10.1016/j.gene.2017.10.007

Islam, S., Sajib, S. D., Jui, Z. S., Arabia, S., Islam, T., and Ghosh, A. (2019). Genome-wide identification of glutathione S-transferase gene family in pepper, its classification, and expression profiling under different anatomical and environmental conditions. *Sci. Rep.* 9:9101. doi: 10.1038/s41598-019-45320-x

Jain, M., Ghanashyam, C., and Bhattacharjee, A. (2010). Comprehensive expression analysis suggests overlapping and specific roles of rice glutathione S-transferase genes during development and stress responses. *BMC Genomics* 11:73. doi: 10.1186/1471-2164-11-73

Jambunathan, N. (2010). Determination and detection of reactive oxygen species (ROS), lipid peroxidation, and electrolyte leakage in plants. *Methods Mol. Biol.* 639, 292–298. doi: 10.1007/978-1-60761-702-0_18

Jefferson, R. A., Kavanagh, T. A., and Bevan, M. W. (1987). GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* 6, 3901–3907.

Jha, B., Sharma, A., and Mishra, A. (2011). Expression of SbGSTU (tau class glutathione S-transferase) gene isolated from *Salicornia brachiata* in tobacco for salt tolerance. *Mol. Biol. Rep.* 38, 4823–4832. doi: 10.1007/s11033-010-0625-x

Jiang, Q., Yang, J., Wang, Q., Zhou, K., Mao, K., and Ma, F. (2019). Overexpression of MdEPF2 improves water use efficiency and reduces oxidative stress in tomato. *Environ. Exp. Bot.* 162, 321–332. doi: 10.1007/s00394-014-0770-4

Jiang, Y., Liang, G., and Yu, D. (2012). Activated expression of WRKY57 confers drought tolerance in *Arabidopsis*. *Mol. Plant* 5, 1375–1388. doi: 10.1093/mp/sss080

Joshi, R., Wani, S. H., Singh, B., Bohra, A., Dar, Z. A., Lone, A. A., et al. (2016). Transcription factors and plants response to drought stress: current understanding and future directions. *Front. Plant Sci.* 7:1029. doi: 10.3389/fpls.2016.01029

Kumar, S., Asif, M. H., Chakrabarty, D., Tripathi, R. D., Dubey, R. S., and Trivedi, P. K. (2013). Expression of a rice Lambda class of glutathione S-transferase, OsGSTL2, in *Arabidopsis* provides tolerance to heavy metal and other abiotic stresses. *J. Hazard. Mater.* 248-249, 228–237. doi: 10.1016/j.jhazmat.2013.01.004

Kumar, S., and Trivedi, P. K. (2018). Glutathione S-Transferases: role in combating abiotic stresses including arsenic detoxification in plants. *Front. Plant Sci.* 9:751. doi: 10.3389/fpls.2018.00751

Liu, D., Liu, Y., Rao, J., Wang, G., Li, H., Ge, F., et al. (2013a). Overexpression of the glutathione S-transferase gene from *Pyrus pyrifolia* fruit improves tolerance to abiotic stress in transgenic tobacco plants. *Mol. Biol.* 47, 515–523. doi: 10.7868/s0026898413040101

Liu, Y. J., Han, X. M., Ren, L. L., Yang, H. L., and Zeng, Q. Y. (2013b). Functional divergence of the glutathione S-transferase supergene family in *Physcomitrella* patens reveals complex patterns of large gene family evolution in land plants. *Plant Physiol.* 161, 773–786. doi: 10.1104/pp.112.205815

Moller, I. M., Jensen, P. E., and Hansson, A. (2007). Oxidative modifications to cellular components in plants. *Annu. Rev. Plant Biol.* 58, 459–481. doi: 10.1146/annurev.arplant.58.032806.103946

Parmar, R., Seth, R., Singh, P., Singh, G., Kumar, S., and Sharma, R. K. (2019). Transcriptional profiling of contrasting genotypes revealed key candidates and nucleotide variations for drought dissection in *Camellia sinensis* (L.) O. Kuntze. *Sci. Rep.* 9:7487. doi: 10.1038/s41598-019-43925-w

Qi, J., Song, C. P., Wang, B., Zhou, J., Kangasjarvi, J., Zhu, J. K., et al. (2018). Reactive oxygen species signaling and stomatal movement in plant responses to drought stress and pathogen attack. *J. Int. Plant Biol.* 60, 805–826. doi: 10.1111/jipb.12654

Ren, Z., Zhang, D., Cao, L., Zhang, W., Zheng, H., Liu, Z., et al. (2020). Functions and regulatory framework of ZmNST3 in maize under lodging and drought stress. *Plant Cell Environ.* 43, 2272–2286. doi: 10.1111/pce.13829

Rezaei, M. K., Shobbar, Z. S., Shahbazi, M., Abedini, R., and Zare, S. (2013). Glutathione S-transferase (GST) family in barley: identification of members, enzyme activity, and gene expression pattern. *J. Plant Physiol.* 170, 1277–1284. doi: 10.1016/j.jplph.2013.04.005

Rong, W., Qi, L., Wang, A., Ye, X., Du, L., Liang, H., et al. (2014). The ERF transcription factor TaERF3 promotes tolerance to salt and drought stresses in wheat. *Plant Biotechnol. J.* 12, 468–479. doi: 10.1111/pbi.12153

Samarina, L. S., Bobrovskikh, A. V., Doroshkov, A. V., Malyukova, L. S., Matskiv, A. O., Rakhmangulov, R. S., et al. (2020). Comparative expression analysis of stress-inducible candidate genes in response to cold and drought in tea plant [*Camellia sinensis* (L.) Kuntze]. *Front. Genet.* 11:611283. doi: 10.3389/fgene.2020.611283

Sasan, M., Maryam, E., Fateme, M., Maryam, S., Babak, S., and Hassan, M. (2011). Plant glutathione S-transferase classification, structure and evolution. *African J. Biotechnol.* 10, 8160–8165.

Shao, H. B., Chu, L. Y., Jaleel, C. A., and Zhao, C. X. (2008). Water-deficit stress-induced anatomical changes in higher plants. *C. R. Biol.* 331, 215–225. doi: 10.1016/j.crvi.2008.01.002

Sharma, R., Sahoo, A., Devendran, R., and Jain, M. (2014). Over-expression of a rice tau class glutathione s-transferase gene improves tolerance to salinity and oxidative stresses in *Arabidopsis*. *PLoS One* 9:e92900. doi: 10.1371/journal.pone.0092900

Sheehan, D., Meade, G., Foley, V. M., and Dowd, C. A. (2001). Structure, function and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily. *Biochem J.* 360, 1–16. doi: 10.1042/bj3600001

Shi, J., Fu, X. Z., Peng, T., Huang, X. S., Fan, Q. J., and Liu, J. H. (2010). Spermine pretreatment confers dehydration tolerance of citrus in vitro plants via modulation of antioxidative capacity and stomatal response. *Tree Physiol.* 30, 914–922. doi: 10.1093/treephys/tpq030

Shui, L. Y., Yan, M. L., Li, H., Wang, P., Zhao, H., Wang, M. L., et al. (2021). *Characterization of the R2R3-MYB Transcription Factor CsMYB113 Regulates Anthocyanin Biosynthesis in Tea Plants (Camellia Sinensis)*. *PREPRINT (Version* 1). Durham, NC: Research Square.

Singh, A., Kumar, A., Yadav, S., and Singh, I. K. (2019). Reactive oxygen species-mediated signaling during abiotic stress. *Plant Gene* 18, 405–413.

Smirnoff, N. (1993). The role of active oxygen in the response of plants to water deficit and desiccation. *New Phytol.* 125, 27–58. doi: 10.1111/j.1469-8137.1993.tb03863.x

Srivastava, D., Verma, G., Chauhan, A. S., Pande, V., and Chakrabarty, D. (2019). Rice (*Oryza sativa* L.) tau class glutathione S-transferase (OsGSTU30) overexpression in *Arabidopsis thaliana* modulates a regulatory network leading to heavy metal and drought stress tolerance. *Metallomics* 11, 375–389. doi: 10.1039/c8mt00204e

Stavridou, E., Voulgari, G., Michailidis, M., Kostas, S., Chronopoulou, E. G., Labrou, N. E., et al. (2021). Overexpression of a biotic stress-inducible pvgstu gene activates early protective responses in tobacco under combined heat and drought. *Int. J. Mol. Sci.* 22:2352. doi: 10.3390/ijms22052352

Sun, J., Qiu, C., Ding, Y., Wang, Y., Sun, L., Fan, K., et al. (2020). Fulvic acid ameliorates drought stress-induced damage in tea plants by regulating the ascorbate metabolism and flavonoids biosynthesis. *BMC Genomics* 21:411. doi: 10.1186/s12864-020-06815-4

Wan, S., Zhang, Y., Duan, M., Huang, L., Wang, W., Xu, Q., et al. (2020). Integrated analysis of long non-coding RNAs (lncRNAs) and mRNAs reveals the regulatory role of lncRNAs associated with salt resistance in *Camellia sinensis*. *Front. Plant Sci.* 11:218. doi: 10.3389/fpls.2020.00218

Wang, R., Ma, J., Zhang, Q., Wu, C., Zhao, H., Wu, Y., et al. (2019). Genome-wide identification and expression profiling of glutathione transferase gene family under multiple stresses and hormone treatments in wheat (*Triticum aestivum* L.). *BMC Genomics* 20:986. doi: 10.1186/s12864-019-6374-x

Wei, W., Liang, D. W., Bian, X. H., Shen, M., Xiao, J. H., Zhang, W. K., et al. (2019). GmWRKY54 improves drought tolerance through activating genes in abscisic acid and Ca2+ signaling pathways in transgenic soybean. *Plant J.* 100, 384–398. doi: 10.1111/tpj.14449

Xia, E. H., Li, F. D., Tong, W., Li, P. H., Wu, Q., Zhao, H. J., et al. (2019). Tea plant information archive: a comprehensive genomics and bioinformatics platform for tea plant. *Plant Biotechnol. J.* 17, 1938–1953. doi: 10.1111/pbi.13111

Xiao, L., Tang, L., Wang, W., Gao, Y., Huang, Y., Meng, Y., et al. (2020). Cloning and functional analysis of CsWRKYIIcs transcription factors in tea plant. *Sci. Agricult. Sinica* 53, 2460–2476. doi: 10.1016/j.gene.2017.05.044

Xu, J., Tian, Y. S., Xing, X. J., Peng, R. H., Zhu, B., Gao, J. J., et al. (2016). Over-expression of AtGSTU19 provides tolerance to salt, drought and methyl viologen stresses in *Arabidopsis*. *Physiol. Plant.* 156, 164–175. doi: 10.1111/ppl.12347

Xu, J., Xing, X. J., Tian, Y. S., Peng, R. H., Xue, Y., Zhao, W., et al. (2015). Transgenic *Arabidopsis* plants expressing tomato glutathione S-Transferase showed enhanced resistance to salt and drought stress. *PLoS One* 10:e0136960. doi: 10.1371/journal.pone.0136960

Yang, G., Wang, Y., Xia, D., Gao, C., Wang, C., and Yang, C. (2014). Overexpression of a GST gene (ThGSTZ1) from Tamarix hispida improves drought and salinity tolerance by enhancing the ability to scavenge reactive oxygen species. *Plant Cell Tissue Organ. Culture* 117, 99–112.

Yang, Q., Liu, Y.-J., and Zeng, Q. Y. (2019). Overexpression of three orthologous glutathione S-transferases from populus increased salt and drought resistance in *Arabidopsis*. *Biochem. Systemat. Ecol.* 83, 57–61.

Zhang, X., Henriques, R., Lin, S. S., Niu, Q. W., and Chua, N. H. (2006). Agrobacterium-mediated transformation of *Arabidopsis thaliana* using the floral dip method. *Nat. Protoc.* 1, 641–646. doi: 10.1038/nprot.2006.97

Zhang, Y., Wan, S., Liu, X., He, J., Cheng, L., Duan, M., et al. (2020). Overexpression of CsSnRK2.5 increases tolerance to drought stress in transgenic *Arabidopsis*. *Plant Physiol. Biochem.* 150, 162–170. doi: 10.1016/j.plaphy.2020.02.035

Zhao, X., Qi, C., Jiang, H., Zhong, M., You, C., Li, Y., et al. (2020). MdWRKY15 improves resistance of apple to *Botryosphaeria dothidea* via the salicylic acid-mediated pathway by directly binding the MdICS1 promoter. *J. Integr. Plant Biol.* 62, 527–543. doi: 10.1111/jipb.12825

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 04 January 2022
doi: 10.3389/fpls.2021.792862

Genome-Wide Investigation of the Cysteine Synthase Gene Family Shows That Overexpression of *CSase* Confers Alkali Tolerance to Alfalfa (*Medicago sativa* L.)

*Yuying Yuan, Tingting Song, Jinqiu Yu, Wenkai Zhang, Xiangyin Hou, Zelai Kong Ling and Guowen Cui**

Department of Grassland Science, College of Animal Science and Technology, Northeast Agricultural University, Harbin, China

OPEN ACCESS

Edited by:
Rajeev K. Varshney, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), India

Reviewed by:
Marco Betti, Seville University, Spain
Fei Chen, Hangzhou Normal University, China

***Correspondence:**
Guowen Cui
cgw603@163.com

Specialty section:
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

Received: *11 October 2021*
Accepted: *13 December 2021*
Published: *04 January 2022*

Citation:
Yuan Y, Song T, Yu J, Zhang W, Hou X, Kong Ling Z and Cui G (2022) Genome-Wide Investigation of the Cysteine Synthase Gene Family Shows That Overexpression of CSase Confers Alkali Tolerance to Alfalfa (Medicago sativa L.). Front. Plant Sci. 12:792862. doi: 10.3389/fpls.2021.792862

Alfalfa is widely grown worldwide as a perennial high-quality legume forage and as a good ecological landcover. The cysteine synthase (CSase) gene family is actively involved in plant growth and development and abiotic stress resistance but has not been systematically investigated in alfalfa. We identified 39 *MsCSase* genes on 4 chromosomes of the alfalfa genome. Phylogenetic analysis demonstrated that these genes were clustered into six subfamilies, and members of the same subfamily had similar physicochemical properties and sequence structures. Overexpression of the *CSase* gene in alfalfa increased alkali tolerance. Compared with control plants, the overexpression lines presented higher proline, soluble sugars, and cysteine and reduced glutathione contents and superoxide dismutase and peroxidase activities as well as lower hydrogen peroxide and superoxide anion contents after alkali stress. The relative expression of γ-glutamyl cysteine synthetase gene (a downstream gene of *CSase*) in the overexpression lines was much higher than that in the control line. The *CSase* gene enhanced alkalinity tolerance by regulating osmoregulatory substances and improving antioxidant capacity. These results provide a reference for studying the CSase gene family in alfalfa and expanding the alkali tolerance gene resources of forage plants.

Keywords: alfalfa, *CSase* gene family, genome-wide analysis, gene overexpression, alkali stress

INTRODUCTION

Cysteine is the first organic substance in plants found to contain both sulfur and nitrogen and is a precursor of sulfur-containing metabolites such as methionine (Takahashi et al., 2011), glutathione (GSH) and Fe-S clusters, which play an important role in plant development and metabolic processes (Droux, 2004; Van Hoewyk et al., 2008). The synthesis of cysteine can be roughly divided into the absorption and reduction of elemental sulfur (Kopriva, 2006; Davidian and Kopriva, 2010). Cysteine synthase (CSase) is involved in the final step of cysteine synthesis; this enzyme catalyzes the synthesis of cysteine from H_2S and O-acetylserine (OAS) (Jez and Dey, 2013; Romero et al., 2014).

The *CSase* gene is often referred to as the O-acetylserine(thiol)lyase gene (*OAS-TL*) and belongs to the CSase gene family (Droux et al., 1998; Wirtz and Hell, 2006; Alvarez et al., 2010a). The CSase gene family was previously identified in Arabidopsis (*Arabidopsis thaliana*) and was found to comprise nine genes divided into five subfamilies, all of whose members contain PLP-binding sites (PXXSVKDR) that are highly conserved across species (Yamaguchi et al., 2000). However, cytosolic *OAS-A1*, plastidial *OAS-B*, and mitochondrial *OAS-C* were identified as the three *OASTLs* that were also expressed at relatively high levels and interacted with ser acetyltransferase (*SAT*) (Bonner et al., 2005; Heeg et al., 2008; Jez and Dey, 2013). In addition, *CYS-D1* and *CYS-D2* also have weak cysteine synthesis functions in mitochondria (Yamaguchi et al., 2000). CSases compose a protein family whose members have multiple functions, and *CSase* genes in different tissue sites may have different functions. For example, L-cysteine desulfhydrase 1 (*DES1*) in the cytoplasm has L-cysteine desulfhydrase activity, sulfocysteine synthase (*SCS)* in the chloroplast encodes S-thiocysteine synthase, and the mitochondrial enzyme *CAS-C1* has β-cyanoalanine synthase (CAS) activity (Alvarez et al., 2010a; Bermúdez et al., 2010). Recently, it has also been shown that *CSase* genes are involved in environmental stress responses such as responses to high-salt conditions and heavy metals, and overexpression of *CSase* genes has been shown to increase the ability of plants to adapt to oxidative stress (Ning et al., 2009; Xie et al., 2012). Moreover, a wide range of defense compounds that can respond to adverse environments use cysteine as a precursor (Alvarez et al., 2010b). Taking the GSH metabolic pathway as an example, the synthesis of GSH as an antioxidant molecule is restricted by cysteine, and in turn GSH is a precursor for the synthesis of phytochelatins (PCs), thiolated peptides involved in the detoxification of heavy metals (Cui et al., 2012, 2014). When plants need to enhance GSH biosynthesis under heavy metal stress, *CSase* can increase cysteine production and subsequently affect the synthesis of downstream substances and achieve improved plant tolerance.

The earliest report of the *CYS-C1* gene in Arabidopsis involved a CAS that converts cyanide and cysteine to β-cyanoalanine and H_2S in mitochondria. *CYS-C1* and *Cys-C* act together to complete the cyclic pathway of cyanide detoxification (García et al., 2010; Álvarez et al., 2012b). However, the activity of *CYS-C1* during cysteine synthesis is also relatively high (Hatzfeld et al., 2000), and *CYS-C1* is considered a member of the CSase gene family (Watanabe et al., 2008). Since *SAT* and *CSase* interact for efficient synthesis of cysteine, authentic *CSase* can interact with *SAT* (Droux et al., 1998; Romero et al., 2014). Moreover, the *SlOAS7* gene in the CYS-C subfamily was found to interact with *SAT* in tomato (Liu et al., 2018). Based on the above information, it is speculated that CYS-C subfamily members may also be true CSases.

Alfalfa (*Medicago sativa* L.), which is widely grown in Asia, Europe, and America, is a high-quality perennial forage plant of the legume family; alfalfa is high yielding and rich in nutrients and is one of the most important forage species for healthy and efficient livestock breeding. The CSase gene family has been extensively studied in many species, and genome-wide analyses have identified members of the CSase gene family in Arabidopsis (Yamaguchi et al., 2000), tomato (*Solanum lycopersicum* L.) (Liu et al., 2018), foxtail millet (*Setaria italica* (L.) P. Beauvois) (Liu et al., 2019), and sorghum (*Sorghum bicolor*) (Akbudak et al., 2018), but our knowledge of the CSase gene family in forage crop species such as alfalfa is still limited. The recently published genome of alfalfa (cultivar Xinjiangdaye) provides an important resource for further molecular studies of this species (Chen et al., 2020). Previous work by our group found that this gene responds to alkali stress (Song et al., 2017, 2021). Based on this information, a total of 39 *CSase* genes were identified and classified into 6 subfamilies in this study, and bioinformatic analyses including phylogenetic analysis, motif composition analysis, and gene duplication analysis were performed to provide a theoretical basis for clarifying the evolutionary history and biological functions of the members of this gene family. In addition, we successfully cloned a *CSase* gene (belonging to the CYS-C1 subfamily) from alfalfa, transferred it into alfalfa, and analyzed its potential function. The results showed that this gene encodes a protein that promotes cysteine synthesis and improves the alkalinity tolerance of overexpression lines by increasing the antioxidant capacity of the plant.

MATERIALS AND METHODS

Identification of Cysteine Synthase Gene Family Members in Alfalfa

The sequences of the nine identified *AtCSase* genes were obtained from the NCBI database.[1] The *M. sativa* Xinjiangdaye genome sequence was downloaded from a website.[2] *MsCSases* were identified by two rounds of BLASTP. A hidden Markov model (HMM) was used by Pfam 31.1[3] to ensure that the PF00291 domain was retained, and DNAMAN was used for sequence comparison searches for the PLP-binding site (PXXSVKDR) in alfalfa.

Phylogenetic Analysis and Multiple Sequence Alignment

A phylogenetic tree was generated by MEGA 5 using the NJ method, with 1,000 bootstrap replicates. Multiple sequence alignments of CSases were created with ClustalX.

Analysis of Conserved Motifs and Conserved Domains

The conserved motif structures within the CSase sequences were identified by MEME Suite Version 5.2.0[4] with the following parameters: zero or one occurrence per sequence of site distribution, a maximum of 10 misfits and a maximum width of motif between 6 and 50. NCBI Batch CD-Search[5] was used to

[1]http://ncbi.nlm.nih.gov/

[2]https://figshare.com/projects/whole_genome_sequencing_and_assembly_of_Medicago_sativa/66380

[3]https://pfam.xfam.org/

[4]http://meme-suite.org/tools/meme

[5]https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi

analyze the conserved domains of the CSase proteins, after which the domains were visualized by TBtools.

Vector Construction and Plant Transformation

We used the *CSase* gene of *M. truncatula* (*Medtr7g078070.1*) as a reference sequence to clone the *CSase* gene of *M. sativa*. Transient expression vectors for tobacco and overexpression vectors for alfalfa were constructed by the one-step cloning method. Then, the expression vectors were transformed into *Agrobacterium rhizogenes* by using the freeze-thaw method. The *CSase* gene was transformed into alfalfa *via* Agrobacterium mediation using the cotyledon method, and regenerated alfalfa plants were obtained. The *bar* gene detection method and fluorescence quantitative analysis technology were used to screen overexpression plants. The sequences of the primers used are shown in **Supplementary Table 1**.

Plant Growth Conditions and Treatments

Nicotiana benthamiana plants were grown in plastic pots filled with vermiculite. Approximately 1-month-old seedlings were used for transient expression. *M. sativa* Longmu 801 was used in this study. Softwood cuttings from the CK line and overexpression lines OV#L11, OV#L12, and OV#L13 were transplanted into plastic pots containing vermiculite (one plant per pot). All the plants were grown under a 16 h light/8 h dark photoperiod under a day/night temperature cycle of 22°C/18°C. Hoagland solution (1/10 strength) was applied to the plants every 3 days.

For $NaHCO_3$ treatment, 150 mM $NaHCO_3$ was applied for 5 d, and a 5 d recovery period was selected as the best condition for identifying stress phenotypes. Samples were taken at 0, 1, 6, 12, 24, 48 h, and 5 d after the beginning of the treatment. Three biological replicates were included per line.

Subcellular Localization Analysis

To explore the subcellular localization of CSase proteins, we constructed a transient expression vector. *CSase* gene was inserted downstream from the double CaMV 35S promoter in the pCAMBIA1300 vector. The pCAMBIA1300 vector carries GFP gene. The sequences of the primers used are listed in **Supplementary Table 1**. The resulting vector was introduced into the *Agrobacterium tumefaciens* strain GV3101. We used a syringe to infiltrate *Agrobacterium tumefaciens* strain GV3101 containing a tobacco transient expression vector into 1-month-old tobacco leaves. After infiltration, the plants were cultivated for 72 h under dark conditions. The fluorescence signal in the infested tobacco leaves was subsequently observed by confocal microscopy.

Determination of Physiological Indicators and Expression Analysis of Related Genes

Physiological traits including GSH, cysteine, Pro, and MDA contents and SOD and POD activities were measured using reagent kits (Nanjing Jiancheng Bioengineering Institute, Nanjing, China). The instructions of the kits were followed for specific test procedures.

For qRT-PCR, total RNA was isolated from alfalfa samples using an RNeasy Plant Mini Kit (CWBIO, Jiangsu, China), and cDNA was synthesized using a kit (Vazyme). qRT-PCR was used to analyze the relative expression levels of alfalfa CSase-responsive genes. The *GADPH* gene was used as a reference. The sequences of the primers used are shown in **Supplementary Table 1**.

Statistical Analyses

To determine significance, all statistical analyses were performed by using Microsoft Excel.

RESULTS

Identification of MsCSase Genes in Alfalfa

First, a total of 39 *MsCSase* gene sequences were retrieved from alfalfa using BLAST, PF00291 domain and PLP-binding site (PXXSVKDR) searches and named *MsCSase01* to *MsCSase39* according to their chromosome locations (**Supplementary Table 2**). They were unevenly mapped onto chromosomes 1, 4, 5, and 7, which contained 14, 9, 7, and 9 genes, respectively (**Supplementary Table 2**). Information about their coding DNA sequences (CDS) and resulting protein sequences are presented in **Supplementary Table 3**. Then, characterization of the proteins revealed that the predicted isoelectric points (pIs) of the MsCSase proteins ranged from 5.17 to 9.08 (**Supplementary Table 2**). Except for *MsCSase01*, *MsCSase31* and *MsCSase36*, the length and molecular mass did not widely vary (**Supplementary Table 2**). The phylogenetic tree results demonstrated that the MsCSase proteins could be divided into 6 subfamilies according to the clades and classification from Arabidopsis, including 14, 9, 5, 4, and 4 members in the CysA subfamily, CysB subfamily, SCS subfamily, CysD subfamily and CysC subfamily, respectively (**Figure 1**). Similar to that which occurred in a study in tomato (Liu et al., 2018), *MsCSase25*, *MsCSase27*, and *MsCSase28* were separated into a separate family and did not belong to the other five subfamilies. These results indicated that the characteristics and patterns of evolution in various species are more likely to differ.

Structural Features and Synteny Analysis

Genetic structural diversity supported the phylogenetic groupings to some extent (Wei et al., 2016). Therefore, we analyzed the relationship between gene structure and phylogenetic clustering to gain insight into the evolution of the MsCSase gene family in alfalfa (**Figures 2A,B**). Gene structure analysis showed that genes within the same subfamily presented similar structures; for instance, the CysA subfamily members contained 10 exons, and the exon distribution of genes on the same branch was largely similar. In addition, the protein motif analysis (the conserved motifs *via* sequence logo are shown in **Supplementary Figure 1**) by MEME found a similar pattern and

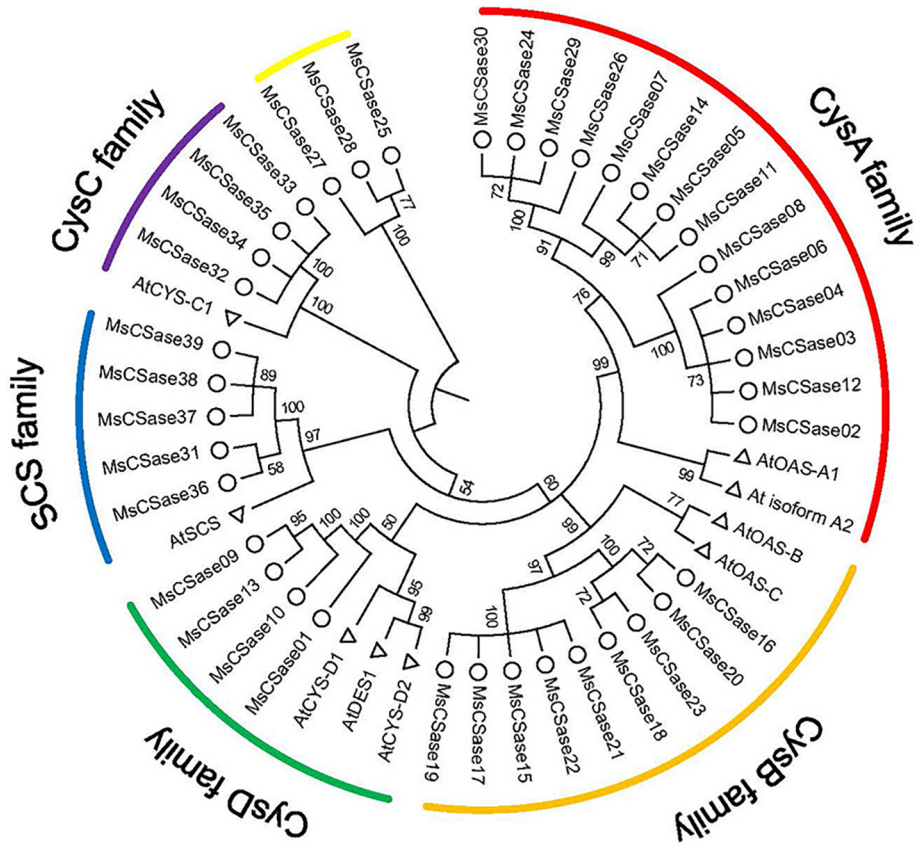


FIGURE 1 | Phylogenetic tree generated using the NJ method. The tree represents relationships among CSases of alfalfa and Arabidopsis. The different colored arcs indicate different CSase subfamilies. Δ and ○ represent *CSase* genes in alfalfa and Arabidopsis, respectively. The number represents the bootstrap value of the node. Larger node values imply a higher confidence level.

the order and distribution of the motifs were roughly similar among the members of the subfamilies (**Figure 2C**). These results supported the close evolutionary relationship of the classification of these MsCSase subfamilies. Moreover, the MsCSase genes had five characterized domains, including PLN02565, the PLN02565 superfamily, Trp-synth-beta_II superfamily and PLN02556 domains, all of which are related to cysteine synthesis (**Figure 2D**). Overall, the domain similarity suggested that these genes may have similar functions, but the differences in the activity of their encoded enzymes may be related to differences in gene structure and motifs between subfamily members.

To elucidate the mechanism through which the MsCSase gene family members in alfalfa expanded, gene duplication events were identified. Two pairs of tandem duplication genes (*MsCSase9/10*, *MsCSase22/23*) and 35 groups of synteny gene pairs in which *MsCSase07/24* were segmentally duplicated genes were identified by TBtools and MCScanX software (**Supplementary Table 4**). Duplicated genes were located on chromosomes 1, 5, and 10 (**Supplementary Figure 2**). Taken together, the results indicated that there was no obvious relationship between chromosome length and the number of genes. Some *MsCSase* genes may have been generated by gene duplications and tandem and segmental duplications contributed to the evolution of *MsCSases* in alfalfa.

Molecular Cloning and Subcellular Localization of CSase

Using the *Medtr7g078070.1* gene of *Medicago truncatula* as a probe, we cloned the gene with accession number MK334208 named *CSase* from alfalfa (cultivar Longmu 801). Sequence analysis showed that the amino acid sequences of *MsCSas32*, *MsCSase34,* and *CSase* were nearly identical (**Supplementary Figure 3**); they were 99.91% similar at the nucleotide level, and only one nucleotide differed between *MsCSase34* and *CSase* (**Supplementary Figure 4**). Combining the results of the phylogenetic evolutionary tree analysis with these results, we determined that the *CSase* gene belonged to the CYS-C1 subfamily and was highly conserved in two different alfalfa varieties.

qRT-PCR analysis of different alfalfa tissue parts revealed that the relative expression of *CSase* was higher in the leaves than in the other tissues (**Figure 3A**). Moreover, the relative expression in mature leaves was much higher than that in other tissues. To further clarify the location of gene activity, we evaluated the subcellular localization of *CSase*. We fused its ORF sequence without the terminal codon to GFP at the N-terminus under the control of the CaMV 35S promoter and ultimately transiently

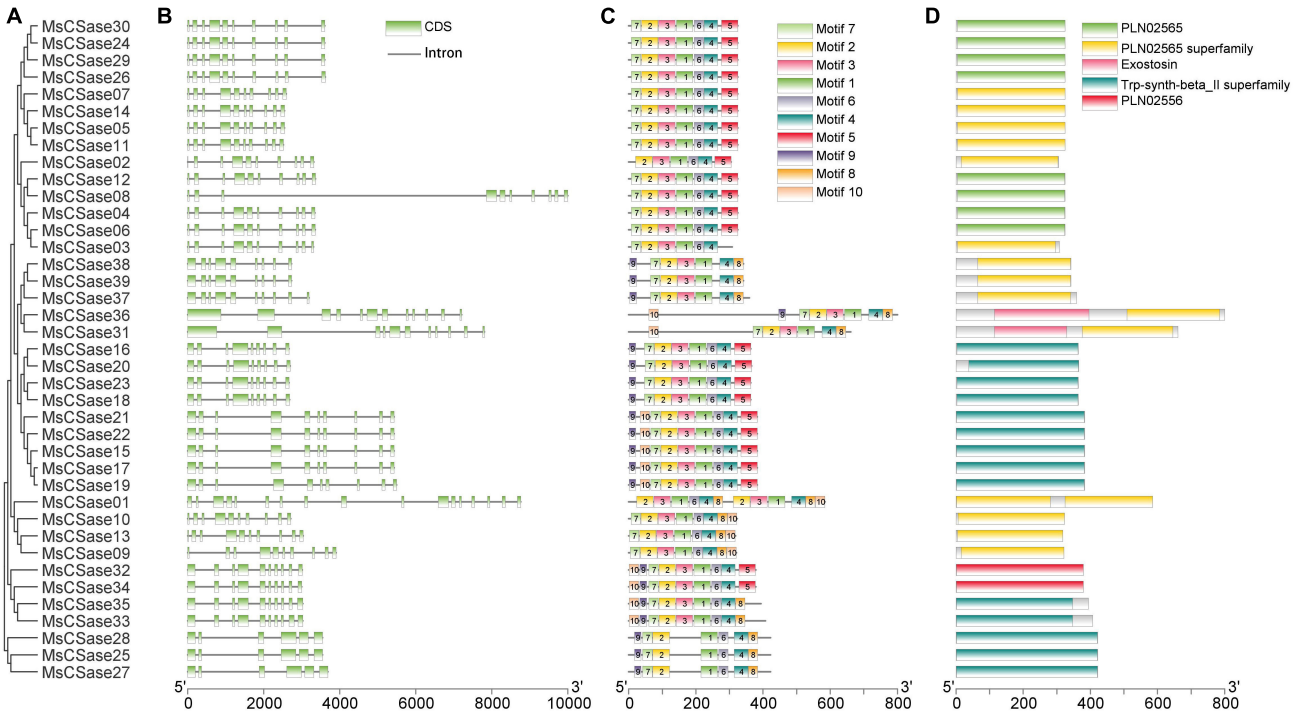


FIGURE 2 | Architecture of conserved protein motifs, domains and gene structures among the MsCSase gene family members in alfalfa. **(A)** Phylogenetic tree of alfalfa MsCSase proteins constructed using MEGA 5 software. **(B)** Gene structure of the MsCSase gene family members. The green boxes indicate exons; the black lines indicate introns. **(C)** Motif composition of MsCSase proteins. The motifs, numbered 1–10, are displayed in different colored boxes. The results of a sequence logo analysis of the conserved motifs are provided in **Supplementary Figure 1**. The length of the protein can be estimated using the scale at the bottom. **(D)** The characterized domains of the *MsCSase* genes were identified. The different colors represent different domains.

expressed it in tobacco (*N. benthamiana*) leaf epidermal cells. The results showed that the presence of the CSase–GFP fusion protein in the chloroplasts of the cells (**Figure 3B**). This finding also directly validates the accuracy of the qRT-PCR results.

Overexpression of Cysteine Synthase Enhances Alkali Tolerance in Transgenic Alfalfa

qRT-PCR analysis under different abiotic stresses revealed that *CSase* could respond positively to salt, alkali and drought stresses in alfalfa, and the response to alkali stress was more pronounced than that to the other stressors (**Figure 4A**). To determine whether alfalfa alkalinity tolerance was altered by up-regulation of *CSase*, we constructed alfalfa overexpression vectors (**Supplementary Figure 5**) and transformed the CDS of *CSase* into the alfalfa cultivar Longmu 801 to obtain *CSase*-overexpressing transgenic lines. The glucosamine gene is a marker gene that is present only in the vector itself and not in alfalfa. *Bar* gene detection revealed successful infestation of alfalfa in response to a bacterial solution (**Supplementary Figure 6**). qRT-PCR analysis showed that the relative expression of the *MsCSase* gene in the control (CK) lines were much lower than that in the OV# L11-, OV# L12-, and OV#L13-overexpressed lines (**Figure 4B**). This indicates that the *CSase* gene was successfully overexpressed in alfalfa. We ultimately selected the CK line and overexpression lines OV#L11, OV#L12, and OV#L13 for subsequent experiments.

Under normal growth conditions, there was no significant difference in phenotype between the CK and overexpression lines, and the differences between the aboveground biomass fresh weight and the relative chlorophyll content were not significant. However, after 5 d of alkali stress, the leaf wilting of the overexpression lines was much lower than that of the CK (**Figure 4C**). The alkali treatment led to degradation of the relative chlorophyll content (**Figure 4D**), but the chlorophyll contents were still higher in the overexpression lines than in the CK line. Moreover, the difference in aboveground biomass between the lines was not significant (**Figure 4E**). Overall, overexpressing *CSase* provided increased tolerance to alkali stress in alfalfa.

To clarify the reason for the increased alkali tolerance of the *CSase* overexpression lines, we also measured the contents of malondialdehyde (MDA), proline (Pro) and soluble sugars (SSs) in the *CSase* overexpression lines and CK plants grown under normal and alkaline conditions. However, after alkali stress, the contents of MDA, Pro and SSs increased within each line, and the increase in MDA content in the overexpression line was significantly lower than that in CK, while the contents of osmoregulatory substances containing Pro and SSs were significantly higher than those in CK (**Figures 5A–C**). It can be hypothesized that, compared with the CK plants, the

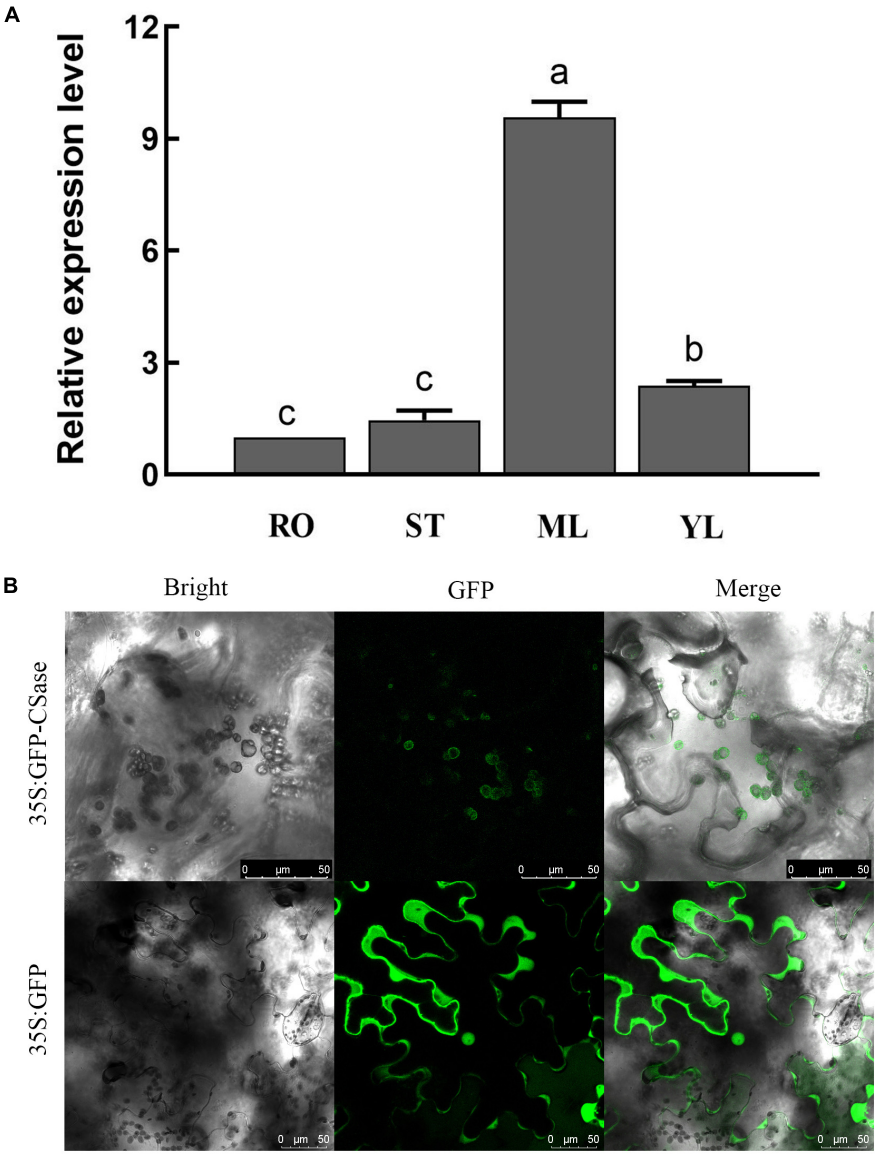


FIGURE 3 | Expression pattern and subcellular localization of CSase. **(A)** Differential expression of representative *CSases* in different tissues according to qRT-PCR. RO, root; ST, stem; ML, mature leaf; and YL, young leaf. **(B)** Confocal laser scanning microscopy images of tobacco leaf cells expressing the CSase protein (35S:GFP-CSase) and GFP protein (35S:GFP). Scale bars = 50 μm.

overexpression lines are alkaline tolerant due to their lower degree of membrane damage and higher accumulation of osmoregulatory substances.

Cysteine Synthase Overexpression Increased the Cysteine and Glutathione Contents to Improve Alkali Tolerance

CSases have been shown to encode CSase proteins (Hatzfeld et al., 2000). The increase in CSase content promotes an increase in cysteine content in plants. In the present study, the cysteine content in the overexpression lines was significantly higher than that in CK under normal growth conditions, and the cysteine content under alkali stress was higher than that in the CK and tended to increase (**Figure 6A**). Accordingly, we speculated that cysteine plays a role in improving alkali tolerance in plants.

Based on our preliminary research, we found that cysteine is a precursor for the synthesis of antioxidant substances such as GSH (Droux, 2004; Wirtz and Droux, 2005; Kopriva, 2006) and that GSH plays an important role in enhancing antioxidant capacity and transducing redox-sensitive signals in plants (Cnubben et al., 2001; Matés et al., 2002; Pastori and Foyer, 2002). Therefore, we determined the GSH content and the relative expression of γ-glutamyl cysteine synthetase (*γ-ECS*) in each line. The results showed that the GSH content in the plants increased in response to alkali stress and that the overexpression lines contained more

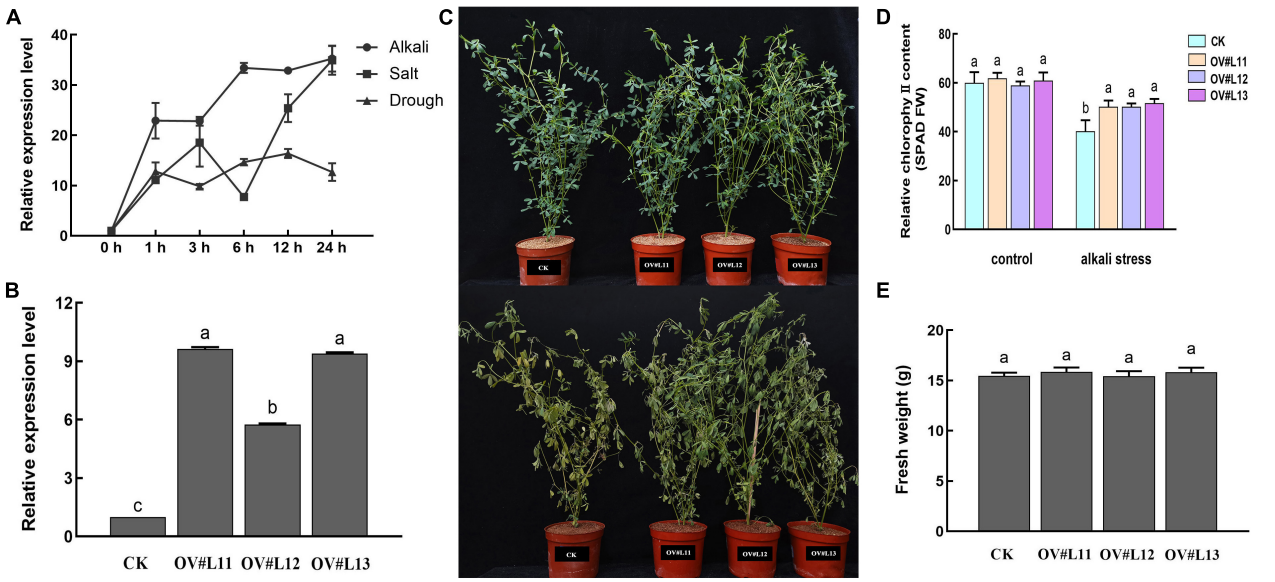


FIGURE 4 | Overexpression of *CSase* confers alkalinity tolerance to alfalfa. **(A)** Relative expression of *CSase* under salt, alkali and drought stress. **(B)** Relative expression of *Bar* in different lines. **(C)** Phenotypic plots of different lines under normal growth and alkali conditions. **(D)** Tissue SPAD value of different lines under normal and alkali stress conditions. **(E)** Fresh weight of each line under normal conditions. The average of three independent samples was calculated and is shown, together with the standard error of the mean (bars) between samples. Different lowercase letters under the same conditions indicate that the difference is significant ($P < 0.05$).

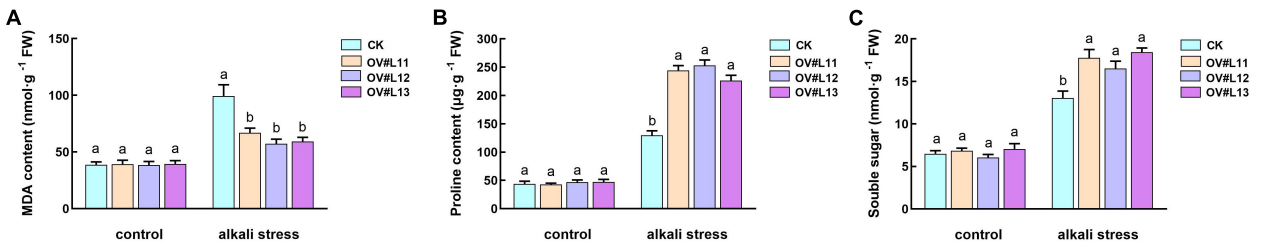


FIGURE 5 | Contents of MDA **(A)**, Pro **(B)**, and SSs **(C)** in each line under normal conditions and after alkali stress. One-month-old alfalfa seedlings were subjected to 150 mM $NaHCO_3$ solution for 5 days. The average of three independent samples was calculated and is shown, together with the standard error of the mean (bars) between samples. Different lowercase letters under the same conditions indicate that the difference is significant ($P < 0.05$).

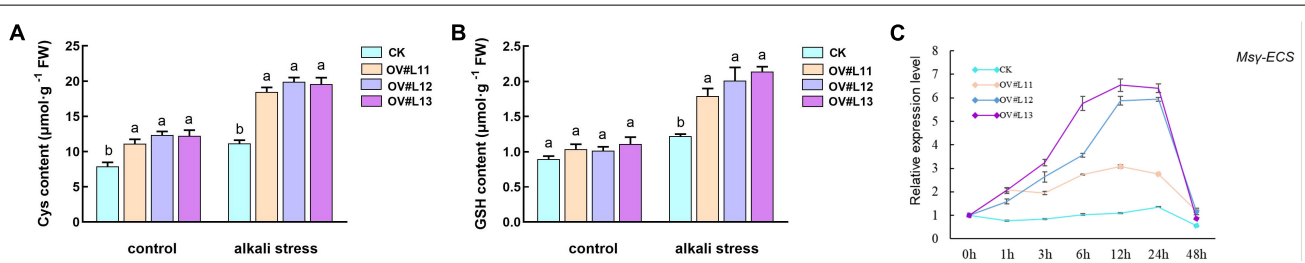


FIGURE 6 | Contents of cysteine **(A)** and GSH **(B)** and the relative expression of the *Msγ-ECS* gene **(C)** in each line under normal conditions and after alkali stress. One-month-old alfalfa seedlings were subjected to 150 mM $NaHCO_3$ solution for 5 days. The average of three independent samples was calculated and is shown, together with the standard error of the mean (bars) between samples. Different lowercase letters under the same conditions indicate that the difference is significant ($P < 0.05$).

GSH (**Figure 6B**). The trends of the relative expression of γ-*ECS* in the overexpression and CK lines were similar, but the relative expression of γ-*ECS* was higher in the overexpression line than in the CK line (**Figure 6C**). Overexpression of *CSase* and *SAT* in tobacco also significantly increased the relative expression of γ-*ECSs* as well as the GSH content in the plants

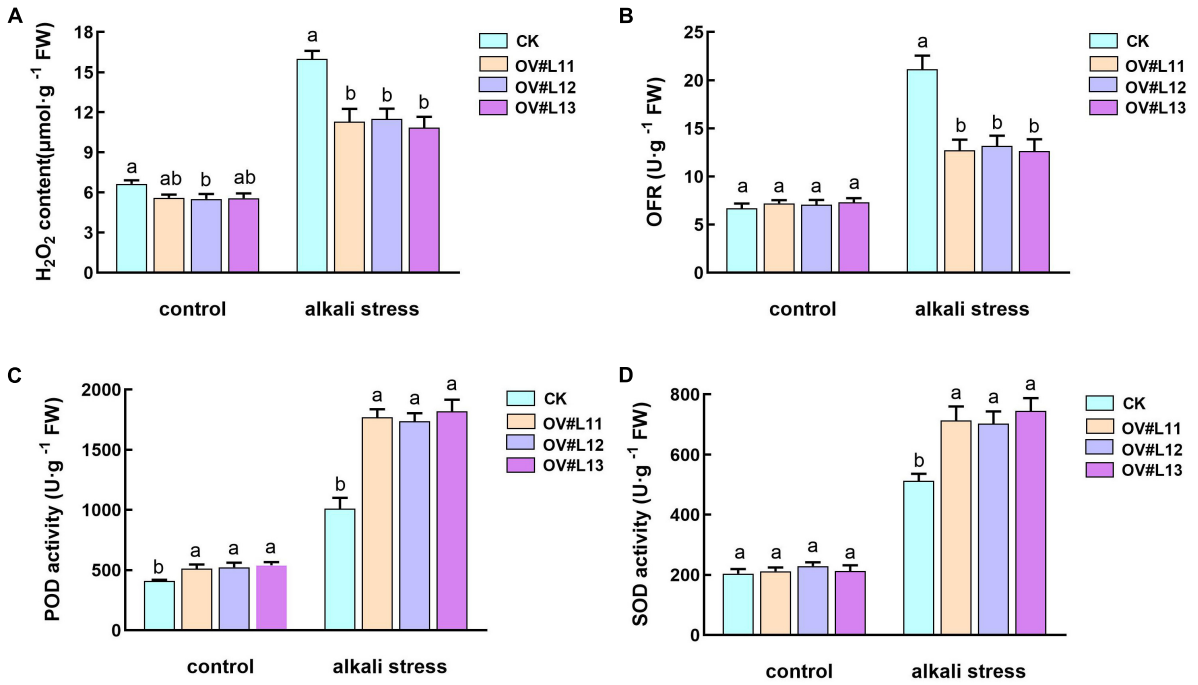


FIGURE 7 | Content of H_2O_2 **(A)** and OFR **(B)** and the activities of POD **(C)** and SOD **(D)** in each line under normal conditions and after alkali stress. One-month-old alfalfa seedlings were subjected to 150 mM $NaHCO_3$ solution for 5 days. The average of three independent samples was calculated and is shown, together with the standard error of the mean (bars) between samples. Different lowercase letters under the same conditions indicate that the difference is significant ($P < 0.05$).

(Nakamura et al., 2014). Thus, the overexpression of the *CSase* gene could regulate downstream metabolic pathways, which led to an increase in the relative expression of the downstream γ-*ECS* gene, and the increase in cysteine content provided the possibility of an increase in GSH content. This also laid the foundation for the improvement in alkali tolerance in the overexpression line.

Cysteine Synthase Overexpression Increased the Antioxidant Capacity of Transgenic Alfalfa

Studies on *DES1* and *OAS-A1* in Arabidopsis showed that cysteine is an important determinant of antioxidant capacity in the cytoplasm (Loìpez-Martiìn et al., 2008; Alvarez et al., 2010a), while *SCS* plays an important role in chloroplast redox (Bermúdez et al., 2010). In addition, cysteine and GSH are also associated with plant antioxidants. Therefore, we measured the superoxide anion (OFR) and hydrogen peroxide (H_2O_2) contents and the enzymatic activity of several antioxidant enzymes. Under normal growth conditions, although the H_2O_2 and OFR contents in the overexpression line were lower than those in the CK line, only the H_2O_2 content differed significantly (**Figures 7A,B**). The levels in both lines increased significantly after alkali stress, but the levels in the overexpression lines were significantly lower than those in the CK line. Thus, we speculated that the overexpression lines may be superior to the CK line in terms of the scavenging ability of reactive oxygen species (ROS).

Based on the above speculation, we also measured the activities of peroxidase (POD) and superoxide dismutase (SOD). The results showed that the activities of both enzymes were much higher in the overexpression lines than in the CK plants under alkali stress, and the activities of POD in the former were also significantly higher than those in the CK plants under normal growth conditions (**Figures 7C,D**). In conclusion, it was clear that the overexpression lines could reduce the accumulation of ROS by increasing the activity of antioxidant enzymes under alkali stress to enhance the alkali tolerance of the overexpression lines.

DISCUSSION

The CSase gene family is widespread in plants, and its members play an important role in cysteine synthesis, cyanide metabolism and other pathways. The synthesis of cysteine is the main function of the members of this gene family, and the CSase they encode is involved in the final step of cysteine synthesis. Cysteine is the first reduced sulfur donor organic molecule synthesized in plants and is involved in the synthesis of a variety of compounds involved in defense, redox, and other functions and occupies a central position in plant metabolism. Most studies on the classification and functional determination of the CSase gene family have focused on Arabidopsis (Barroso et al., 1995; Domínguez-Solís et al., 2004; Alvarez et al., 2010a; Álvarez et al., 2012a; Birke et al., 2012). However, studies on the identification, classification and related functions of CSase gene family members in specific species are lacking. For this reason, we conducted the present study on the MsCSase gene family in alfalfa.

The 39 *MsCSase* genes that were identified were unevenly distributed across only a few chromosomes. Although previous studies have also shown that *CSase* genes are unevenly distributed on only a few specific chromosomes, they have not shown that the genes are subject to fragment duplication or tandem duplication events (Yamaguchi et al., 2000; Akbudak et al., 2018; Liu et al., 2018, 2019). However, we identified two pairs of tandem repeat genes and one pair of gene-generating fragment duplication genes in alfalfa (Bennetzen et al., 2005; Zhu et al., 2014). These results suggested that tandem and fragment duplication events play a role in expansion of the alfalfa MsCSase gene family.

Based on phylogenetic analysis and previous Arabidopsis studies, we divided the alfalfa MsCSase gene family members (excluding *MsCSase25*, *MsCSase27*, and *MsCSase28*) into 5 subfamilies. The gene structures, motifs and domains of each subfamily member were somewhat similar. The results of the structural analysis support the reliability of the phylogenetic analysis. Differences in the physicochemical properties and structures of different subfamilies may result in differences in the activity of the enzymes or the function of the genes (Liszewska et al., 2007; Noda et al., 2016). The identified *MsCSase25*, *MsCSase27*, and *MsCSase28,* which do not belong to other subfamilies, although they have PLP-binding sites, still need to be verified whether they encode active cysteine synthases.

In terms of gene function validation, we successfully cloned the *CSase* gene (*MsCSase32*, *MsCSase34*) and produced *CSase*-overexpressing alfalfa. Stress tests showed that overexpression of *CSase* significantly improved alkali tolerance in alfalfa. Alkali stress signals induce the biosynthesis and accumulation of compatible osmotic solutes, including SSs and Pro, to improve tolerance. When the plants were subjected to alkali stress, compared with the CK plants, the overexpression lines accumulated more of these substances to reduce the intracellular osmotic potential, and the tolerance of the overexpression plants was improved by the accumulation of these substances.

CSase activity and the cysteine content increase under metal and salt stresses and that overexpression of *CSase* improves the antioxidant capacity and tolerance of plants (Domìnguez-Solìs et al., 2001; Youssefian et al., 2001; Domínguez-Solís et al., 2004; Fediuc et al., 2005; Pajuelo et al., 2007; Gotor et al., 2015). GSH content increases in overexpression lines under environmental stresses (Sabetta et al., 2017) and GSH plays important roles in scavenging ROS and transducing stress signals (Meyer, 2008; Zagorchev et al., 2013; Leng et al., 2015). In addition, redox- and ROS-dependent regulatory networks are important for photosynthesis in chloroplasts (Strand et al., 2015; Gütle et al., 2016). In the present study, we found that the increased resistance of the transgenic lines are due to the overexpression of the *CSase* gene leading to an increase in the content of antioxidants such as cysteine and downstream GSH, which in turn leads to enhanced antioxidant capacity. This was also evidenced by the increase in SOD and POD activities and the decrease in H_2O_2 and OFR contents in overexpression plants under alkali stress. Moreover, the enhanced antioxidant capacity of the transgenic lines may make their photosynthesis less affected by alkali stress and subsequently have higher SPAD values. The enhanced stress resistance resulted in good phenotypes of overexpression lines under alkali stress. In conclusion, CSase overexpression lines enhance plant tolerance by increasing the antioxidant capacity of plants. This conclusion is also consistent with the results of the previous group, which showed that *CSase* can respond to alkali stress and that increased cysteine content can improve the antioxidant capacity of the plants (Song et al., 2017, 2021).

Interestingly, numerous studies have shown that *CYS-C1* has dual functions in synthesizing cysteine and β-cyanoalanine (Ikegami et al., 1988; Watanabe et al., 2008; Marrero-Degro et al., 2010). However, *CYS-C1* currently synthesizes CAS only in mitochondria. The encoded *CSase* product (CAS) is localized in the mitochondria of Arabidopsis and is involved in the detoxification of HCN (Álvarez et al., 2012b). In the present study, this protein was localized in chloroplasts and enhanced the alkalinity tolerance of plants by synthesizing cysteine. From this, we hypothesized that *CSase* could have the ability to encode both CSase and CAS but selectively encodes one of the enzymes depending on the expression location. The above hypothesis needs to be further investigated. Whether *CSase* in alfalfa can also encode CAS also needs to be further investigated.

CONCLUSION

In this study, we focused on 39 alfalfa MsCSase family members and classified them into 6 subfamilies first on the basis of the results of a phylogenetic analysis followed by a gene structure analysis, conserved domain characterization and a synteny analysis and on the basis of the high similarity in these aspects of members within the same subfamily. Subsequently, we cloned *CSase* and successfully overexpressed it in alfalfa. Evidence from both physiological experiments and the determination of the relative expression of downstream genes indicated that the overexpression lines can significantly improve alkali stress tolerance in alfalfa by increasing oxidative stress protection and the levels of osmoregulatory substances. These findings set the stage for the study of the CSase gene family. We will focus our future work on the associated metabolic pathways to further clarify the molecular mechanism of basal tolerance.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

GC and YY designed the experiments. YY wrote the first draft of the article and GC revised it. All authors participated in the experiments and read and approved the final manuscript.

FUNDING

This study was supported in part by the National Natural Science Foundation of China (31872998).

ACKNOWLEDGMENTS

The authors thank to lab members for assistance. And we extend our thanks to the reviewers and editor for their careful reading and helpful comments on this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.792862/full#supplementary-material

REFERENCES

Akbudak, M. A., Filiz, E., and Uylas, S. (2018). Identification of O-acetylserine(thiol)lyase (OASTL) genes in sorghum (*Sorghum bicolor*) and gene expression analysis under cadmium stress. *Mol. Biol. Rep.* 46, 343–354. doi: 10.1007/s11033-018-4477-0

Alvarez, C., Calo, L., Romero, L. C., García, I., and Gotor, C. (2010a). An O-acetylserine(thiol)lyase homolog with l-cysteine desulfhydrase activity regulates cysteine homeostasis in *Arabidopsis*. *Plant Physiol.* 152, 656–669. doi: 10.1104/pp.109.147975

Álvarez, C., García, I., Moreno, I., Pérez-Pérez, M. E., Crespo, J. L., Romero, L. C., et al. (2012a). Cysteine-generated sulfide in the cytosol negatively regulates autophagy and modulates the transcriptional profile in *Arabidopsis*. *Plant Cell* 24, 4621–4634. doi: 10.1105/tpc.112.105403

Álvarez, C., García, I., Romero, L. C., and Gotor, C. (2012b). Mitochondrial sulfide detoxification requires a functional isoform o-acetylserine(thiol)lyase c in *Arabidopsis thaliana*. *Mol. Plant* 5, 1217–1226.

Alvarez, C., Lozano-Juste, J., Romero, L. C., García, I., Gotor, C., León, J., et al. (2010b). Inhibition of *Arabidopsis* O-acetylserine(thiol)lyase A1 by tyrosine nitration. *J. Biol. Chem.* 286, 578–586.

Barroso, C., Vega, J., and Gotor, C. (1995). A new member of the cytosolic O-acetylserine(thiol)lyase gene family in *Arabidopsis thaliana*. *FEBS Lett.* 363:1.

Bennetzen, J., Yu, J., Wang, J., Lin, W., Li, S., Li, H., et al. (2005). The genomes of *Oryza sativa*: a history of duplications. *PLoS Biol.* 3:e38. doi: 10.1371/journal.pbio.0030038

Bermúdez, M. A., Páez-Ochoa, M. A., Gotor, C., and Romero, L. C. (2010). Arabidopsis S-sulfocysteine synthase activity is essential for chloroplast function and long-day light-dependent redox control. *Plant Cell* 22, 403–416. doi: 10.1105/tpc.109.071985

Birke, H., Haas, Florian, H., De Kok, Luit, J., Balk, J., et al. (2012). Cysteine biosynthesis, in concert with a novel mechanism, contributes to sulfide detoxification in mitochondria of *Arabidopsis thaliana*. *Biochem. J.* 445, 275–283. doi: 10.1042/bj20120038

Bonner, E. R., Cahoon, R. E., Knapke, S. M., and Jez, J. M. (2005). Molecular basis of cysteine biosynthesis in plants. *J. Biol. Chem.* 280, 38803–38813. doi: 10.1074/jbc.M505313200

Chen, H., Zeng, Y., Yang, Y., Huang, L., Tang, B., Zhang, H., et al. (2020). Allele-aware chromosome-level genome assembly and efficient transgene-free genome editing for the autotetraploid cultivated alfalfa. *Nat. Commun.* 11:2494. doi: 10.1038/s41467-020-16338-x

Cnubben, N., Rietjens, I., Wortelboer, H., Zanden, J. V., and Bladeren, P. (2001). The interplay of glutathione-related processes in antioxidant defense. *Environ. Toxicol. Pharmacol.* 10, 141–152.

Cui, W., Chen, H., Zhu, K., Jin, Q., Xie, Y., Jin, C., et al. (2014). Cadmium-induced hydrogen sulfide synthesis is involved in cadmium tolerance in *Medicago sativa* by reestablishment of reduced (Homo)glutathione and reactive oxygen species Homeostases. *PLoS One* 9:e109669.

Cui, W., Li, L., Gao, Z., H, Wu, Xie, Y., and Shen, W. (2012). Haem oxygenase-1 is involved in salicylic acid-induced alleviation of oxidative stress due to cadmium stress in *Medicago sativa*. *J. Exp. Bot.* 63, 5521–5534.

Davidian, J.-C., and Kopriva, S. (2010). Regulation of sulfate uptake and assimilation – the same or not the same? *Mol. Plant* 3, 314–325. doi: 10.1093/mp/ssq001

Domìnguez-Solìs, J. R., Gutiérrez-Alcalá, G., Romero, L. C., and Gotor, C. (2001). The cytosolic O-acetylserine(thiol)lyase gene is regulated by heavy metals and can function in cadmium tolerance. *J. Biol. Chem.* 276, 9297–9302. doi: 10.1074/jbc.M009574200

Domínguez-Solís, J. R., López-Martín, M. C., Ager, F. J., Ynsa, M. D., Romero, L. C., and Gotor, C. (2004). Increased cysteine availability is essential for cadmium tolerance and accumulation in *Arabidopsis thaliana*. *Plant Biotechnol. J.* 2, 469–476. doi: 10.1111/j.1467-7652.2004.00092.x

Droux, M. (2004). Sulfur assimilation and the role of sulfur in plant metabolism: a survey. *Photosynthesis Res.* 79, 331–348. doi: 10.1023/b:pres.0000017196.95499.11

Droux, M., Ruffet, M. L., Douce, R., and Job, D. (1998). Interactions between serine acetyltransferase and O-acetylserine (thiol) lyase in higher plants–structural and kinetic properties of the free and bound enzymes. *Eur. J. Biochem.* 255, 235–245.

Fediuc, E., Lips, S. H., and Erdei, L. (2005). O-acetylserine (thiol) lyase activity in Phragmites and Typha plants under cadmium and NaCl stress conditions and the involvement of ABA in the stress response. *J. Plant Physiol.* 162, 865–872. doi: 10.1016/j.jplph.2004.11.015

García, I., Castellano, J. M., Vioque, B., Solano, R., Gotor, C., and Romero, L. C. (2010). Mitochondrial β-cyanoalanine synthase is essential for root hair formation in *Arabidopsis thaliana*. *Plant Cell* 22, 3268–3279. doi: 10.1105/tpc.110.076828

Gotor, C., Laureano-Marín, A. M., Moreno, I., Aroca, Á., García, I., and Romero, L. C. (2015). Signaling in the plant cytosol: cysteine or sulfide? *Amino Acids* 47, 2155–2164.

Gütle, D. D., Roret, T., Müller, S. J., Couturier, J., Lemaire, S. D., Hecker, A., et al. (2016). Chloroplast FBPase and SBPase are thioredoxin-linked enzymes with similar architecture but different evolutionary histories. *Proc. Natl. Acad. Sci. U.S.A.* 113, 6779–6784. doi: 10.1073/pnas.1606241113

Hatzfeld, Y., Maruyama, A., Schmidt, A., Noji, M., Ishizawa, K., and Saito, K. (2000). β-cyanoalanine synthase is a mitochondrial cysteine synthase-like protein in spinach and Arabidopsis1. *Plant Physiol.* 123, 1163–1172. doi: 10.1104/pp.123.3.1163

Heeg, C., Kruse, C., Jost, R., Gutensohn, M., Ruppert, T., Wirtz, M., et al. (2008). Analysis of the *Arabidopsis* O-Acetylserine(thiol)lyase gene family demonstrates compartment-specific differences in the regulation of cysteine synthesis. *Plant Cell* 20, 168–185. doi: 10.1105/tpc.107.056747

Ikegami, F., Takayama, K., and Murakoshi, I. (1988). Purification and properties of β-cyano-l-alanine synthase fromVicia angustifolia. *Phytochemistry* 27, 2011–2016.

Jez, J. M., and Dey, S. (2013). The cysteine regulatory complex from plants and microbes: what was old is new again. *Curr. Opin. Struct. Biol.* 23, 302–310. doi: 10.1016/j.sbi.2013.02.011

Kopriva, S. (2006). Regulation of sulfate assimilation in *Arabidopsis* and beyond. *Ann. Bot.* 97, 479–495. doi: 10.1093/aob/mcl006

Leng, X., Jia, H., Sun, X., Shangguan, L., Mu, Q., Wang, B., et al. (2015). Comparative transcriptome analysis of grapevine in response to copper stress. *Sci. Rep.* 5:17749. doi: 10.1038/srep17749

Liszewska, F., Lewandowska, M., Pochocka, D., and Sirko, A. (2007). Mutational analysis of O-acetylserine (thiol) lyase conducted in yeast two-hybrid system. *Biochim. Biophys. Acta* 1774, 450–455.

Liu, D., Li, J., Lu, J., Tian, B., Liu, X., Yang, G., et al. (2019). Cloning and functional analysis of four O-Acetylserine (thiol) lyase family genes from foxtail millet. *Plant Physiol. Biochem.* 139, 325–332. doi: 10.1016/j.plaphy.2019.03.032

Liu, D., Lu, J., Li, H., Wang, J., and Pei, Y. (2018). Characterization of the O-acetylserine(thiol)lyase gene family in *Solanum lycopersicum* L. *Plant Mol. Biol.* 99, 123–134. doi: 10.1007/s11103-018-0807-9

Loìpez-Martiìn, M. C., Becana, M., Romero, L. C., and Gotor, C. (2008). Knocking out cytosolic cysteine synthesis compromises the antioxidant capacity of the cytosol to maintain discrete concentrations of hydrogen peroxide in *Arabidopsis*. *Physiology* 147, 562–572. doi: 10.1104/pp.108.117408

Marrero-Degro, J., Marcano-Velázquez, J., and Siritunga, D. (2010). Isolation and characterization of novel β-cyanoalanine synthase and cysteine synthase genes from cassava. *Plant Mol. Biol. Rep.* 29, 514–524. doi: 10.1007/s11105-010-0255-4

Matés, J. M., Pérez-Gómez, C., Núñez de Castro, I., Asenjo, M., and Márquez, J. (2002). Glutamine and its relationship with intracellular redox status, oxidative stress and cell proliferation/death. *Int. J. Biochem. Cell Biol.* 34, 439–458.

Meyer, A. J. (2008). The integration of glutathione homeostasis and redox signaling. *J. Plant Physiol.* 165, 1390–1403. doi: 10.1016/j.jplph.2007.10.015

Nakamura, M., Ochiai, T., Noji, M., Ogura, Y., and Saito, K. (2014). An improved tolerance to cadmium by overexpression of two genes for cysteine synthesis in tobacco. *Plant Biotechnol.* 31, 141–147.

Ning, H., Zhang, C., Yao, Y., and Yu, D. (2009). Overexpression of a soybean O-acetylserine (thiol) lyase-encoding gene GmOASTL4 in tobacco increases cysteine levels and enhances tolerance to cadmium stress. *Biotechnol. Lett.* 32, 557–564. doi: 10.1007/s10529-009-0178-z

Noda, M., Nakamura, M., Takamiya, R., Tamura, T., Ito, T., and Kodama, H. (2016). A spinach O-acetylserine(thiol)lyase homologue, SoCSaseLP, suppresses cysteine biosynthesis catalysed by other enzyme isoforms. *Biochim. Open* 2, 24–32.

Pajuelo, E., Carrasco, J. A., Romero, L. C., Chamber, M. A., and Gotor, C. (2007). Evaluation of the metal phytoextraction potential of crop legumes. Regulation of the expression of O-Acetylserine (Thiol)lyase under metal stress. *Plant Biol.* 9, 672–681. doi: 10.1055/s-2007-965439

Pastori, G. M., and Foyer, C. H. (2002). common components, networks, and pathways of cross-tolerance to stress. The central role of "redox" and abscisic acid-mediated controls. *Plant Physiol.* 129, 460–468. doi: 10.1104/pp.011021

Romero, L. C., Aroca, M. A., Laureano-Marín, A., Moreno, I., García, I., and Gotor, C. (2014). Cysteine and cysteine-related signaling pathways in *Arabidopsis thaliana*. *Mol. Plant* 7, 264–276.

Sabetta, W., Paradiso, A., Paciolla, C., and de Pinto, M. C. (2017). "Chemistry, biosynthesis, and antioxidative function of glutathione in plants," in *Glutathione in Plant Growth, Development, and Stress Tolerance*, eds M. Hossain, M. Mostofa, P. Diaz-Vivancos, D. Burritt, M. Fujita, and L. S. Tran (Cham: Springer), 1–27. doi: 10.1007/978-3-319-66682-2_1

Song, T., Sun, N., Dong, L., and Cai, H. (2021). Enhanced alkali tolerance of rhizobia-inoculated alfalfa correlates with altered proteins and metabolic processes as well as decreased oxidative damage. *Plant Physiol. Biochem.* 159, 301–311.

Song, T., Xu, H., Sun, N., Jiang, L., Tian, P., Yong, Y., et al. (2017). Metabolomic analysis of alfalfa (*Medicago sativa* L.) root-symbiotic rhizobia responses under alkali stress. *Front. Plant Sci.* 8:1208. doi: 10.3389/fpls.2017.01208

Strand, D. D., Livingston, A. K., Satoh-Cruz, M., Froehlich, J. E., Maurino, V. G., and Kramer, D. M. (2015). Activation of cyclic electron flow by hydrogen peroxide in vivo. *Proc. Natl. Acad. Sci. U.S.A.* 112, 5539–5544. doi: 10.1073/pnas.1418223112

Takahashi, H., Kopriva, S., Giordano, M., Saito, K., and Hell, R. (2011). Sulfur assimilation in photosynthetic organisms: molecular functions and regulations of transporters and assimilatory enzymes. *Annu. Rev. Plant Biol.* 62, 157–184. doi: 10.1146/annurev-arplant-042110-103921

Van Hoewyk, D., Pilon, M., and Pilon-Smits, E. A. H. (2008). The functions of NifS-like proteins in plant sulfur and selenium metabolism. *Plant Sci.* 174, 117–123. doi: 10.1016/j.plantsci.2007.10.004

Watanabe, M., Kusano, M., Oikawa, A., Fukushima, A., Noji, M., and Saito, K. (2008). Physiological roles of the β-substituted alanine synthase gene family in *Arabidopsis*. *Plant Physiol.* 146, 310–320. doi: 10.1104/pp.107.106831

Wei, Y., Shi, H., Xia, Z., Tie, W., Ding, Z., Yan, Y., et al. (2016). Genome-wide identification and expression analysis of the WRKY gene family in cassava. *Front. Plant Sci.* 7:25. doi: 10.3389/fpls.2016.00025

Wirtz, M., and Droux, M. (2005). Synthesis of the sulfur amino acids: cysteine and methionine. *Photosynthesis Res.* 86, 345–362. doi: 10.1007/s11120-005-8810-9

Wirtz, M., and Hell, R. (2006). Functional analysis of the cysteine synthase protein complex from plants: structural, biochemical and regulatory properties. *J. Plant Physiol.* 163, 273–286. doi: 10.1016/j.jplph.2005.11.013

Xie, Y., Lai, D., Mao, Y., Zhang, W., Shen, W., and Guan, R. (2012). Molecular cloning, characterization, and expression analysis of a novel gene encoding l-cysteine desulfhydrase from *Brassica napus*. *Mol. Biotechnol.* 54, 737–746. doi: 10.1007/s12033-012-9621-9

Yamaguchi, Y., Nakamura, T., Kusano, T., and Sano, H. (2000). Three *Arabidopsis* genes encoding proteins with differential activities for cysteine synthase and β-cyanoalanine synthase. *Plant Cell Physiol.* 41, 465–476.

Youssefian, S., Nakamura, M., Orudgev, E., and Kondo, N. (2001). Increased cysteine biosynthesis capacity of transgenic tobacco overexpressing an O-acetylserine(thiol) lyase modifies plant responses to oxidative stress. *Plant Physiol.* 126, 1001–1011.

Zagorchev, L., Seal, C., Kranner, I., and Odjakova, M. (2013). A central role for thiols in plant tolerance to abiotic stress. *Int. J. Mol. Sci.* 14, 7405–7432. doi: 10.3390/ijms14047405

Zhu, Y., Wu, N., Song, W., Yin, G., Qin, Y., and Yan, Y. (2014). Soybean (*Glycine max*) expansin gene superfamily origins: segmental and tandem duplication events followed by divergent selection among subfamilies. *BMC Plant Biol.* 14:93.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 28 January 2022
doi: 10.3389/fpls.2021.806949

Genome-Wide Association Study Reveals Complex Genetic Architecture of Cadmium and Mercury Accumulation and Tolerance Traits in *Medicago truncatula*

Timothy Paape[1], Benjamin Heiniger[2], Miguel Santo Domingo[3†], Michael R. Clear[1], M. Mercedes Lucas[3] and José J. Pueyo[3]*

[1] Brookhaven National Laboratory, Upton, NY, United States, [2] Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland, [3] Department of Soil, Plant and Environmental Quality, Institute of Agricultural Sciences, ICA-CSIC, Madrid, Spain

OPEN ACCESS

Edited by:
Rajeev K. Varshney, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), India

Reviewed by:
Nevin D. Young, University of Minnesota Twin Cities, United States
Xin-Yuan Huang, Nanjing Agricultural University, China

***Correspondence:**
Timothy Paape
tpaape@bnl.gov

†Present address:
Miguel Santo Domingo, Centre for Research in Agricultural Genomics, CRAG (CSIC-IRTA-UAB-UB), Barcelona, Spain

Specialty section:
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

Received: *01 November 2021*
Accepted: *13 December 2021*
Published: *28 January 2022*

Citation:
Paape T, Heiniger B, Santo Domingo M, Clear MR, Lucas MM and Pueyo JJ (2022) Genome-Wide Association Study Reveals Complex Genetic Architecture of Cadmium and Mercury Accumulation and Tolerance Traits in Medicago truncatula. Front. Plant Sci. 12:806949. doi: 10.3389/fpls.2021.806949

Heavy metals are an increasing problem due to contamination from human sources that and can enter the food chain by being taken up by plants. Understanding the genetic basis of accumulation and tolerance in plants is important for reducing the uptake of toxic metals in crops and crop relatives, as well as for removing heavy metals from soils by means of phytoremediation. Following exposure of *Medicago truncatula* seedlings to cadmium (Cd) and mercury (Hg), we conducted a genome-wide association study using relative root growth (RRG) and leaf accumulation measurements. Cd and Hg accumulation and RRG had heritability ranging 0.44 – 0.72 indicating high genetic diversity for these traits. The Cd and Hg trait associations were broadly distributed throughout the genome, indicated the traits are polygenic and involve several quantitative loci. For all traits, candidate genes included several membrane associated ATP-binding cassette transporters, *P*-type ATPase transporters, oxidative stress response genes, and stress related UDP-glycosyltransferases. The *P*-type ATPase transporters and ATP-binding cassette protein-families have roles in vacuole transport of heavy metals, and our findings support their wide use in physiological plant responses to heavy metals and abiotic stresses. We also found associations between Cd RRG with the genes *CAX3* and *PDR3,* two linked adjacent genes, and leaf accumulation of Hg associated with the genes *NRAMP6* and *CAX9*. When plant genotypes with the most extreme phenotypes were compared, we found significant divergence in genomic regions using population genomics methods that contained metal transport and stress response gene ontologies. Several of these genomic regions show high linkage disequilibrium (LD) among candidate genes suggesting they have evolved together. Minor allele frequency (MAF) and effect size of the most significant SNPs was negatively correlated with large effect alleles being most rare. This is consistent with purifying selection against alleles that increase toxicity and abiotic stress.

Conversely, the alleles with large affect that had higher frequencies that were associated with the exclusion of Cd and Hg. Overall, macroevolutionary conservation of heavy metal and stress response genes is important for improvement of forage crops by harnessing wild genetic variants in gene banks such as the Medicago HapMap collection.

Keywords: cadmium, mercury, polygenic, standing variation, genetic architecture, *Medicago truncatula* (Medicago)

INTRODUCTION

Heavy metals are high-density elements that can cause toxic effects when present in excess quantities. Cadmium (Cd) and mercury (Hg) are two of the most toxic heavy metals to humans as Cd poisoning can cause kidney damage and osteoporosis (Järup and Åkesson, 2009), while Hg poisoning is associated with lung, kidney, muscle and brain damage (Vallee and Ulmer, 1972; Bernhoft, 2012). Heavy metals may occur naturally at low concentrations in soils, often originating from volcanic soils and weathered rocks. However, the predominant source of heavy metal contamination are mines, foundries, and smelters, which are often associated with particularly high levels of contamination in surrounding soils (Tchounwou et al., 2012; Alloway, 2013). This pollution can negatively affect surrounding agricultural and natural ecosystems. Heavy metal accumulation in agricultural soils is worsened by atmospheric deposition, sewage irrigation practices, and the extensive use of soil amendments, livestock manures, pesticides and agrophytochemicals (Nicholson et al., 2003; Peng et al., 2019). Plants grown on contaminated soils may accumulate heavy metals in aerial parts such as leaf tissues and seeds and can result in severe health consequences for foraging animals and humans if these metals enter the food supply (Peralta-Videa et al., 2009).

In plants, heavy metals such as Cd and Hg are non-essential ions that can be taken up from the environment by essential micronutrient transporters (Clemens, 2006) and subsequently assimilated into the aerial parts of the plants *via* the roots. High intracellular accumulation of metal ions can lead to the denaturation of proteins, the displacement of essential metals from biomolecules, problems in membrane integrity, and the formation of reactive oxygen species (ROS). Excess metals in leaves can result in chlorosis, disruption of photosynthetic pathways, and breakdown of basic metabolic processes often leading to the death of the plant. Managing heavy metal toxicity by plants requires several genetic loci for integrated transport and tissue detoxification that operate at both the cellular and molecular level. Thus, traits associated with tolerance and accumulation are expected to be polygenic.

Molecular studies have shown that plants bind Cd and Hg with phytochelatins that are synthesized from glutathione and cysteine through the activation of the phytochelatin synthase (PCS) enzymatic pathway (Cobbett, 2001; Hossain et al., 2012). Binding and chelation of Cd and Hg ions enable their transport across membranes where they are sequestered into vacuoles for sub-cellular compartmentalization (Martinoia et al., 2018). The ATP-binding cassette subfamily of transporters (ABC-transporters), specifically *ABCC1* and *ABCC2* in *A. thaliana* and poplar, transport chelated Cd and Hg ions into vacuoles (Park et al., 2012; Brunetti et al., 2015; Sun et al., 2018) which reduces cellular toxicity and limits dissemination throughout the plant. Additionally, unchelated Cd ions can be directly sequestered into the vacuole and mediated by the heavy metal ATPase *HMA3* (Morel et al., 2008; Chao et al., 2012) and the cation exchange (CAX) type antiporters such as *AtCAX2* and *AtCAX4* (Cheng et al., 2005; Punshon et al., 2012). In rice (*Oryza sativa*), the natural resistance-associated macrophage protein 5 (*NRAMP5*, also a manganese (Mn) transporter), is the main transporter of Cd into roots through the apoplast (Sasaki et al., 2012; Clemens and Ma, 2016). *NRAMP6* also plays a role in Cd accumulation in *A. thaliana* as mutants confer greater Cd tolerance (Cailliatte et al., 2009). Translocation of heavy metals from the root into the shoot occurs by loading ions into the xylem often using heavy metal ATPases (HMA's), such as *HMA2* and *HMA4*, in *A. thaliana* and rice (Hanikenne et al., 2008; Takahashi et al., 2012). While far less is known about intracellular transport of Hg, there may be conservation between Cd and Hg transporters at the protein-family level given the similar vacuole transport mechanism for both Cd and Hg in *A. thaliana* using the same ABC-transporters (Park et al., 2012).

Identifying the physiological processes and genetic mechanisms that plants use to limit the accumulation of toxic heavy metals is important for human health and agriculture. The wild relatives of crop species possess standing variation that can be useful for identifying the molecular mechanisms associated with heavy metal tolerance and accumulation (Chao et al., 2012; Wu et al., 2015; Zhao et al., 2017; Chen et al., 2018). For example, allelic variants for reduced accumulation of heavy metals (excluder alleles) can minimize toxicity in plants grown in natural and agricultural ecosystems. Mining these genetic variants in species-wide germplasm collections of crop relatives can help us identify single nucleotide polymorphisms (SNPs) for both excluder and accumulator alleles and their associated genes. Conversely, accumulator alleles in highly tolerant plant genotypes may possess useful genetic adaptations that can facilitate phytoremediation of toxic soils (Swartjes, 2011; Yan et al., 2020) by increasing the capacity of plants to absorb high levels of toxic metals.

We are interested in the genetic architecture of tolerance to Cd and Hg and leaf accumulation of these metals in the model legume species *Medicago truncatula,* which is a Mediterranean forage legume and close relative of important crop legumes. Legumes such as *M. truncatula* have evolved symbiotic interactions with nitrogen-fixing bacteria (rhizobia) that can enrich plants with nitrogen. Symbiosis begins when the roots of the host plant come into contact with soil-borne

rhizobia, resulting in nodule formation on the roots (Kevei et al., 2002; Van de Velde et al., 2010). Inside the nodules, the rhizobia fix nitrogen for the host-plant in exchange for carbon nutrients from host-plants to rhizobia. Therefore, understanding genetic mechanisms that the host-plants use to detoxify plant tissues, particularly roots, is essential for ensuring successful rhizobia infection of roots to form nodules (León-Mediavilla et al., 2018; Arregui et al., 2021). Detoxification of roots involves both compartmentalization of toxic ions (i.e., in vacuoles) and transport of the ions away from roots often into aerial tissues such as leaves. However, excessive transport of toxic ions to aerial tissues would come at a high cost to plant fitness due to physiological stress and disruption of photosynthesis.

Considerable genomic resources exist for *M. truncatula* including high quality reference genomes and gene annotations (Young et al., 2011; Tang et al., 2014), genome-wide polymorphism data (Branca et al., 2011; Paape et al., 2013), a HapMap panel of resequenced germplasm for conducting genome wide association studies (GWAS) (Stanton-Geddes et al., 2013; Bonhomme et al., 2014; Kang et al., 2019), and a large mutant collection (Lee et al., 2018). These resources allow us to map phenotypes to very fine genomic regions and identify candidate genes associated with traits of interest, providing the basis for functional genetics studies (Curtin et al., 2017). Moreover, in previous studies Hg- and Cd-sensitive and tolerant cultivars were identified by phenotyping *M. truncatula* germplasm resources, and differential responses to metal stress between tolerant and sensitive cultivars were reported (García de la Torre et al., 2013, 2021). Both cultivar selection and transgenic approaches can be used to obtain legumes with increased tolerance to abiotic stress (Coba de la Peña and Pueyo, 2012).

Standing genetic variation for quantitative traits, such as phenotypic responses to heavy metals, is useful for detecting allelic variation (i.e., SNPs) or genomic divergence associated with trait variation. The variation present in species-wide samples from natural populations (such as the Medicago HapMap panel) may be comprised of adaptive and deleterious alleles that have been tested in highly variable, natural habitats. In addition to identifying genetic loci associated with traits, GWAS can be used to estimate allele frequencies of associated SNPs and their effect size, which may reveal the forces of selection that contributed to the genetic architecture of a trait (Stinchcombe and Hoekstra, 2007; Josephs et al., 2017). In GWAS data, we expect that mutations (SNPs) with larger effect sizes will most often be deleterious (Eyre-Walker and Keightley, 2007) and be subject to purifying or negative selection (Trotter, 2014) under mutation-selection balance (Crow and Kimura, 2010). These processes can be revealed by negative correlations between SNP effect size (estimated in the GWAS) and minor allele frequencies (MAFs), where SNPs with smaller effect on a trait will be at higher frequency and those with a larger effect will be at lower frequencies (Stanton-Geddes et al., 2013; Josephs et al., 2015). In addition, the phenotypic distributions exhibited by standing variation may also correspond to genomic divergence. By treating genotypic groups at opposite ends of phenotypic distributions as populations, it may be possible to detect signals of genomic divergence using population differentiation methods such as F-statistics (F_{st}) and composite likelihood ratio (CLR) tests. F-statistics are traditionally used to test for natural selection between populations by comparing allele frequencies within and between groups (Holsinger and Weir, 2009), while CLR tests detect selective sweeps between two populations (XP-CLR; Chen et al., 2010). These tests may provide complementary information to GWAS regarding genomic regions underlying divergent trait values.

Toxic levels of Cd and Hg in plants can be estimated based on relative root growth (RRG) and leaf accumulation. RRG acts as a measure of tolerance or root growth inhibition (García de la Torre et al., 2013, 2021) while leaf accumulation measured using ionomics characterizes leaf-level responses such as root-to-shoot transport and leaf capacity. In this study we set out to (1) quantify the phenotypic variation of Cd and Hg tolerance and accumulation in the Medicago HapMap panel, (2) identify genes that are associated with these traits using complementary methods to locate genomic regions, (3) characterize the genetic architecture of tolerance and accumulation of these two metals, and (4) estimate whether the genetic architecture has been shaped by selection acting upon natural variation. When orthologous genes are detected in GWAS experiments across multiple plant species it suggests conservation of metal transporters and detoxification mechanisms across the plant kingdom. Overall, the expected results of the present study will contribute to our understanding of the mechanisms involved in heavy metal tolerance, translocation and accumulation in plants, and will provide candidate targets to modify heavy metal accumulation or tolerance by traditional breeding or transgenic approaches.

METHODS

Plant Growth and Measurements of Phenotypes

Seeds from 236 resequenced *M. truncatula* genotypes were obtained from the University of Minnesota, Medicago HapMap project[1]. Two separate heavy metal treatments were applied in parallel to the set of *M. truncatula* HapMap genotypes at the seedling stage. A modified Hoagland nutrient solution (based on García de la Torre et al., 2013, 2021) was used to grow plants in hydroponic conditions using the following nutrient concentrations: 2.02 gL^{-1} KNO_3, 0.68 gL^{-1} KH_2PO_4, 0.182 gL^{-1} $CaCl_2{\cdot}2H_2O$, 0.615 gL^{-1} $MgSO_4{\cdot}7H_2O$, 0.109 gL^{-1} K_2SO_4, 0.205 gL^{-1} Hampiron (Rhône Poulenc), and 1.35 mL of a solution containing: 11 gL^{-1} H_3BO_3, 6.2 gL^{-1} $MnSO_4 \cdot H_2O$, 10 gL^{-1} KCl, 1 gL^{-1} $ZnSO_4 \cdot 7H_2O$, 1 gL^{-1} $(NH_4)_6$ Mo_7 O_{24} $\cdot 4H_2O$, 0.5 gL^{-1} $CuSO_4{\cdot}5H_2O$ and 0.5 mLL^{-1} H_2SO_4. One set of plants was treated with mercury using 4 μM $HgCl_2$ added to the Hoagland solution. A second set of plants was treated with cadmium using 10 sμM $CdCl_2$ added to the Hoagland solution. A third set of plants was given no heavy metal treatment and was used as a control group. Each treatment began with fifteen replicate plant seedlings for each *M. truncatula* genotype. Four traits were measured on the *M. truncatula* Hapmap panel

[1]http://www.medicagohapmap.org/hapmap/germplasm

following the heavy metal treatments: (1) Relative Root Growth (RRG) in plants treated with Hg and (2) RRG in plants treated with Cd, (3) accumulation of Cd in leaf tissues, (4) and accumulation of Hg in leaf tissues. Seedlings were placed for 24 h in the hydroponic system with 250 mL of nutrient solution for acclimatizing in growth chamber conditions (24/20°C, 16/8 h photoperiod), and then 48 h in the same conditions with or without the heavy metal treatment. RRG is a reliable indicator of metal tolerance in *M. truncatula* (García de la Torre et al., 2013, 2021). Root length was measured at 24 h of plants growing in the hydroponic medium before metal was added, and 48 h later, after applying the metal treatments. Photo images of the seedlings were taken at 24 h and again after the 48 h treatment. Roots were measured using the software ImageJ. For calculating the RRG of the seedlings, the increase in length is normalized by the increase in control seedlings (Equation 1):

$$RRG = \left(\frac{\Delta length_{treatment}}{\Delta length_{control}}\right) \times 100$$

For metal concentration measurements in leaves, cotyledons were harvested after plants had been exposed to Cd or Hg treatment for 48 h. The leaf tissues were washed with 10 mM Na_2EDTA to remove traces of metals on their surface. Washed and dried leaves were digested using concentrated HNO_3 and H_2O_2. After the tissue digestion, distilled water was added, then the mixture was filtered. Cd and Hg concentrations were measured using inductively coupled plasma atomic emission spectroscopy (ICP-AES). For leaf accumulation, three samples (biological replicates) were analyzed per genotype. We used R version 3.3.3 to conduct statistical on phenotypic distributions and the lmer4 package to estimate variance components to calculate broad sense heritability (H^2). Metal concentrations were determined as previously reported (García de la Torre et al., 2013, 2021).

Genome-Wide Association Study

Genome-wide SNP data (BCF files) called against the Mt4.0 version of the reference genome from 262 *Medicago truncatula* accessions was obtained from the Medicago HapMap project[2]. Missing SNPs (due to low coverage in some genotypes) were imputed using BEAGLE version 4.1 (Browning and Browning, 2016) with default parameters. The imputed dataset was split by chromosome and converted to hapmap format (.hmp) using TASSEL 5 (Bradbury et al., 2007). We used the software Admixture (Alexander et al., 2009) to generate a population structure co-variance matrix. The following criteria was used to select independent SNPs for the Admixture analysis, based on similar criteria used by Gentzbittel et al. (2019) to remove SNPs that were linked or below allele frequency thresholds. SNPs were filtered by genotyping rate and minor allele frequency, then converted to bed format using PLINK 1.9 beta 6.5 (Chang et al., 2015) using the command line input parameters: –geno 0.05 –maf 0.01 –make-bed. Independent sites were extracted using PLINK using the following input parameters: -indep 300 60 1.22. Admixture v1.3.0 was run on 843,307 SNPs and *k*-values ranging from 1 to 10 were estimated by running 10 iterations per K with a different seed value for each iteration (–seed = 1 to –seed = 10). For each k, average cross validation errors were calculated and the iteration with lowest CV error was plotted in R. The SNP dataset containing 262 HapMap accessions was filtered to contain accessions with phenotype data present in the current study, resulting in a SNP dataset containing 236 genotypes.

We ran the GWAS using GAPIT (Lipka et al., 2012; Tang et al., 2016) version 20160323, once with the population structure covariance matrix (with lowest cross validation error, $k = 5$), and a second time without the structure covariance matrix. A kinship (K) matrix, to control for genetic relatedness was estimated by the software. The parameters used in the GAPIT model were KI = NULL, PCA.total = 3, SNP.MAF = 0.02, SNP.fraction = 0.6, Major.allele.zero = TRUE and Geno.View.output = FALSE. We also ran the GWAS using GEMMA (Zhou and Stephens, 2012). The imputed dataset was filtered to only contain phenotyped individuals with bcftools 1.2 and subsequently converted to bed format using PLINK 2.0 alpha while the phenotype data was converted to fam format using the PLINK option: –make-just-fam. We then used GEMMA 0.98.1 to calculate relatedness matrices for each trait and to perform GWAS iterations using a multivariate linear mixed model and a minor allele frequency cutoff of 2 percent (parameters: −lmm -maf 0.02). As with GAPIT, GEMMA was run with and without a population structure covariance matrix for comparison.

We avoided using strict *p*-value thresholds because they would be less informative for characterizing the genetic architecture of polygenic traits. Instead, we explored the larger landscape of SNPs by utilizing a range of *p*-values. Because we began with several million SNPs, the top 100 to 1000 SNPs with the lowest *p*-values for each trait are potentially relevant based on functional annotations of genes in the vicinity of these SNPs. The 1000 most significant SNPs across all chromosomes were annotated using a custom Python script for the genes within 1 kb distance from these SNPs using the gene context files available from the HapMap project[3]. Among these 1000, we can select any subset such as the top 100 for each trait. Information including distance between SNP and gene, or the nucleotide substitution type (i.e., synonymous, missense), were extracted from the gene context files. Additionally, the blastn tool provided by NCBI (Camacho et al., 2009) was used to perform BLAST of the genes against *Arabidopsis thaliana*. Gene descriptions, gene ontologies (GO-terms) and tissue-specific RNAseq expression levels (from six tissue types in the *M. truncatula* A17 genotype) were further annotated using information from MedicMine (Krishnakumar et al., 2015). Initially, genes with relevance to heavy metal tolerance were categorized as ATPase, metal ion, ion transport, and stress, based on their functional annotation in *M. truncatula*, GO-terms, and the annotation of the closest BLAST ortholog in *A. thaliana*. A statistical test for GO-term enrichment using the genes with the 1000 most significant SNPs was conducted using AgriGO (Du et al., 2010) using false discovery adjusted *p*-value ≤ 0.05. Regions with a high density of candidate SNPs ("GWAS peaks") were manually identified using IGV 2.6.3

[2]http://www.medicagohapmap.org/downloads/mt40

[3]http://www.medicagohapmap.org

(Robinson et al., 2011; Thorvaldsdóttir et al., 2013). To calculate pairwise linkage disequilibrium (LD) between the 100 most significant SNPs in the peaks, vcftools was used with the –geno-r2 flag and the resulting values were plotted using the R package LDHeatmap (Shin et al., 2006).

Quantifying Genomic Divergence Using Phenotypic Divergence

The phenotypic distributions of each trait (**Figure 1** and **Supplementary Figure 1**) were used to define two groups, high and low, (per trait) based on trait measurements in the opposite tails of the distributions. For each of the traits measured, the two groups consisted of 30 individuals with the lowest phenotype values and the other consisted of 30 individuals with the highest phenotype values. We used *vcftools* (Danecek et al., 2011) to generate F_{st} statistics with the previously mentioned groups and a window and step size of 100 kb. Similarly, xpclr 1.1[4] was used to calculate XP-CLR (Chen et al., 2010) with identical window and step size. Genes in the top 2 percent of windows with highest values of F_{st} or XP-CLR were annotated in the same way as the GWAS SNPs and were further tested for GO-term enrichment. For some genomic regions, overlap between GWAS peaks and the two statistics were determined.

[4]https://github.com/hardingnj/xpclr/

Correlation Between Minor Allele Frequency and Effect Size

The correlation between minor allele frequency (MAF) and effect size of the 100 and 1000 most significant GWAS SNPs was determined using Pearson's correlation coefficient for both GAPIT and GEMMA output datasets. We permuted genotype-phenotype values to make 100 phenotype datasets, then ran 100 iterations of GWAS iterations to make null distributions of MAF and effect sizes. To run the permutations, we used GEMMA, as it performs far better than GAPIT in terms of computational time. For each iteration, Pearson's correlation coefficients were calculated for minor allele frequency (MAF) and effect size using the 100 and 1000 most significant SNPs. The resulting distribution of correlation coefficients was compared to the correlation of the unpermuted results for each phenotype.

To test whether the most significant SNPs show significant differences than the genomic background, Tajima's D was calculated for the 1000 most significant SNPs of each trait using

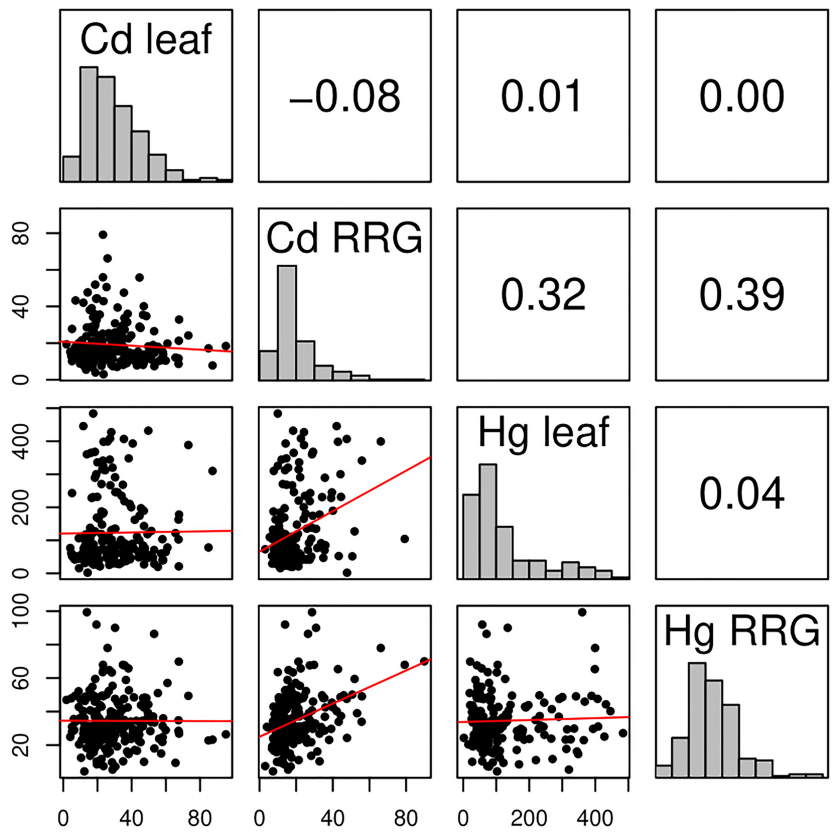


FIGURE 1 | Correlation matrix of Cd leaf accumulation, Cd relative root growth (RRG), Hg leaf accumulation, and Hg relative root growth (RRG) measured in the *M. truncatula* HapMap panel. Distributions of the four phenotypes are shown in the diagonals. The upper, off-diagonal panels contain pairwise Pearson's correlation coefficients (*r*). Lower, off-diagonals are the plotted trait values of each *M. truncatula* genotype measured in our experiment. The red line shows the slope of the correlation between any pair of traits. The *x*- and *y*-axes are values of the phenotype measurements [in μg^{-1} dry weight for leaf, and percent for relative root growth (RRG)].

variscan (Vilella et al., 2005; Hutter et al., 2006) with a window size of 50 bp. We used SNPs that were selected for the Admixture analysis as the background empirical distribution. A *t*-test was used to determine whether the differences between Tajima's D of the SNPs identified in the GWAS and those in the genomic background were significant.

RESULTS

High Phenotypic Variability, Heritability, and Correlations Among Traits

We measured seedlings from the Medicago HapMap collection for relative root growth (RRG) and metal ion accumulation in leaf tissues following 48 h exposure to Cd and Hg treatments (**Supplementary Table 1**). We found nearly 30-fold difference for Cd RRG and 23-fold difference for Hg RRG between the lowest and highest tolerant genotypes (**Figure 1**). Broad sense heritability (H^2) was equal to 0.61 and 0.72 for the two RRG traits, respectively (**Supplementary Figure 1**). We found 50-fold and 270-fold differences in Cd leaf and Hg leaf accumulation, respectively; H^2 was equal to 0.54 and 0.44 for Cd and Hg leaf accumulation, respectively. No correlation was found between leaf accumulation and RRG for either metal treatment (i.e., plants treated with Cd that had high RRG did not have high leaf accumulation of Cd, and similarly for Hg) (**Figure 1** and **Supplementary Figure 1**). The highest correlation was between the two root growth traits, Cd and Hg RRG (Pearson's $r = 0.39$). The heritability estimates and large amount of standing variation suggest that GWAS can identify SNPs and genes underlying these traits due to sufficient genetic differences between genotypes, and that both tolerance and susceptibility alleles are present in the *M. truncatula* HapMap panel.

Population Structure

We conducted a GWAS to identify genomic regions associated with leaf and root responses to Cd and Hg treatments in *M. truncatula* using a dataset of ∼12 million SNPs after imputing missing data. Among these, we sampled 843 k SNPs to generate a covariance matrix (Q) to control for confounding effects of population structure in our GWAS. Based on our analysis of population structure, we determined that $k = 5$ was the most strongly supported clustering (**Supplementary Figures 2, 3**) using the cross-validation method in Admixture (Alexander et al., 2009). Proportions of each genotype assigned to each of the five clusters were then used in the covariance matrix Q. Including only kinship (K) co-variance in the GWAS had similar empirical and expected distributions of *p*-values (i.e., similar model fit) in the quantile-quantile plots as did K and Q together ($K + Q$, kinship plus population structure) for each of the four traits (**Figure 2**). This indicated that kinship largely controlled for confounding genotype relationships with only slight improvement of the model fit by adding structure covariance (e.g., Cd RRG and Hg leaf accumulation). We found that GAPIT and GEMMA performed similarly, but the quantile-quantile plots from GAPIT showed slightly better fitting models (**Figure 2** and **Supplementary Figure 4**). There were no obvious differences in model fit between either software when the structure covariance matrix was included or excluded. Visual comparison of quantile-quantile plots showed an equal ability of both programs to correct for population structure and kinship.

Toxic Metal Leaf Accumulation and Relative Root Growth Are Polygenic Traits

Candidate gene lists were first generated by selecting the 50 most significant SNPs (**Table 1** and **Supplementary Table 2**) within 1 kb proximity of a gene for each trait SNPs were assumed to be in LD with genes closer than 1 kb, Paape et al., 2012; Paape, 2020). We then broadened our search for genes to include the top 1000 SNPs for each trait and identified genes with functional annotations based on Blast results to *A. thaliana,* and manual curation of relevant gene annotations and gene ontologies. This resulted in about half of the SNPs for each trait being associated with a gene, which is a useful number of genes for GO-term enrichment analysis. We first grouped genes with functional annotations for ATPase, ion transport, metal ion, or stress (**Figure 3A**). Among the traits, numbers ranged from 29 - 42 genes that had any of these four functional assignments (**Supplementary Table 3**). The two RRG traits had the most genes, with metal ion and ATPase being the most common in all four of the traits. To determine whether any biological processes or molecular functions were enriched among the candidate genes, we performed tests for gene ontology (GO-term) enrichment for each trait. Response to stress and defense response GO-terms were enriched for Cd RRG, defense response was enriched for Hg leaf, and nucleotide and ATP binding were enriched for Hg RRG (**Supplementary Table 4**). No significant GO-terms were found for Cd leaf.

As a complementary approach to GWAS, we used phenotype values of plant genotypes in the high vs. low ends of the phenotypic distributions, and F_{st} and XP-CLR population genetics statistics to identify genomic regions with high divergence between genotypes within these two groups. We were able to detect significant divergence in genomic regions that contained genes with annotations relevant for heavy metal accumulation and tolerance. The number of genes with ATPase, ion transport, metal ion, or stress related annotations ranged from 69 to 99 total genes per trait (**Figures 3B,C**). Consistent with the GWAS results, the metal ion category was the most represented across all traits for both population genetics statistics. Among the four traits, GO-term enrichment was most noteworthy for Cd leaf accumulation using the two population genetics methods. For genes identified using F_{st}, enriched GO-terms were found for ion binding, cation binding, metal ion binding, response to chemical stimulus, and response to oxidative stress/oxidoreductase activity (**Supplementary Table 5**). For genes identified using XP-CLR, we found GO-term enrichment for oxidoreductase activity, carbon-oxygen lyase activity, FAD binding, and manganese ion binding. The distribution of gene categories and gene ontologies illustrate that by selecting plant genotypes with divergent phenotypes, we were able to detect

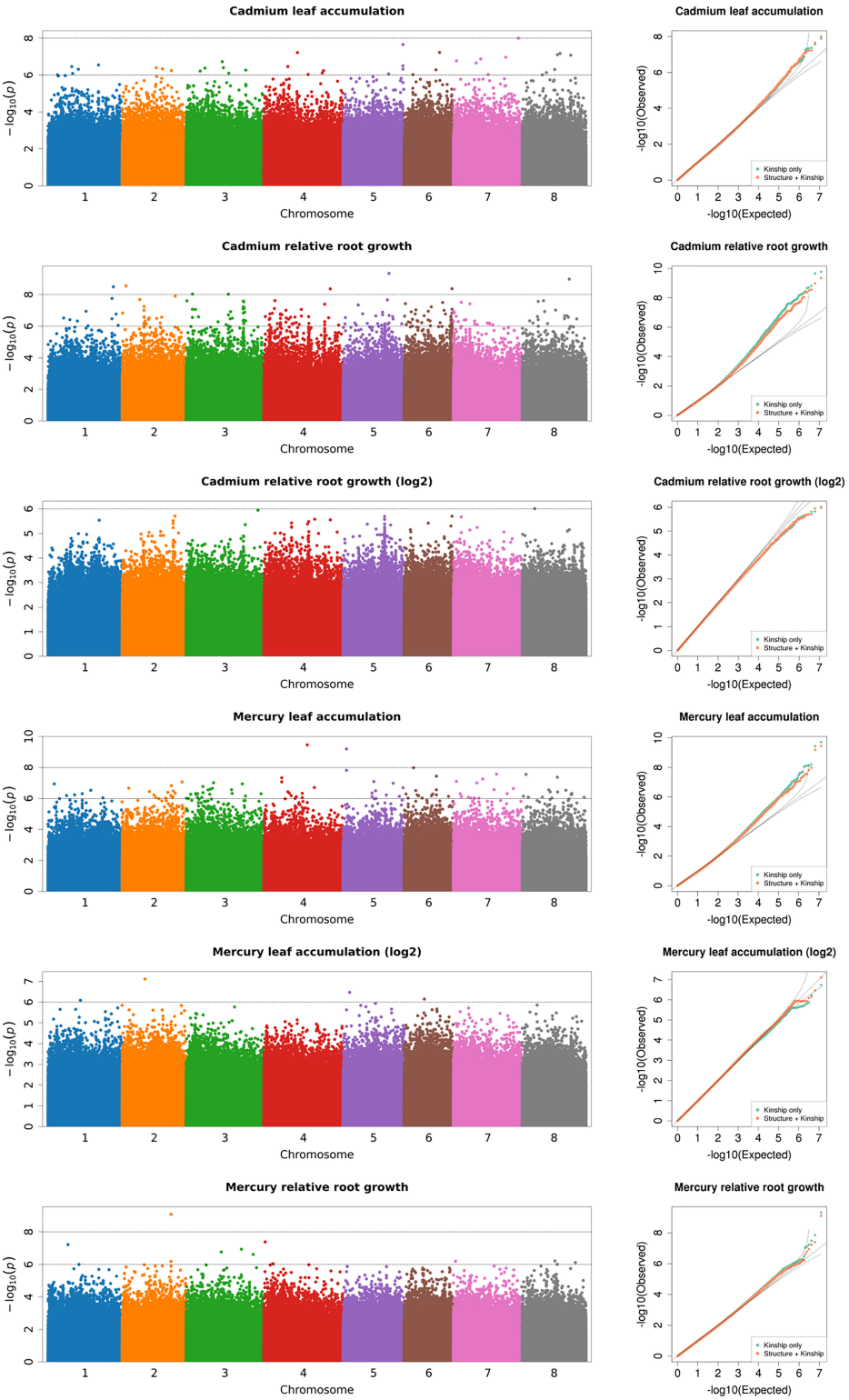


FIGURE 2 | Genome-wide association analysis of heavy metal accumulation and tolerance traits (Manhattan plots shown for each trait; output from GAPIT). Each point represents a SNP at each position in the genome. The number on the *x*-axis is the chromosome (each color corresponds to one of the eight Mt4.0 chromosomes). The *y*-axis is the negative $\log_{10}$ of the *p*-values. The higher *y*-axis values indicate smaller *p*-values. Next to each Manhattan plot are quantile-quantile (Q-Q) plots which show the model fit with population structure covariance included (orange points), and without population structure covariance (green points). The *x*-axis of the Q-Q plots is the expected distribution, the *y*-axis is the empirical distribution.

TABLE 1 | Candidate genes for each trait within 1 kb proximity to top 50 SNPs and a subset of the top 1000 SNPs identified by GWAS.

Rank	Chr	Gene Name	Position	*P*-value	Location	Gene Description
Cd leaf						
5	8	Medtr8g064490	27050814	6.72E-08	intergenic	23S rRNA m2A2503 methyltransferase
6	8	Medtr8g469310	25178759	7.84E-08	intergenic	peptide/nitrate transporter plant
10	7	Medtr7g009740	2221211	1.68E-07	intron	translation factor GUF1-like protein
22	3	Medtr3g036200	13449844	4.18E-07	downstream	transmembrane protein
26	8	Medtr8g465310	23246354	4.94E-07	synonymous	transmembrane amino acid transporter family protein
29	3	Medtr3g093290	42636439	5.37E-07	downstream	MFS transporter
32	3	Medtr3g031940	10032134	6.01E-07	synonymous	LRR and NB-ARC domain disease resistance protein
40	6	Medtr6g016640	6335718	9.40E-07	intron	proline dehydrogenase
43	1	Medtr1g033940	12271020	1.07E-06	intergenic	transmembrane protein
47	8	Medtr8g090295	37943857	1.48E-06	intergenic	HCC2, HOMOLOGUE OF COPPER CHAPERONE SCO1 2
52	7	Medtr7g092070	36458974	1.58E-06	downstream	OXS2, OXIDATIVE STRESS 2
605	5	Medtr5g094830	41452129	3.34E-05	synonymous	ABCC3; transmembrane ATPase activity
Cd RRG						
4	1	Medtr1g103200	46692340	3.17E-09	downstream	transmembrane protein, putative
5	6	Medtr6g090260	34282023	4.18E-09	missense	Fe(II)-dependent oxygenase superfamily protein
7	3	Medtr3g015500	4499608	9.09E-09	downstream	NBS-LRR type disease resistance protein
11	1	Medtr1g100787	45717235	1.75E-08	intergenic	LRR receptor-like kinase family protein
24	7	Medtr7g033125	11718492	3.78E-08	intron	2-isopropylmalate synthase
25	4	Medtr4g104990	43495277	4.00E-08	intergenic	PB1 domain protein
26	5	Medtr5g026510	10889054	4.47E-08	downstream	LRR receptor-like kinase
27	2	Medtr2g438720	15656689	5.61E-08	missense	ankyrin repeat plant-like protein
28	3	Medtr3g089940	40881882	5.83E-08	intron	zinc-binding alcohol dehydrogenase family protein
36	3	Medtr3g088845	40664333	7.88E-08	intron	2-oxoisovalerate dehydrogenase subunit alpha
41	3	Medtr3g088955	40721568	7.88E-08	synonymous	flavin containing monooxygenase YUCCA10-like protein
45	3	Medtr3g015500	4499656	8.43E-08	downstream	NBS-LRR type disease resistance protein
148	8	Medtr8g015980	5289479	4.56E-07	missense	ABCC8; transmembrane ATPase activity
192	5	Medtr5g070320	29773843	6.99E-07	intron	PDR3; pleiotropic drug resistance 3
340	5	Medtr5g070330	29783407	1.96E-06	intron	CAX3; cation exchanger 3
363	2	Medtr2g069090	28697944	2.08E-06	downstream	Fe(II)-dependent oxygenase superfamily protein
612	3	Medtr3g090170	40983523	4.49E-06	intron	ATVPS45; vacuolar protein sorting 45
797	2	Medtr2g019020	6102461	7.32E-06	synonymous	ABCC2; transmembrane ATPase activity
944	2	Medtr2g095480	40788857	9.99E-06	upstream	HIPP3; Heavy metal transport/detoxification superfamily protein
Hg leaf						
1	4	Medtr4g080140	31078608	3.47E-10	intergenic	ASP1; aspartate aminotransferase
3	6	Medtr6g017165	6799845	1.05E-08	intron	allene oxide cyclase
5	7	Medtr7g080880	30810684	2.68E-08	downstream	SPO22/ZIP4-like meiosis protein
19	3	Medtr3g088460	40127485	1.15E-07	intron	NRAMP metal ion transporter 6
25	2	Medtr2g015690	4673621	2.16E-07	downstream	phospholipase A1
27	7	Medtr7g105570	42817932	2.28E-07	intergenic	enoyl-(acyl carrier) reductase
33	8	Medtr8g056790	18970685	2.81E-07	intergenic	photosystem II CP43 chlorophyll apoprotein
35	8	Medtr8te073680	31194688	2.96E-07	downstream	Serine/Threonine-kinase Cx32, related protein
43	8	Medtr8g469600	25322097	3.80E-07	downstream	LRR receptor-like kinase family protein
513	2	Medtr2g036380	15751342	8.27E-06	intron	HMA2; heavy metal atpase 2
917	5	Medtr5g033320	14374444	1.96E-05	intron	ABCC3; transmembrane ATPase activity
Hg RRG						
1	2	Medtr2g083420	35018740	8.10E-10	missense	UDP-glucosyltransferase family protein
4	3	Medtr3g087150	39520330	1.15E-07	intergenic	plant/T5J17-70 protein, putative
7	8	Medtr8g465580	23392640	6.07E-07	synonymous	S-locus lectin kinase family protein
10	8	Medtr8g091330	38065096	7.59E-07	downstream	phosphoglycerate/bisphosphoglycerate mutase family protein
12	4	Medtr4g020850	6714547	9.31E-07	missense	disease resistance protein (TIR-NBS-LRR class)
15	2	Medtr2te036520	15806363	1.04E-06	downstream	galactose oxidase/kelch repeat protein
18	3	Medtr3g040680	14343459	1.08E-06	intergenic	non-specific phospholipase C4
31	3	Medtr3g064650	29135942	1.64E-06	synonymous	transmembrane protein, putative
34	7	Medtr7g094500	37617280	1.69E-06	intron	u6 snRNA-associated-like-Sm protein
58	4	Medtr4g116010	47960277	2.51E-06	intergenic	Fe(II)-dependent oxygenase superfamily protein
81	2	Medtr2te008720	1586136	3.51E-06	downstream	Peroxidase superfamily protein
108	5	Medtr5g073020	31072945	6.78E-06	intron	Cytochrome P450, family 71, subfamily B, polypeptide 23
270	3	Medtr3g014705	4234570	1.27E-05	intron	Zinc finger (CCCH-type) family protein / RNA recognition motif
307	1	Medtr1g084780	37746862	1.47E-05	downstream	H+ ATPase 2
351	1	Medtr1g063920	28084622	1.66E-05	intron	non-intrinsic ABC protein 12
357	4	Medtr4g076900	29426248	1.67E-05	intron	ABC-2 type transporter family protein

Columns from left to right are Rank (rank order assigned based on the top 50 and top 1000 SNPs), chromosome (Chr), Gene Name (Mt4.0 gene ID), Position on the chromosome, P-value, Location in proximity of the gene, and Gene Description. Within each trait, SNPs are sorted smallest to largest p-value and rank.

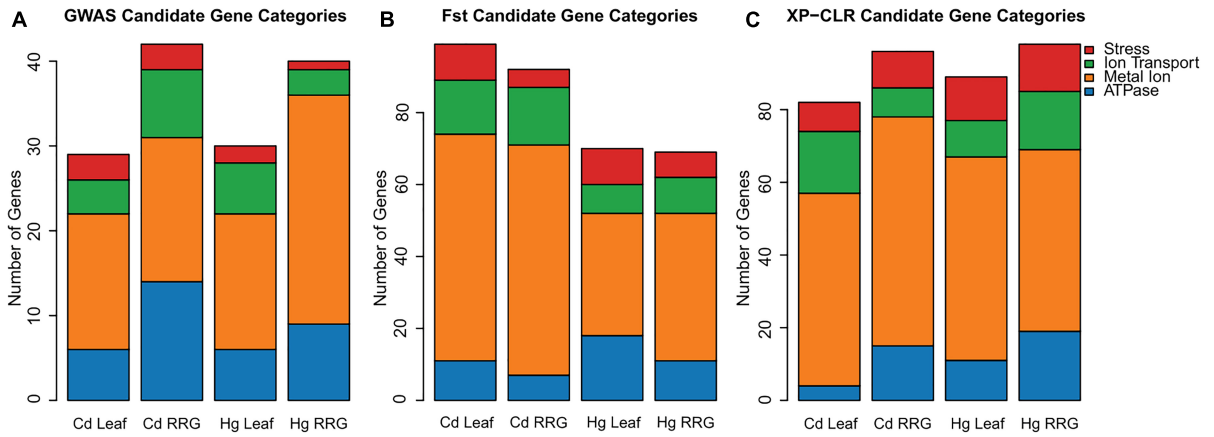


FIGURE 3 | The number of genes with annotations in categories for ATPase, ion transport, metal ion, and stress related genes within 1 kb of the 1000 most significant SNPs using GAPIT for each of the four traits **(A)**. The number of genes in these four categories in the upper 2% of the highest F_{st} sliding windows **(B)**, and the number of genes with these four categories in the 2% of sliding windows with highest XP-CLR-values **(C)**. The *y*-axis represents the number of genes in each category.

regions of genomic divergence with biological relevance for metal ion stress.

Candidate Gene Identification

To further identify candidate genes detected by GWAS, we looked at gene ontologies that mentioned metal ion transport, trans-membrane proteins, and cellular detoxification, and orthologous genes and gene families based on Blast hits to *A. thaliana* that have some role in heavy metal tolerance/accumulation (**Table 1** and **Supplementary Table 2**). For Cd leaf accumulation, among the top 50 SNPs, we found a GTPase membrane protein (Medtr7g009740) and multiple transmembrane proteins (Medtr3g036200, Medtr3g093290, Medtr8g465310). An oxidative stress protein, *MtOXS2* (Medtr7g092070) was among the top 100 SNPs and an ATP-binding cassette protein, *MtABCC3* [Medtr5g094830, orthologous to *AtABCC3* (Brunetti et al., 2015)] was among the top 1000 SNPs. For Cd RRG, among the top 100 SNPs, we found a non-synonymous SNP in a Fe (II)-dependent oxygenase superfamily protein (Medtr6g090260), a leucine-rich transmembrane protein kinase (Medtr1g100787) with ATPase gene ontology, and a non-synonymous SNP in an ankyrin repeat plant-like protein (Medtr2g438720), which shows highest expression in roots in six tissue-specific comparisons in the A17 ecotype/genotype (**Supplementary Table 2**). The top 100-1000 SNPs for this trait included the ABC-transporters *MtABCC2* (Medtr2g019020) and *MtABCC8* (Medtr8g015980), and Gamma-Glutamyl Transpeptidase 3, *MtGGT3* (Medtr6g090280), which is localized on tonoplasts in *A. thaliana* (Ohkama-Ohtsu et al., 2007), an ortholog of the vacuolar protein *AtVPS45* (Medtr3g090170), and another Fe (II)-dependent oxygenase superfamily protein. We also found associations in the heavy metal transport/detoxification superfamily protein *MtHIPP3* (Medtr2g095480). Heavy-metal-associated Isoprenylated Plant Proteins (HIPP's) are a family of metallochaperones found only in plants (de Abreu-Neto et al., 2013) that undergo transcriptomic and functional responses to Cd and Zn in *A. thaliana* (Tehseen et al., 2010) and rice (Khan et al., 2019).

Using GWAS and population genomic scans (F_{st}, and XP-CLR), we found a large region (∼1.3 Mb) of interest associated with Cd RRG on chromosome 2. A highly significant non-synonymous SNP was found at position 15,656,689 (p-value = 5.61×10^{-8}) in an ankyrin repeat protein (Medtr2g438720), and the surrounding region contained multiple ankyrin repeat genes within 1 kb of many top SNPs (**Figure 4A** and **Table 2**). Eleven of those SNPs were in exons, nine of which were non-synonymous SNPs, while two others were synonymous (**Supplementary Table 6**). The high LD in this region suggested the ankyrin genes were linked. While these ankyrin genes did not show homology to *A. thaliana* orthologs, some ankyrin repeat proteins have been shown to mediate membrane bound protein-protein interactions that facilitate heavy metal transport in *A. thaliana* roots (Gao et al., 2009). Genomic divergence between high and low Cd RRG genotypes was found by scanning for high F_{st} in a region upstream from the GWAS peak (**Figure 4B**). This region contains a tandemly triplicated ABC transporter (*MtABCC14*, Medtr2g436680; **Supplementary Figure 5**), an ortholog of Heavy Metal Transporter 10 (*MtHMP10*, Medtr2g436830), a Fe (II)-dependent oxygenase superfamily protein, ATPase transporter *MtHMA7* (Medtr2g035840), and a potassium transporter (Medtr2g438160; **Table 2**). A significant peak was also detected using XP-CLR for the window containing the three *MtABCC14* copies and *MtHMP10* (Medtr2g436830; **Figure 4C** and **Table 2**), which was the same genomic region identified using F_{st}.

Another cluster of low p-values associated with Cd RRG was found on chromosome 5 (**Supplementary Table 7**) with SNPs in high LD (**Figure 5A**). When we ran the GWAS following $\log_2$ transformation of the Cd RRG phenotype, the model fit was somewhat improved, and this peak became more pronounced (**Figure 2** and **Supplementary Figure 4**). In this genomic region, five SNPs (four intronic SNPs and one in the 3′-UTR)

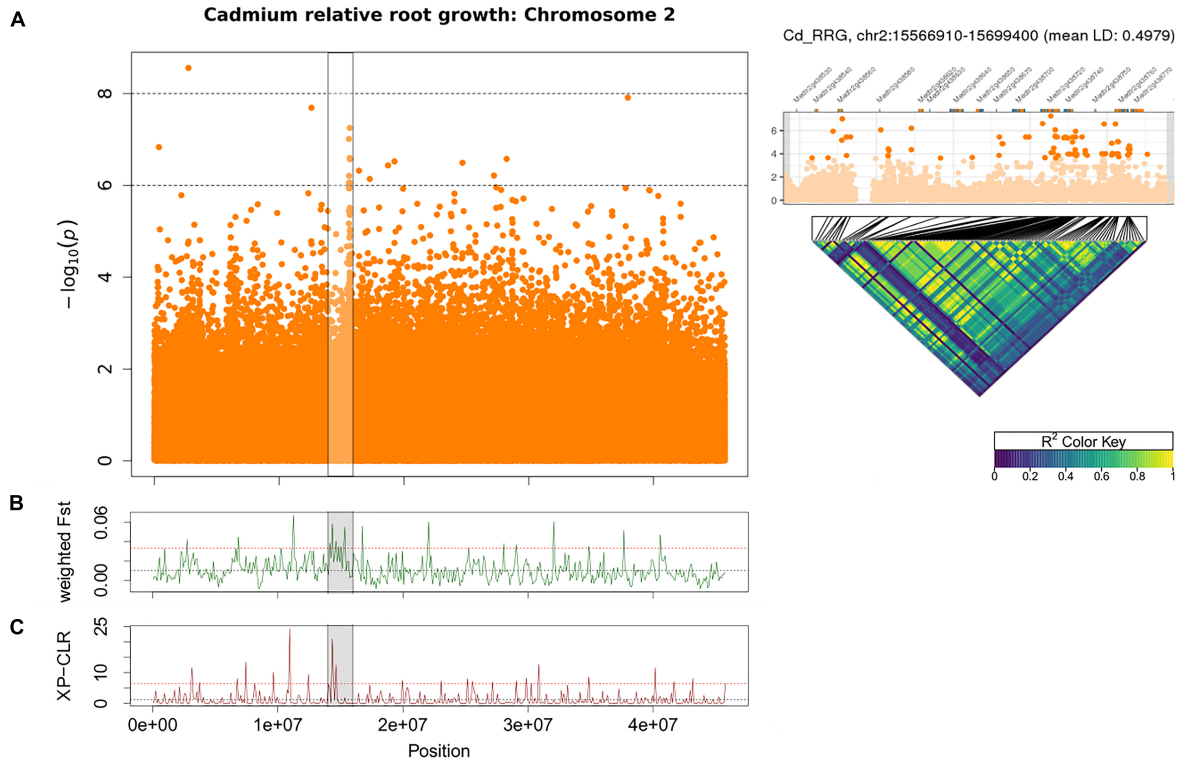


FIGURE 4 | Manhattan plot of Cd relative root growth (RRG) associations on chromosome 2 using GAPIT **(A)**. The highlighted cluster of SNP's (horizontal black bars) contains multiple ankyrin repeat genes (see **Table 2**). Each point represents a SNP, the *x*-axis signifies the position on the chromosome while the *y*-axis is the $-\log_{10}$ of the *p*-value. Pairwise linkage disequilibrium (LD) among the 100 most significant SNPs in the peak is shown as a heatmap on the right, the most significant SNPs and Mt4.0 genes in this region shown above the heatmap. The black lines mark the position of each SNP in the peak, dark blue color signifies no LD, green to yellow is high to maximum LD. Weighted F_{st} statistics from sliding window analysis (100 kb windows) on chromosome 2 with GWAS peak marked in gray **(B)**. Higher F_{st} indicates genomic differentiation based on the groups of plant genotypes with the highest vs. lowest Cd RRG-values. XP-CLR is a two-group comparison of selective sweeps using plant genotypes with the highest vs. lowest Cd RRG-values **(C)**. The dashed black line represents the mean F_{st} or CLR-values for the chromosome, the dashed red line represents windows ≥ 5% of the highest F_{st} or CLR-values for the chromosome.

were in *MtCAX3* (Cation Exchanger 3, Medtr5g070330), and three SNPs in the Pleiotropic Drug Resistance 3 gene, *MtPDR3* (Medtr5g070320). In the Mt4.0 reference genome, *MtCAX3* and *MtPDR3* are adjacent (5878 bp apart; **Supplementary Figure 6**). The peak also contained a non-synonymous, synonymous, and several intronic SNPs in a damaged DNA-binding protein, *MtDDB2* (Medtr5g070310), and an undecaprenyl pyrophosphate synthase *MtCPT7* (Medtr5g070270), which is orthologous to *AtCPT7* (also known as *AtCPT4*). We found several intergenic SNPs with low *p*-values in high LD in the region containing a cluster of duplicated undecaprenyl pyrophosphate synthase genes (Medtr5g070210, Medtr5g070220, Medtr5g070230; **Figure 5**) that are upstream from *MtCPT7*, all of which appear to be linked. The window containing *MtCAX3* and *MtPRD3* is also in the upper tail of F_{st} on chromosome 5 (**Figure 5B**), and the XP-CLR statistic showed a significant peak just upstream from the GWAS peak (**Figure 5C**).

For Hg leaf accumulation, the most significant SNP was in an intergenic region within 1 kb of an aspartate aminotransferase (Medtr4g080140), and among the top 20 SNPs was the Natural Resistance Associated Macrophage Protein 6 gene (*MtNRAMP6*, Medtr3g088460). In the *M. truncatula* A17 genotype, *MtNRAMP6* is expressed most highly in the roots (**Supplementary Table 2**). Genes found among the top 1000 SNPs were the heavy-metal ATPase *MtHMA2* (Medtr2g036380) and *MtABCC3* (Medtr5g033320). We found a large region (∼2.3 Mb) on chromosome 8 (**Figure 6**) that contains a non-synonymous SNP in *MtCAX9* (Medtr8g085260) and several surrounding genes with ATP binding/ATPases, and transmembrane GO-terms (**Supplementary Table 8**). The SNPs in this region showed very high LD (r^2 = 0.62), suggesting a large locus with several linked genes that may be involved in Hg accumulation. This region overlaps with high F_{st} windows spanning the same chromosomal region, which indicates higher Hg accumulating genotypes are genetically differentiated in this large genomic region. Among the genes in the highest F_{st} peak included a duplicated cation amino acid transporter (four homologs: Medtr8g089320, Medtr8g089340, Medtr8g089342, Medtr8g089360) a magnesium-translocating *P*-type ATPase (Medtr8g089870), an ATPase amine-terminal autoinhibitory domain protein (Medtr8g090120) and an adjacent membrane calcium-translocating *P*-type ATPase (Medtr8g090125) (**Supplementary Table 9**). The population genomics methods provided complementary support that the

TABLE 2 | Genes in chromosome 2 genomic regions associated with Cd relative root growth (Cd RRG) that were identified using GWAS, F_{st} and XP-CLR.

Mt4.0 gene ID	BLAST gene function	Mt4.0 location
GWAS		
Medtr2g438720	ankyrin repeat plant-like protein	chr2:15654113..15660513
Medtr2g438740	ankyrin repeat plant-like protein	chr2:15661862..15664982
Medtr2g438760	ankyrin repeat plant-like protein	chr2:15678476..15683145
Medtr2g438700	ankyrin repeat plant-like protein	chr2:15644925..15649215
Medtr2g438560	hypothetical protein	chr2:15587523..15589330
Medtr2g438580	ankyrin repeat protein	chr2:15601177..15601431
Medtr2g438670	hypothetical protein	chr2:15639667..15639882
F_{st}		
Medtr2g436680	ABC transporter protein (ABCC14, AT3G62700.1)	chr2:14273804..14283871
Medtr2g436710	ABC transporter protein (ABCC14, AT3G62700.1)	chr2:14295748..14305736
Medtr2g436730	ABC transporter protein (ABCC14, AT3G62700.1)	chr2:14312635..14321998
Medtr2g436830	heavy metal transport (ATHMP10, AT1G56210.1)	chr2:14380972..14382489
Medtr2g437380	2OG-Fe(II) oxygenase family oxidoreductase	chr2:14380972..14382489
Medtr2g035840	heavy metal P-type ATPase (HMA7, AT5G44790.1)	chr2:15206300..15212503
Medtr2g438160	potassium transporter 12 (AT1G31120.1)	chr2:15394059..15398083
XP-CLR		
Medtr2g436680	ABC transporter protein (ABCC14, AT3G62700.1)	chr2:14273804..14283871
Medtr2g436710	ABC transporter protein (ABCC14, AT3G62700.1)	chr2:14295748..14305736
Medtr2g436730	ABC transporter protein (ABCC14, AT3G62700.1)	chr2:14312635..14321998
Medtr2g436830	heavy metal transport (ATHMP10, AT1G56210.1)	chr2:14380972..14382489
Medtr2g437380	2OG-Fe(II) oxygenase family oxidoreductase	chr2:14380972..14382489

Columns from left to right are Mt4.0 gene ID, gene function identified by BLAST to A. thaliana, and gene location with start and stop coordinates based on the Mt4.0 genome.

regions on chromosomes 2, 5, and 8 identified by GWAS contain a genomic signal that could be detected using only the genotypes with the most extreme phenotypes.

For Hg RRG, the top SNP was a non-synonymous substitution on chromosome 2 in a UDP-glycosyltransferase gene (Medtr2g083420). This gene was shown to be most highly expressed in roots in the A17 genotype (**Supplementary Table 2**). Among the other top SNPs for Hg RRG was a Fe (II)-dependent oxygenase superfamily protein (Medtr4g116010) and a peroxidase superfamily protein, both with metal ion binding and oxidoreductase activity GO-terms, and a cytochrome P450 gene (Medtr5g073020) with heme binding, membrane, and metal ion GO-terms. Three ATP binding genes all of which have membrane/transport ontologies, including H^+ ATPase 2 (Medtr1g084780), non-intrinsic ABC protein 12 (Medtr1g063920), and an ABC-2 type transporter family protein (Medtr4g076900) (**Table 1**), were also associated with Hg RRG.

Minor Allele Frequency and Effect Size

The distribution of effect sizes of all SNPs is centered around zero, while those with the lowest *p*-values are shifted significantly toward larger effect sizes (**Supplementary Figure 7**). For all four trait associations, we found negative correlations between MAF's and effect sizes (**Figure 7**), where SNPs with larger effect sizes had the tendency to be at low frequency. This is consistent with purifying selection acting to remove alleles that may be deleterious if they result in elevated toxicity after exposure to Cd or Hg. Because SNPs with the lowest *p*-values are inherently biased toward large effect sizes, we ran GWAS on 100 permuted phenotype datasets to generate a null distribution for each trait. For Cd leaf accumulation and RRG, the correlation coefficients from the empirical datasets were below 99 and 86 of the 100 permuted datasets, respectively (**Supplementary Figure 8**). For Hg leaf accumulation and RRG, the correlation coefficients from the empirical dataset coefficients were below only 37 and 52 of the permuted datasets. Among the four traits, we would only consider the correlation coefficient from Cd leaf accumulation to be statistically significant ($p = 0.01$). Nevertheless, the correlations between MAF and effect size in all four traits were consistently negative. We used the Tajima's D statistic to test for significant differences in the top SNPs detected in GWAS compared with the rest of the genome (i.e., "background SNPs"). We found significantly lower mean Tajima's D for the top SNPs from the GWAS compared with background SNPs for three of the four traits (**Figure 8**) due to an excess of low frequency alleles (i.e., low MAF's), consistent with purifying selection (Paape et al., 2013).

DISCUSSION

Phenotypic and Genetic Diversity

Heritability was sufficiently high for GWAS and our heritability estimates showed similar levels of phenotypic/genetic variability for the two RRG traits (about 10% difference) as well as for the leaf accumulation traits. Overall, heritability was higher for RRG than for leaf accumulation. Variability in Cd accumulation found in *M. truncatula* leaves was comparable to findings in *A. thaliana* (Chao et al., 2012), barley (Wu et al., 2015), *Brassica napus* (Chen et al., 2018), and rice (Zhao et al., 2018). Interestingly,

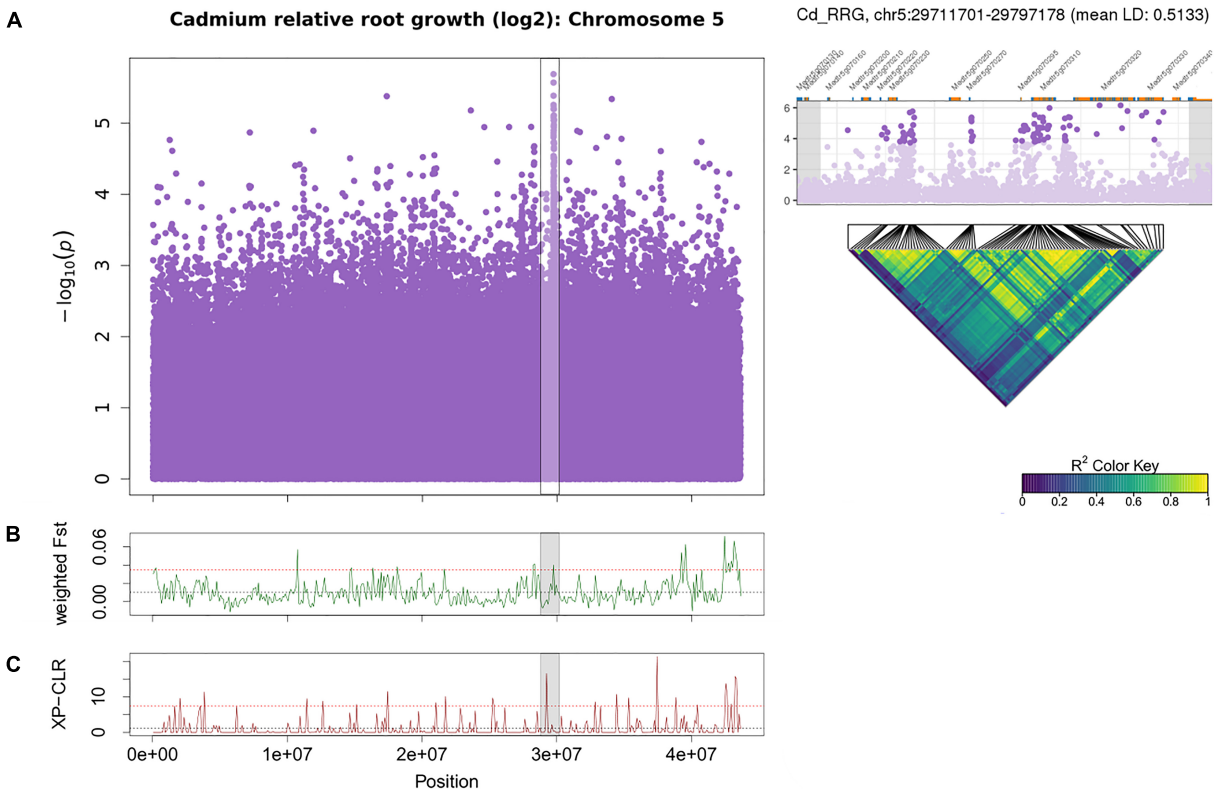


FIGURE 5 | Manhattan plot of Cd relative root growth (RRG) associations for chromosome 5 using GAPIT **(A)**. The region on chromosome 5 associated with Cd RRG contained *MtCAX3* (Medtr5g070330), *MtPDR3* (Medtr5g070320), *MtCPT7* (Medtr5g070270) and *MtDDB2* (Medtr5g070310). The SNPs in genic and intergenic regions are in high LD (pairwise linkage disequilibrium (LD) among the 100 most significant SNPs in the peak). Each point represents a SNP, the x-axis signifies the position on the chromosome while the *y*-axis is the $-\log_{10}$ of the *p*-value. The Cd RRG phenotype was $\log_2$ transformed to normalize the data prior to running the GWAS. The GWAS region is highlighted in gray the F_{st} **(B)** and XP-CLR **(C)** sliding window (100 kb windows) analyses. The *x*-axis represents the chromosomal position, the *y*-axis the F_{st} or XP-CLR-value of the window at the corresponding position. Higher F_{st} indicates population differentiation, higher XP-CLR indicates the presence of a selective sweep. The dashed black line represents the mean F_{st} or CLR-values for the chromosome, the dashed red line represents windows ≥ 5% of the highest F_{st} or CLR-values for the chromosome.

variation in Hg accumulation in leaves was about 5 times wider than Cd accumulation in *M. truncatula*. To our knowledge, the only other study that has measured Hg accumulation in a wild plant or crop association panel was in maize (Zhao et al., 2017). Our study showed significantly higher variation in leaf Hg accumulation than the maize study (∼270-fold versus ∼3.5-fold). However, Zhao et al. (2017) collected mature plant leaves under significantly different conditions, making direct comparisons between studies difficult. Our phenotyping approach also revealed no correlation between RRG and leaf accumulation in *M. truncatula*, indicating that high metal tolerance does not translate to higher metal accumulation in leaves (see **Figure 1**). It is therefore important to consider that resilience to the toxic heavy metals may not require that plant roots take up or accumulate these metals in the roots. Relative root growth is a measure of tolerance, while accumulation is a measure of root to shoot transport and storage. Moreover, roots serve to protect shoots from excess toxicity which is evident by the accumulation of higher levels of heavy metals in roots compared to shoots in non-hyperaccumulating species (Krämer, 2010; Paape et al., 2020). The strongest correlation was between the RRG of plants treated with Cd and Hg. Our findings are similar to a recent study comparing RRG responses to Cd and Hg treatments in a different collection of *M. truncatula* genotypes (r = 0.48) which did not include sequence data (García de la Torre et al., 2021). The positive correlation in RRG between the two metal treatments suggests that both heavy metals elicit a common stress response that inhibits root growth while retaining other metal-specific responses (i.e., the correlation, $r < 0.5$), but existing genetic diversity for root length may also be a contributing factor.

Candidate Genes Are Broadly Dispersed Across the Genome

In many cases the most significant SNPs were found in functionally relevant genes that were broadly dispersed throughout the genome. For this reason, we found it more meaningful to examine a wide range of *p*-values for each trait association because we expected RRG and leaf accumulation of Cd and Hg to be highly polygenic. This allowed us to look at gene annotations and biologically relevant GO-terms

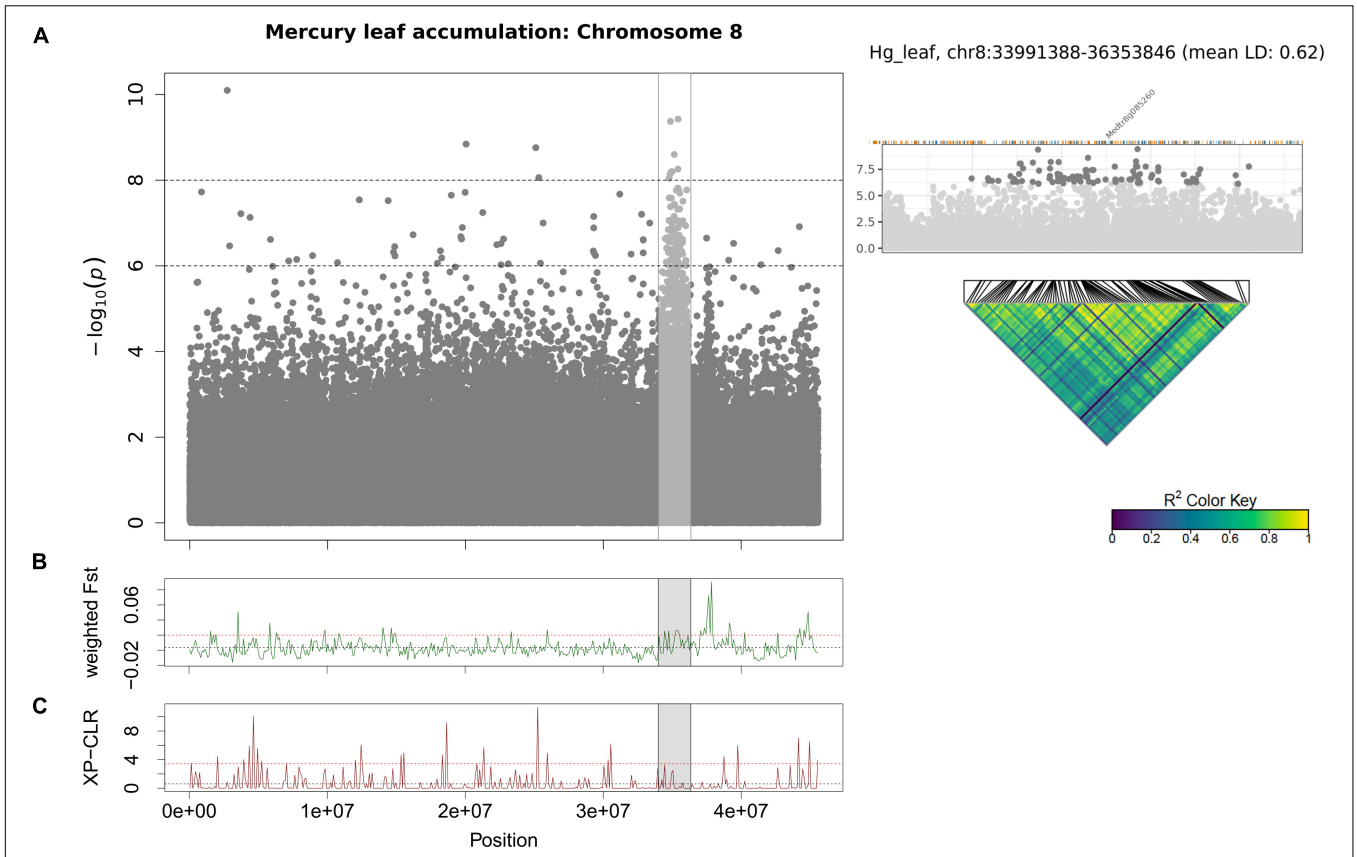


FIGURE 6 | Manhattan plot of Hg leaf accumulation for chromosome 8 using GEMMA **(A)**. Each point represents a SNP, the *x*-axis signifies the position on the chromosome while the *y*-axis is the $-\log_{10}$ of the *p*-value. The Cd RRG phenotype was $\log_2$ transformed to normalize the data prior to running the GWAS. The GWAS region is highlighted in gray the F_{st} **(B)** and XP-CLR **(C)** sliding window (100 kb windows) analyses. The *x*-axis represents the chromosomal position, the *y*-axis the F_{st} or XP-CLR-value of the window at the corresponding position. Higher F_{st} indicates population differentiation, higher XP-CLR indicates the presence of a selective sweep. The dashed black line represents the mean F_{st} or CLR-values for the chromosome, the dashed red line represents windows ≥ 5% of the highest F_{st} or CLR-values for the chromosome.

among a list of several hundred genes for each trait. Our findings for Cd accumulation are considerably different than a similar GWAS in *A. thaliana*, where a single genomic peak containing the *HMA3* locus was identified, which drowned out the signal at any other genomic region (Chao et al., 2012). We found clusters of SNPs in multiple genomic regions, all of which appeared to have several linked genes with functional relevance to metal transport or metal tolerance. In cases such as on chromosomes 2, 5, and 8, the genomic region of interest appears to constitute a large locus containing many functionally relevant genes. If these genes are functionally linked, they likely evolved together as a single locus, such as *MtCAX3* and *MtPDR3* on chromosome 5 which are clearly linked. Moreover, the detection of orthologs that have functionally relevant annotations indicates that our GWAS provided a realistic picture of the genetic architecture underlying metal tolerance traits in *M. truncatula*. Included among the genes identified in our genome wide statistical analyses were several stress-related genes, metal ion transporters, membrane transporters, and vacuole proteins with ion binding gene ontologies. Many of these genes are orthologous to known metal transporters studied in other plant species suggesting they are evolutionarily conserved.

A large category of candidate genes were ATP-dependent transporters that were associated with all traits, suggesting that ATP binding genes are important for Cd and Hg tolerance in *M. truncatula*. Included among these were multiple ATP binding cassette (ABC) proteins associated with each of the traits. Many of these are orthologous to ABC-transporters in *A. thaliana* (e.g., *AtABCC2*, *AtABCC3*, *AtABCC8*) and rice that are known to interact with metal-binding phytochelatins that chelate Cd and Hg ions allowing for transportation into vacuoles by membrane bound ABC-proteins (Park et al., 2012; Brunetti et al., 2015; Sun et al., 2018). Another class of ATP-transporters are the P_{1B}-type heavy metal ATPases (HMA's) which are well-known for their roles in root to shoot transport (Hanikenne et al., 2008) and compartmentalization of unchelated heavy metals (Morel et al., 2008). HMA's have been shown to play essential roles in Cd, Co, Pb and Zn transport in *A. thaliana* and rice (Chao et al., 2012; Takahashi et al., 2012). Thus far, HMAs have not been associated with Hg tolerance and transport in plants, instead relying on ABC-transporters (Park et al., 2012). In our study, we found

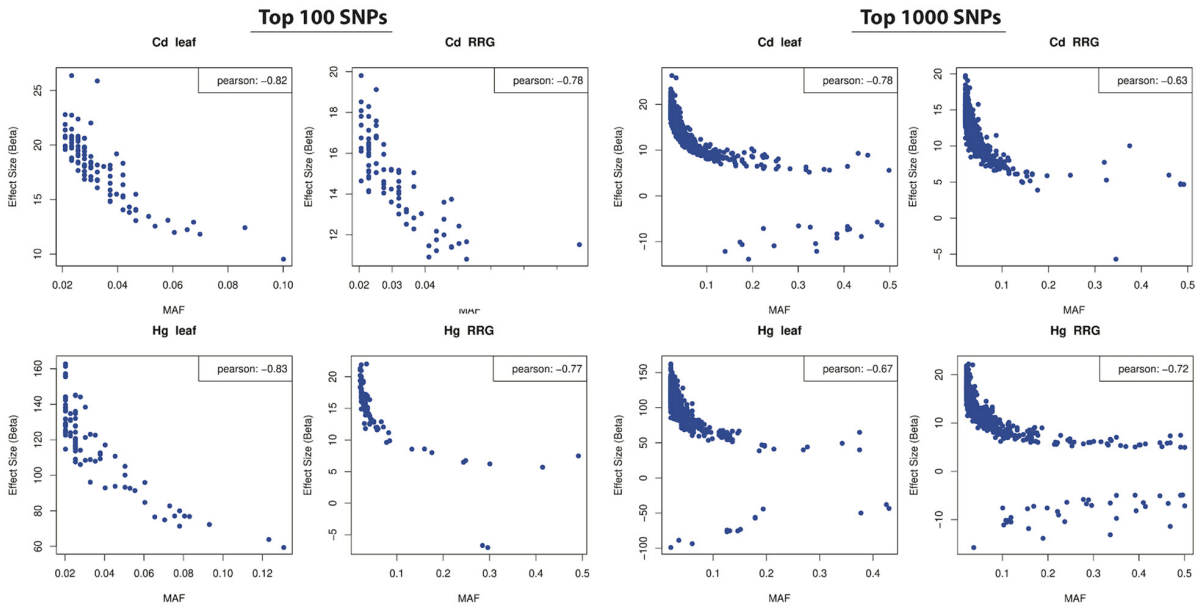


FIGURE 7 | Correlations between minor allele frequency (MAF; *x*-axis) and SNP effect size (*y*-axis) in the top 100 (left four panels) and top 1000 (right four panels) most significant SNPs (effect sizes estimated from GAPIT). Pearson's correlation coefficients shown in upper right corners.

MtHMA2 associated with leaf accumulation of Hg and a P_{1B}-type heavy metal ATPase (H^+-ATPase 2) associated with Hg RRG. Interestingly, no HMA's were associated with either of the Cd traits. Clearly, ABC- and ATPase transporters have widespread roles that range from general to specific metal responses in plants, which include transport of essential and non-essential metal ions (Takahashi et al., 2012).

The two most strongly associated candidate genes for Hg leaf accumulation were chloroplast and mitochondria membrane transporters, one of which was an aspartate aminotransferase ortholog previously identified in *A. thaliana* to be associated with transition metal ion binding on mitochondria and oxidative stress (Tan et al., 2010). Another highly significant SNP was in the Natural Resistance-Associated Macrophage Protein 6, *MtNRAMP6*. In rice and *A. thaliana*, *NRAMP's* are involved in aluminum (Xia et al., 2010) and manganese transport (Cailliatte et al., 2010). This gene has the highest expression in roots which may result in greater availability of Hg to be transported to the leaf leading to higher concentrations in the high Hg-accumulating genotypes. In *A. thaliana,* overexpressed *NRAMP6* resulted in Cd hypersensitivity through increased Cd uptake, while *nramp6* loss of function mutants resulted in reduced Cd uptake leading to increased plant tolerance when exposed to high levels of Cd (Cailliatte et al., 2009). Similarly, in rice, knock-out mutations of *NRAMP5* resulted in reduced Cd accumulation in roots and leaves (Sasaki et al., 2012). Previous GWAS studies in barley and rapeseed found that homologs of *NRAMP5* and *NRAMP6* were associated with Cd accumulation (Wu et al., 2015; Chen et al., 2018). While our study found an association between *NRAMP6* and Hg, but not Cd accumulation, our results provide supporting evidence that this family of genes contributes to tolerance and transport of an array of metal ions using similar mechanisms (Nevo and Nelson, 2006).

Large Genomic Regions Identified With Multiple Methods

We also found genomic regions with clusters of SNPs and genes that showed linkage for Cd tolerance and leaf accumulation of Hg, which aligned with regions also identified using population genomics methods. Using genotypes that were divergent for Cd RRG, the F_{st}, and XP-CLR statistics identified a region on chromosome 2 that containing a tandemly triplicated ABC-transporter (*MtABCC14*), the heavy metal transport/detoxification superfamily protein *AtHMP10*, and a Fe-oxidoreductase gene. Interestingly, this region has very high LD and contains several gene families with known metal tolerance roles and gene ontologies that support metal transport mechanisms or metal-induced stress responses. Expression enhancement through the triplication of genes is a common mechanism in metal ion transport in hyperaccumulating plants (Hanikenne et al., 2008; Shahzad et al., 2010) suggesting that the three linked *MtABCC14* copies may jointly contribute to tolerance to Cd. A second F_{st} peak containing the P_{1B}-type heavy metal ATPase *MtHMA7,* and a potassium transporter was found closer to the GWAS peak that contained ankyrin repeats. While these ankyrin genes do not show homology to *A. thaliana* orthologs, ankyrin repeat proteins were shown to mediate membrane bound protein-protein interactions that facilitated heavy metal transport in *A. thaliana* roots (Gao et al., 2009). In a GWAS study on salinity tolerance in *M. truncatula,* chromosome 2 was previously shown to have many loci associated with ion stress, with the closest peak being approximately 800 kb away

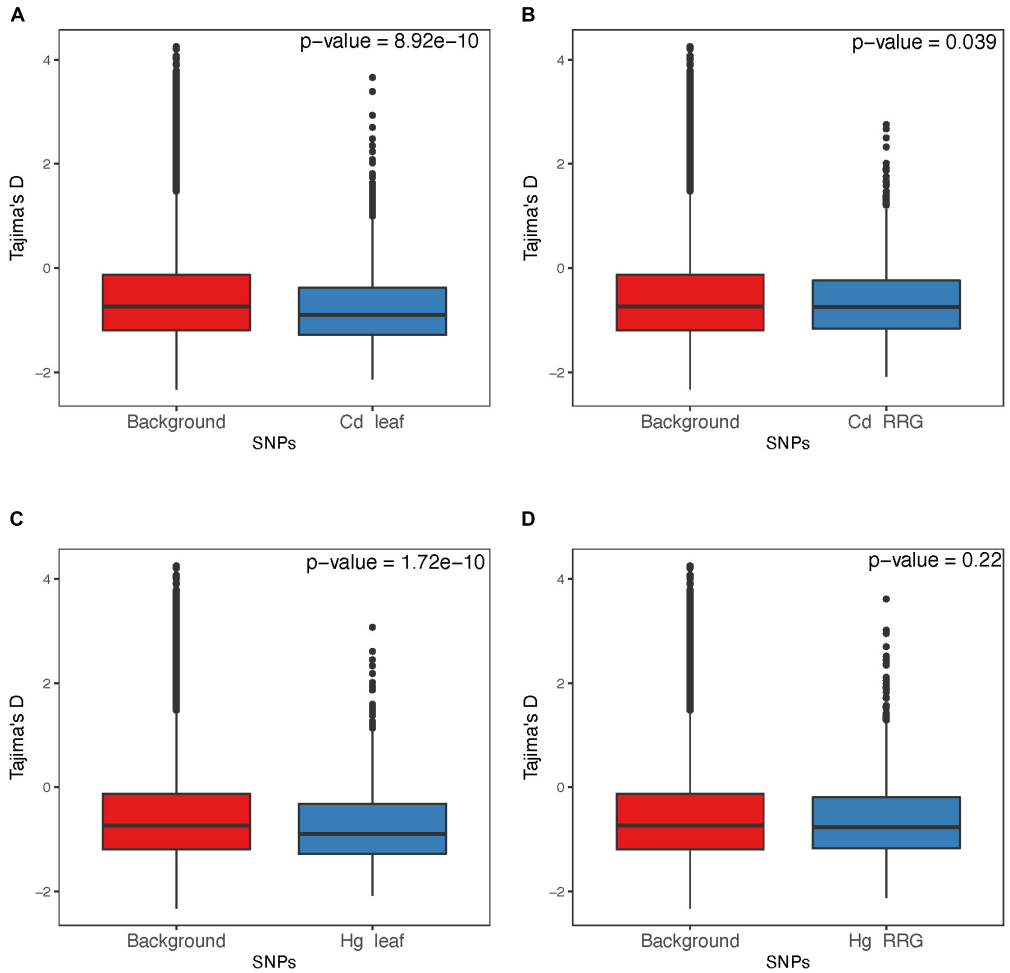


FIGURE 8 | Comparison of Tajima's D from the genomic background and regions containing SNPs from the GWAS for the traits Cd leaf accumulation **(A)**, Cd relative root growth (RRG) **(B)**, Hg leaf accumulation **(C)**, and Hg RRG **(D)**. Tajima's D was estimated using 10 bp regions surrounding focal SNPs.

from the peak identified here (Kang et al., 2019). We propose that this arm of chromosome 2 has clusters of specific and general ion transporters that influence both toxic and essential ions in *M. truncatula*.

The region on chromosome 5 associated with Cd RRG was identified using both GWAS and F_{st} statistics, which contains the genes *MtCAX3* and *MtPDR3* which are located < 6 kb apart. Homologs of these two genes have been shown to have substantial roles in heavy metal tolerance. The cation exchanger genes *CAX1* and *CAX3* are transporters located on the tonoplast. Upregulation of *CAX3* was reported to increase Cd tolerance in *A. thaliana* by sequestering Cd into the vacuole (Cheng et al., 2005). Expression levels of *CAX3* are upregulated by Hb1 (class 1 hemoglobin), which was shown to inhibit the expression of *IRT1* and *PDR8* (Bahmani et al., 2019). The ATP-binding cassette transmembrane transporter *PDR3* is an ortholog of *PDR8*. Inhibition of *PDR8* using RNAi was shown to increase sensitivity to Cd and overexpression increased resistance to Cd in *A. thaliana* (Kim et al., 2007). Because *PDR3* and *PDR8* are both ABC-transmembrane proteins, we assume their function is similar, as we have seen for other ABC-proteins involved in heavy metal transport to vacuoles. Another gene that is near the linked *MtCAX3* and *MtPDR3* genes is *MtCPT7,* which is homologous to *AtCPT7* (also known as *AtCPT4*). *AtCPT7* is involved in dolichol synthesis and was shown to be upregulated after exposure to Cd in *A. thaliana* (Jozwiak et al., 2017). Because *MtCAX3*, *MtPDR3,* and *MtCPT7* have biologically plausible roles in Cd transport, their tight linkage suggests they may be collectively involved in Cd tolerance.

To date, few QTL associated with Hg accumulation have been identified in crops or crop relatives (Wang et al., 2013; Zhao et al., 2017). Using GWAS, we found a large genomic region of interest on chromosome 8 in *M. truncatula* that had several transmembrane and ATPase transporters in a region surrounding *MtCAX9*. As discussed above, the *CAX* transporters have been found to be located on vacuole membranes, and the large number of membrane proteins and transmembrane related gene ontologies (24 genes spanning ∼1.2 Mb) suggests this genomic region possesses biologically relevant candidate genes involved in transport of Hg across membranes, including

into vacuoles. On this chromosome, the highest F_{st} peak that is upstream from the GWAS peak includes multiple calcium and magnesium related P-type ATPases. This finding fits the recurring theme that *P*-type ATPases are involved in multiple metal ion transport mechanisms, including for highly toxic ions, such as Hg. Chromosome 8 contains some highly promising candidate genes to pursue using functional genetics approaches such as knock-out mutations using Hg accumulation to quantify the effects of mutants.

Stress Related Genetic Responses

Together, cellular responses that increase tolerance or accumulation to heavy metals are complex and involve not only ion transporters, but also mechanisms that deal with toxicity such as oxidative stress and ROS production. In excess amounts, all heavy metals trigger oxidative stress responses, resulting in increased levels of ROS in cells and tissues. As storage capacity is exceeded and toxic ions reach the leaves, plants alter leaf structure and chemistry leading to an accumulation of hydroxyl radicals. These hydroxyl radicals will indiscriminately oxidize biomolecules, including vital proteins and plant enzymes. In *M. truncatula*, Cd tolerance has been associated with the efficiency of the plant's antioxidant defenses (Rahoui et al., 2014). The strong association with multiple Fe-oxidoreductases with all four traits suggests that plant tissues are likely under oxidative stress in our study. Moreover, *MtOSX2,* which has metal ion binding gene ontology, is associated with Cd leaf accumulation and may also play a role in dealing with oxidative stress responses to heavy metals (Kawai et al., 2014). In *A. thaliana,* the *AtOXS2* gene (also known as *DEG9*) is a zinc-finger transcription factor shown to be triggered by salt stress and in maize (*Zea mays*), *ZmOXS2* homologs are able to enhance Cd tolerance (He et al., 2016; Jing et al., 2019). The most strongly associated SNP for Hg RRG was a non-synonymous substitution in a UDP-glycosyltransferase gene (Medtr2g083420), which is known to have highest expression in the roots of the A17 ecotype. UDP-glycotransferase mutants of *AtUGGT* in *A. thaliana* were shown to have delayed root growth induced by ion stress, which involved disruption of endoplasmic reticulum transmembrane integrity (Blanco-Herrera et al., 2015). In *M. truncatula,* UDP-glycosyltransferases are involved in the biosynthesis of plant secondary metabolites, including flavonoids and isoflavonoids (Modolo et al., 2007). Flavonoids are crucial antioxidants that can efficiently scavenge the harmful ROS generated by heavy metal toxicity. They have also been reported to act as metal chelators (Kumar and Prasad, 2018).

Standing Variation and Genetic Architecture

The top SNPs detected in our GWAS showed a strong negative relationship between MAF and effect size, consistent with expectations of mutation-selection balance (Barton and Keightley, 2002) and stabilizing selection (Josephs et al., 2017). While some mutations may result in alleles that confer greater accumulation or tolerance to toxic heavy metals that could be beneficial, we expect that most mutations that result in excess uptake of heavy metals to be deleterious. This explains why alleles with large effect that were associated with high tolerance or accumulation were at low frequencies in the population/species because we expect purifying selection to act on these mutations. The SNPs with large effect that are associated with high tolerance or accumulation but are at higher frequencies, would be expected to be more beneficial because they resulted in exclusion of Cd or Hg, and would therefore be more common in the species. However, a large amount of standing genetic variation, including SNPs associated with higher tolerance or accumulation, may provide the necessary polymorphism for species to colonize or adapt to new habitats (Hermisson, 2005; Barrett and Schluter, 2008), suggesting potential benefits of these alleles in some environments. Hence, trade-offs between beneficial and deleterious mutations are expected under models of stabilizing or balancing selection which favors intermediate phenotypes and allele frequencies. It is also likely that alleles conferring greater tolerance or accumulation of Cd or Hg have been tested in natural or contaminated environments at some point throughout the species history and may have enabled a broader species distribution. Moreover, many alleles found in our GWAS to be associated with these specific heavy metals, may be generally associated with stressful levels of ion contamination due to poor quality soils and not necessarily due to direct heavy metal contamination from mining in the very recent past.

CONCLUSION

By combining GWAS analysis with population genomics, we were able to find signals of trait associations that overlapped with genomic regions detected by GWAS. Our population genomics approach used the tails of quantitative trait distributions to provide a contrast for genome-wide scans of genomic divergence, similar to the QTL approach described by Stinchcombe and Hoekstra (2007), but using the natural variation in the Medicago HapMap panel. With this in mind, it could be feasible to sequence only those lines in the ends of the phenotypic distributions in very large collections of genotypes (e.g., García de la Torre et al., 2013, 2021), and find meaningful genetic loci underlying the trait differences without sequencing an entire GWAS panel. Our approach to identify genomics regions using F_{st} and XP-CLR was similar to Long et al. (2013) but unlike their study, we were agnostic to the sub-population assignment and we selected plant genotypes based solely on extreme phenotype categories. In this sense, we feel that the signal that we detected using these methods, represented selection at the species level rather than the population level because the signal was present using genotypes in the same phenotype category that may belong to multiple sub-population assignments. Our approach likely depends on phenotyping a rather large collection of individual genotypes to ensure that the ends of the phenotypic distributions are truly representative of the species level ranges in trait values. Genome-wide association mapping and population genomic scans revealed that leaf and root responses to toxic heavy metals are complex polygenic traits involving many quantitative trait loci. We found ABC-transporters and P-type ATPases

to be commonly associated with Cd and Hg tolerance, which is supported by previous functional studies among members of these gene families playing large roles in heavy metal transport across membranes and isolation in vacuoles. While genetic architect may vary widely across plant species to due to the spatial and genomic distribution of genes involved in metal ion transport, accumulation, and detoxification, gene function of the families of transporters and stress response genes seems to be widely conserved. Low minor allele frequencies of large effect SNPs associated with higher tolerance or accumulation of Cd and Hg are consistent with purifying selection against deleterious mutations. Large effect alleles associated with low tolerance or accumulation were more common in our dataset, suggesting excluder alleles may be advantageous. We suspect similar evolutionary principles and processes would apply to other species as well. Our findings therefore suggest functionally conserved roles in these gene families, and potential conservation at macroevolutionary scales in plants, which is important for breeding and transgenic modifications in crop species.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

TP and JP conceived of the study. JP supervised laboratory and phenotypic experiments. MS and ML conducted phenotypic measurements and data analysis. TP and BH conducted bioinformatics, GWAS, population genomics, and statistical analyses and drafted the manuscript, with editorial contributions from all authors. MC conducted bioinformatics and statistical analyses.

FUNDING

This work was supported by Brookhaven National Laboratory's Program Development and by the U.S. Department of Energy, Office of Science, Office of Biological, and Environmental Research, as part of the Quantitative Plant Science Initiative. The phenotyping work was supported by grants from CSIC (i-LINK1093), MINECO (AGL2013-40758-R), and Agencia Estatal de Investigación, AEI, Spain (AGL2017-88381-R) awarded to JP.

ACKNOWLEDGMENTS

We thank Nevin Young and Roxanne Denny at the University of Minnesota for germplasm, and their support and discussions regarding the Medicago HapMap collection. We also thank Kentaro Shimizu for useful discussions and the University of Zurich Functional Genomics Center for computational support. We also thank Gwyneth Halstead-Nussloch for helpful discussion and assistance with software and scripts.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.806949/full#supplementary-material

REFERENCES

Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19, 1655–1664. doi: 10.1101/gr.094052.109

Alloway, B. J. (2013). "Sources of heavy metals and metalloids in soils," in *Heavy Metals in Soils: Trace Metals and Metalloids in Soils and their Bioavailability Environmental Pollution*, ed. B. J. Alloway (Dordrecht: Springer), 11–50. doi: 10.1007/978-94-007-4470-7_2

Arregui, G., Hipólito, P., Pallol, B., Lara-Dampier, V., García-Rodríguez, D., Varela, H. P., et al. (2021). Mercury-tolerant *Ensifer medicae* strains display high mercuric reductase activity and a protective effect on nitrogen fixation in *Medicago truncatula* nodules under mercury stress. *Front. Plant Sci.* 11:560768. doi: 10.3389/fpls.2020.560768

Bahmani, R., Kim, D., Na, J., and Hwang, S. (2019). Expression of the tobacco non-symbiotic class 1 hemoglobin gene Hb1 reduces cadmium levels by modulating Cd transporter expression through decreasing nitric oxide and ROS level in *Arabidopsis*. *Front. Plant Sci.* 10:201. doi: 10.3389/fpls.2019.00201

Barrett, R., and Schluter, D. (2008). Adaptation from standing genetic variation. *Trends Ecol. Evol.* 23, 38–44. doi: 10.1016/j.tree.2007.09.008

Barton, N. H., and Keightley, P. D. (2002). Understanding quantitative genetic variation. *Nat. Rev. Genet.* 3, 11–21. doi: 10.1038/nrg700

Bernhoft, R. A. (2012). Mercury toxicity and treatment: a review of the literature. *J. Environ. Public Health* 2012, 1–10. doi: 10.1155/2012/460508

Blanco-Herrera, F., Moreno, A. A., Tapia, R., Reyes, F., Araya, M., D'Alessio, C., et al. (2015). The UDP-glucose: glycoprotein glucosyltransferase (UGGT), a key enzyme in ER quality control, plays a significant role in plant growth as well as biotic and abiotic stress in *Arabidopsis thaliana*. *BMC Plant Biol.* 15:127. doi: 10.1186/s12870-015-0525-2

Bonhomme, M., Andr, O., Badis, Y., Ronfort, J., Burgarella, C., Chantret, N., et al. (2014). High-density genome-wide association mapping implicates an F-box encoding gene *in Medicago truncatula* resistance to *Aphanomyces euteiches*. *New Phytol.* 201, 1328–1342. doi: 10.1111/nph.12611

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308

Branca, A., Paape, T. D., Zhou, P., Briskine, R., Farmer, A. D., Mudge, J., et al. (2011). PNAS Plus: whole-genome nucleotide diversity, recombination, and linkage disequilibrium in the model legume *Medicago truncatula*. *Proc. Natl. Acad. Sci. U.S.A.* 108, E864–E870. doi: 10.1073/pnas.1104032108

Browning, B. L., and Browning, S. R. (2016). Genotype imputation with millions of reference samples. *Am. J. Hum. Genet.* 98, 116–126. doi: 10.1016/j.ajhg.2015.11.020

Brunetti, P., Zanella, L., De Paolis, A., Di Litta, D., Cecchetti, V., Falasca, G., et al. (2015). Cadmium-inducible expression of the ABC-type transporter AtABCC3 increases phytochelatin-mediated cadmium tolerance in *Arabidopsis*. *J. Exp. Bot.* 66, 3815–3829. doi: 10.1093/jxb/erv185

Cailliatte, R., Lapeyre, B., Briat, J.-F., Mari, S., and Curie, C. (2009). The NRAMP6 metal transporter contributes to cadmium toxicity. *Biochem. J.* 422, 217–228. doi: 10.1042/BJ20090655

Cailliatte, R., Schikora, A., Briat, J.-F., Mari, S., and Curie, C. (2010). High-affinity manganese uptake by the metal transporter NRAMP1 is essential for

Arabidopsis growth in low manganese conditions. *Plant Cell* 22, 904–917. doi: 10.1105/tpc.109.073023

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. doi: 10.1186/1471-2105-10-421

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., and Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 4:7. doi: 10.1186/s13742-015-0047-8

Chao, D.-Y., Silva, A., Baxter, I., Huang, Y. S., Nordborg, M., Danku, J., et al. (2012). Genome-wide association studies identify heavy metal ATPase3 as the primary determinant of natural variation in leaf cadmium in *Arabidopsis thaliana*. *PLoS Genet.* 8:e1002923. doi: 10.1371/journal.pgen.1002923

Chen, H., Patterson, N., and Reich, D. (2010). Population differentiation as a test for selective sweeps. *Genome Res.* 20, 393–402. doi: 10.1101/gr.100545.109

Chen, L., Wan, H., Qian, J., Guo, J., Sun, C., Wen, J., et al. (2018). Genome-wide association study of cadmium accumulation at the seedling stage in rapeseed (*Brassica napus* L.). *Front. Plant Sci.* 9:375. doi: 10.3389/fpls.2018.00375

Cheng, N.-H., Pittman, J. K., Shigaki, T., Lachmansingh, J., LeClere, S., Lahner, B., et al. (2005). Functional association of *Arabidopsis* CAX1 and CAX3 is required for normal growth and ion homeostasis. *Plant Physiol.* 138, 2048–2060. doi: 10.1104/pp.105.061218

Clemens, S. (2006). Toxic metal accumulation, responses to exposure and mechanisms of tolerance in plants. *Biochimie* 88, 1707–1719. doi: 10.1016/j.biochi.2006.07.003

Clemens, S., and Ma, J. F. (2016). Toxic heavy metal and metalloid accumulation in crop plants and foods. *Annu. Rev. Plant Biol.* 67, 489–512. doi: 10.1146/annurev-arplant-043015-112301

Coba de la Peña, T., and Pueyo, J. J. (2012). Legumes in the reclamation of marginal soils, from cultivar and inoculant selection to transgenic approaches. *Agron. Sustain. Dev.* 32, 65–91. doi: 10.1007/s13593-011-0024-2

Cobbett, C. S. (2001). Heavy metal detoxification in plants: phytochelatin biosynthesis and function. *IUBMB Life* 51, 183–188. doi: 10.1080/152165401753544250

Crow, J. F., and Kimura, M. (2010). *An Introduction to Population Genetics Theory*. Jodhpur: Scientific Publisher.

Curtin, S. J., Tiffin, P., Guhlin, J., Trujillo, D. I., Burghardt, L. T., Atkins, P., et al. (2017). Validating genome-wide association candidates controlling quantitative variation in nodulation. *Plant Physiol.* 173, 921–931. doi: 10.1104/pp.16.01923

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158. doi: 10.1093/bioinformatics/btr330

de Abreu-Neto, J. B., Turchetto-Zolet, A. C., de Oliveira, L. F. V., Bodanese Zanettini, M. H., and Margis-Pinheiro, M. (2013). Heavy metal-associated isoprenylated plant protein (HIPP): characterization of a family of proteins exclusive to plants. *FEBS J.* 280, 1604–1616. doi: 10.1111/febs.12159

Du, Z., Zhou, X., Ling, Y., Zhang, Z., and Su, Z. (2010). agriGO: a GO analysis toolkit for the agricultural community. *Nucleic Acids Res.* 38, W64–W70. doi: 10.1093/nar/gkq310

Eyre-Walker, A., and Keightley, P. D. (2007). The distribution of fitness effects of new mutations. *Nat. Rev.s Genet.* 8, 610–618. doi: 10.1038/nrg2146

Gao, W., Xiao, S., Li, H.-Y., Tsao, S.-W., and Chye, M.-L. (2009). *Arabidopsis thaliana* acyl-CoA-binding protein ACBP2 interacts with heavy-metal-binding farnesylated protein AtFP6. *New Phytol.* 181, 89–102. doi: 10.1111/j.1469-8137.2008.02631.x

García de la Torre, V. S., Coba de la Peña, T., Lucas, M. M., and Pueyo, J. J. (2013). Rapid screening of *Medicago truncatula* germplasm for mercury tolerance at the seedling stage. *Environ. Exp. Bot.* 91, 90–96. doi: 10.1016/j.envexpbot.2013.03.004

García de la Torre, V. S., Coba de la Peña, T., Pueyo, J. J., and Lucas, M. M. (2021). Cadmium-tolerant and -sensitive cultivars identified by screening of *Medicago truncatula* germplasm display contrasting responses to cadmium stress. *Front. Plant Sci.* 12:595001. doi: 10.3389/fpls.2021.595001

Gentzbittel, L., Ben, C., Mazurier, M., Shin, M.-G., Lorenz, T., Rickauer, M., et al. (2019). WhoGEM: an admixture-based prediction machine accurately predicts quantitative functional traits in plants. *Genome Biol.* 20:106. doi: 10.1186/s13059-019-1697-0

Hanikenne, M., Talke, I. N., Haydon, M. J., Lanz, C., Nolte, A., Motte, P., et al. (2008). Evolution of metal hyperaccumulation required cis-regulatory changes and triplication of HMA4. *Nature* 453, 391–395. doi: 10.1038/nature06877

He, L., Ma, X., Li, Z., Jiao, Z., Li, Y., and Ow, D. W. (2016). Maize OXIDATIVE STRESS2 homologs enhance cadmium tolerance in *Arabidopsis* through activation of a putative SAM-dependent methyltransferase gene. *Plant Physiol.* 171, 1675–1685. doi: 10.1104/pp.16.00220

Hermisson, J. (2005). Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics* 169, 2335–2352. doi: 10.1534/genetics.104.036947

Holsinger, K. E., and Weir, B. S. (2009). Genetics in geographically structured populations: defining, estimating and interpreting FST. *Nat. Rev. Genet.* 10, 639–650. doi: 10.1038/nrg2611

Hossain, M. A., Piyatida, P., da Silva, J. A. T., and Fujita, M. (2012). Molecular mechanism of heavy metal toxicity and tolerance in plants: central role of glutathione in detoxification of reactive oxygen species and methylglyoxal and in heavy metal chelation. *J. Bot.* 2012, 1–37. doi: 10.1155/2012/872875

Hutter, S., Vilella, A. J., and Rozas, J. (2006). Genome-wide DNA polymorphism analyses using VariScan. *BMC Bioinformatics* 7:409. doi: 10.1186/1471-2105-7-409

Järup, L., and Åkesson, A. (2009). Current status of cadmium as an environmental health problem. *Toxicol. Appl. Pharmacol.* 238, 201–208. doi: 10.1016/j.taap.2009.04.020

Jing, Y., Shi, L., Li, X., Zheng, H., Gao, J., Wang, M., et al. (2019). OXS2 is required for salt tolerance mainly through associating with salt inducible genes, CA1 and Araport11, in *Arabidopsis*. *Sci. Rep.* 9:20341. doi: 10.1038/s41598-019-56456-1

Josephs, E. B., Lee, Y. W., Stinchcombe, J. R., and Wright, S. I. (2015). Association mapping reveals the role of purifying selection in the maintenance of genomic variation in gene expression. *Proc. Natl. Acad. Sci.U.S.A.* 112, 15390–15395. doi: 10.1073/pnas.1503027112

Josephs, E. B., Stinchcombe, J. R., and Wright, S. I. (2017). What can genome-wide association studies tell us about the evolutionary forces maintaining genetic variation for quantitative traits? *New Phytol.* 214, 21–33. doi: 10.1111/nph.14410

Jozwiak, A., Lipko, A., Kania, M., Danikiewicz, W., Surmacz, L., Witek, A., et al. (2017). Modeling of dolichol mass spectra isotopic envelopes as a tool to monitor isoprenoid biosynthesis. *Plant Physiol.* 174, 857–874. doi: 10.1104/pp.17.00036

Kang, Y., Torres-Jerez, I., An, Z., Greve, V., Huhman, D., Krom, N., et al. (2019). Genome-wide association analysis of salinity responsive traits in *Medicago truncatula*. *Plant Cell Environ.* 42, 1513–1531. doi: 10.1111/pce.13508

Kawai, Y., Ono, E., and Mizutani, M. (2014). Evolution and diversity of the 2-oxoglutarate-dependent dioxygenase superfamily in plants. *Plant J.* 78, 328–343. doi: 10.1111/tpj.12479

Kevei, Z., Vinardell, J. M., Kiss, G. B., Kondorosi, A., and Kondorosi, E. (2002). Glycine-rich proteins encoded by a nodule-specific gene family are implicated in different stages of symbiotic nodule development in *Medicago* Spp. *Mol. Plant Microbe Interact.* 15, 922–931. doi: 10.1094/MPMI.2002.15.9.922

Khan, I. U., Rono, J. K., Zhang, B. Q., Liu, X. S., Wang, M. Q., Wang, L. L., et al. (2019). Identification of novel rice (*Oryza sativa*) HPP and HIPP genes tolerant to heavy metal toxicity. *Ecotoxicol. Environ. Saf.* 175, 8–18. doi: 10.1016/j.ecoenv.2019.03.040

Kim, D.-Y., Bovet, L., Maeshima, M., Martinoia, E., and Lee, Y. (2007). The ABC transporter AtPDR8 is a cadmium extrusion pump conferring heavy metal resistance: role of AtPDR8 in cadmium resistance. *Plant J.* 50, 207–218. doi: 10.1111/j.1365-313X.2007.03044.x

Krämer, U. (2010). Metal hyperaccumulation in plants. *Annu. Rev. Plant Biol.* 61, 517–534. doi: 10.1146/annurev-arplant-042809-112156

Krishnakumar, V., Kim, M., Rosen, B. D., Karamycheva, S., Bidwell, S. L., Tang, H., et al. (2015). MTGD: the *Medicago truncatula* genome database. *Plant Cell Physiol.* 56:e1. doi: 10.1093/pcp/pcu179

Kumar, A., and Prasad, M. N. V. (2018). Plant-lead interactions: transport, toxicity, tolerance, and detoxification mechanisms. *Ecotoxicol. Environ. Saf.* 166, 401–418. doi: 10.1016/j.ecoenv.2018.09.113

Lee, H.-K., Mysore, K. S., and Wen, J. (2018). "Tnt1 insertional mutagenesis in *Medicago truncatula*," in *Functional Genomics in Medicago truncatula Methods in Molecular Biology*, eds L. A. Cañas and J. P. Beltrán (New York, NY: Springer), 107–114. doi: 10.1007/978-1-4939-8633-0_7

León-Mediavilla, J., Senovilla, M., Montiel, J., Gil-Díez, P., Saez, Á, Kryvoruchko, I. S., et al. (2018). MtMTP2-facilitated zinc transport into intracellular compartments is essential for nodule development in *Medicago truncatula*. *Front. Plant Sci.* 9:990. doi: 10.3389/fpls.2018.00990

Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., et al. (2012). GAPIT: genome association and prediction integrated tool. *Bioinformatics* 28, 2397–2399. doi: 10.1093/bioinformatics/bts444

Long, Q., Rabanal, F. A., Meng, D., Huber, C. D., Farlow, A., Platzer, A., et al. (2013). Massive genomic variation and strong selection in *Arabidopsis thaliana* lines from Sweden. *Nat. Genet.* 45, 884–890. doi: 10.1038/ng.2678

Martinoia, E., Mimura, T., Hara-Nishimura, I., and Shiratake, K. (2018). The multifaceted roles of plant vacuoles. *Plant Cell Physiol.* 59, 1285–1287. doi: 10.1093/pcp/pcy113

Modolo, L. V., Blount, J. W., Achnine, L., Naoumkina, M. A., Wang, X., and Dixon, R. A. (2007). A functional genomics approach to (iso)flavonoid glycosylation in the model legume *Medicago truncatula*. *Plant Mol. Biol.* 64, 499–518. doi: 10.1007/s11103-007-9167-6

Morel, M., Crouzet, J., Gravot, A., Auroy, P., Leonhardt, N., Vavasseur, A., et al. (2008). AtHMA3, a P1B-ATPase allowing Cd/Zn/Co/Pb vacuolar storage in *Arabidopsis*. *Plant Physiol.* 149, 894–904. doi: 10.1104/pp.108.130294

Nevo, Y., and Nelson, N. (2006). The NRAMP family of metal-ion transporters. *Biochim. Biophys. Acta Mol. Cell Res.* 1763, 609–620. doi: 10.1016/j.bbamcr.2006.05.007

Nicholson, F. A., Smith, S. R., Alloway, B. J., Carlton-Smith, C., and Chambers, B. J. (2003). An inventory of heavy metals inputs to agricultural soils in England and Wales. *Sci. Total Environ.* 311, 205–219. doi: 10.1016/S0048-9697(03)00139-6

Ohkama-Ohtsu, N., Zhao, P., Xiang, C., and Oliver, D. J. (2007). Glutathione conjugates in the vacuole are degraded by γ-glutamyl transpeptidase GGT3 in *Arabidopsis*: GGT3 degrades GSH conjugates in the vacuole. *Plant J.* 49, 878–888. doi: 10.1111/j.1365-313X.2006.03005.x

Paape, T. (2020). "Patterns of polymorphism, recombination, and selection in *Medicago truncatula*," in *The Model Legume Medicago truncatula*, ed. F. de Bruijn (Hoboken, NJ: Wiley), 903–910. doi: 10.1002/9781119409144.ch115

Paape, T., Akiyama, R., Cereghetti, T., Onda, Y., Hirao, A. S., Kenta, T., et al. (2020). Experimental and field data support range expansion in an allopolyploid *Arabidopsis* owing to parental legacy of heavy metal hyperaccumulation. *Front. Genet.* 11:565854. doi: 10.3389/fgene.2020.565854

Paape, T., Bataillon, T., Zhou, P., Kono, J. Y., Briskine, R., Young, N. D., et al. (2013). Selection, genome-wide fitness effects and evolutionary rates in the model legume *Medicago truncatula*. *Mol. Ecol.* 22, 3525–3538. doi: 10.1111/mec.12329

Paape, T., Zhou, P., Branca, A., Briskine, R., Young, N., and Tiffin, P. (2012). Fine-scale population recombination rates, hotspots, and correlates of recombination in the *Medicago truncatula* genome. *Genome Biol. Evol.* 4, 726–737. doi: 10.1093/gbe/evs046

Park, J., Song, W.-Y., Ko, D., Eom, Y., Hansen, T. H., Schiller, M., et al. (2012). The phytochelatin transporters AtABCC1 and AtABCC2 mediate tolerance to cadmium and mercury: ABC transporters for PC-dependent Cd and Hg tolerance. *Plant J.* 69, 278–288. doi: 10.1111/j.1365-313X.2011.04789.x

Peng, H., Chen, Y., Weng, L., Ma, J., Ma, Y., Li, Y., et al. (2019). Comparisons of heavy metal input inventory in agricultural soils in North and South China: a review. *Sci. Total Environ.* 660, 776–786. doi: 10.1016/j.scitotenv.2019.01.066

Peralta-Videa, J. R., Lopez, M. L., Narayan, M., Saupe, G., and Gardea-Torresdey, J. (2009). The biochemistry of environmental heavy metal uptake by plants: implications for the food chain. *Int. J. Biochem.Cell Biol.* 41, 1665–1677. doi: 10.1016/j.biocel.2009.03.005

Punshon, T., Hirschi, K., Yang, J., Lanzirotti, A., Lai, B., and Guerinot, M. L. (2012). The role of CAX1 and CAX3 in elemental distribution and abundance in *Arabidopsis* seed. *Plant Physiol.* 158, 352–362. doi: 10.1104/pp.111.184812

Rahoui, S., Ben, C., Chaoui, A., Martinez, Y., Yamchi, A., Rickauer, M., et al. (2014). Oxidative injury and antioxidant genes regulation in cadmium-exposed radicles of six contrasted *Medicago truncatula* genotypes. *Environ. Sci. Pollut. Res.* 21, 8070–8083. doi: 10.1007/s11356-014-2718-x

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., et al. (2011). Integrative genomics viewer. *Nat. Biotechnol.* 29, 24–26. doi: 10.1038/nbt.1754

Sasaki, A., Yamaji, N., Yokosho, K., and Ma, J. F. (2012). Nramp5 Is a major transporter responsible for manganese and cadmium uptake in rice. *Plant Cell* 24, 2155–2167. doi: 10.1105/tpc.112.096925

Shahzad, Z., Gosti, F., Frérot, H., Lacombe, E., Roosens, N., Saumitou-Laprade, P., et al. (2010). The five AhMTP1 zinc transporters undergo different evolutionary fates towards adaptive evolution to zinc tolerance in *Arabidopsis halleri*. *PLoS Genet.* 6:e1000911. doi: 10.1371/journal.pgen.1000911

Shin, J.-H., Blay, S., McNeney, B., and Graham, J. (2006). LDheatmap: an R function for graphical display of pairwise linkage disequilibria between single nucleotide polymorphisms. *J. Stat. Softw.* 16, 1–9. doi: 10.18637/jss.v016.c03

Stanton-Geddes, J., Paape, T., Epstein, B., Briskine, R., Yoder, J., Mudge, J., et al. (2013). Candidate genes and genetic architecture of symbiotic and agronomic traits revealed by whole-genome, sequence-based association genetics in *Medicago truncatula*. *PLoS One* 8:e65688. doi: 10.1371/journal.pone.0065688

Stinchcombe, J. R., and Hoekstra, H. E. (2007). Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity* 100, 158–170. doi: 10.1038/sj.hdy.6800937

Sun, L., Ma, Y., Wang, H., Huang, W., Wang, X., Han, L., et al. (2018). Overexpression of PtABCC1 contributes to mercury tolerance and accumulation in *Arabidopsis* and poplar. *Biochem. Biophys. Res. Commun.* 497, 997–1002. doi: 10.1016/j.bbrc.2018.02.133

Swartjes, F. A. (ed.) (2011). *Dealing with Contaminated Sites: From Theory Towards Practical Application*. Dordrecht: Springer.

Takahashi, R., Bashir, K., Ishimaru, Y., Nishizawa, N. K., and Nakanishi, H. (2012). The role of heavy-metal ATPases, HMAs, in zinc and cadmium transport in rice. *Plant Signal. Behav.* 7, 1605–1607. doi: 10.4161/psb.22454

Tan, Y.-F., O'Toole, N., Taylor, N. L., and Millar, A. H. (2010). Divalent metal ions in plant Mitochondria and their role in interactions with proteins and oxidative stress-induced damage to respiratory function. *Plant Physiol.* 152, 747–761. doi: 10.1104/pp.109.147942

Tang, H., Krishnakumar, V., Bidwell, S., Rosen, B., Chan, A., Zhou, S., et al. (2014). An improved genome release (version Mt4.0) for the model legume *Medicago truncatula*. *BMC Genomics* 15:312. doi: 10.1186/1471-2164-15-312

Tang, Y., Liu, X., Wang, J., Li, M., Wang, Q., Tian, F., et al. (2016). GAPIT version 2: an enhanced integrated tool for genomic association and prediction. *Plant Genome* 9, 1–9. doi: 10.3835/plantgenome2015.11.0120

Tchounwou, P. B., Yedjou, C. G., Patlolla, A. K., and Sutton, D. J. (2012). Heavy metals toxicity and the environment. *Exp. Suppl.* 101, 133–164. doi: 10.1007/978-3-7643-8340-4_6

Tehseen, M., Cairns, N., Sherson, S., and Cobbett, C. S. (2010). Metallochaperone-like genes in *Arabidopsis thaliana*. *Metallomics* 2:556. doi: 10.1039/c003484c

Thorvaldsdóttir, H., Robinson, J. T., and Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform.* 14, 178–192. doi: 10.1093/bib/bbs017

Trotter, M. V. (2014). "Mutation-selection balance," in *eLS*, ed. John Wiley & Sons Ltd (Chichester: John Wiley & Sons, Ltd). doi: 10.1002/9780470015902.a0001768.pub2

Vallee, B. L., and Ulmer, D. D. (1972). Biochemical effects of mercury, cadmium, and lead. *Annu. Rev. Biochem.* 41, 91–128. doi: 10.1146/annurev.bi.41.070172.000515

Van de Velde, W., Zehirov, G., Szatmari, A., Debreczeny, M., Ishihara, H., Kevei, Z., et al. (2010). Plant peptides govern terminal differentiation of bacteria in symbiosis. *Science* 327, 1122–1126. doi: 10.1126/science.1184057

Vilella, A. J., Blanco-Garcia, A., Hutter, S., and Rozas, J. (2005). VariScan: analysis of evolutionary patterns from large-scale DNA sequence polymorphism data. *Bioinformatics* 21, 2791–2793. doi: 10.1093/bioinformatics/bti403

Wang, C., Wang, T., Mu, P., Li, Z., and Yang, L. (2013). Quantitative trait loci for mercury tolerance in rice seedlings. *Rice Sci.* 20, 238–242. doi: 10.1016/S1672-6308(13)60124-9

Wu, D., Sato, K., and Ma, J. F. (2015). Genome-wide association mapping of cadmium accumulation in different organs of barley. *New Phytol.* 208, 817–829. doi: 10.1111/nph.13512

Xia, J., Yamaji, N., Kasai, T., and Ma, J. F. (2010). Plasma membrane-localized transporter for aluminum in rice. *Proc. Natl. Acad. Sci.U.S.A.* 107, 18381–18385. doi: 10.1073/pnas.1004949107

Yan, A., Wang, Y., Tan, S. N., Mohd Yusof, M. L., Ghosh, S., and Chen, Z. (2020). Phytoremediation: a promising approach for revegetation of heavy metal-polluted land. *Front. Plant Sci.* 11:359. doi: 10.3389/fpls.2020.00359

Young, N. D., Debellé, F., Oldroyd, G. E. D., Geurts, R., Cannon, S. B., Udvardi, M. K., et al. (2011). The *Medicago* genome provides insight into the evolution of rhizobial symbioses. *Nature* 480, 520–524. doi: 10.1038/nature10625

Zhao, J., Yang, W., Zhang, S., Yang, T., Liu, Q., Dong, J., et al. (2018). Genome-wide association study and candidate gene analysis of rice cadmium accumulation in grain in a diverse rice collection. *Rice* 11:61. doi: 10.1186/s12284-018-0254-x

Zhao, Z., Fu, Z., Lin, Y., Chen, H., Liu, K., Xing, X., et al. (2017). Genome-wide association analysis identifies loci governing mercury accumulation in maize. *Sci. Rep.* 7:247. doi: 10.1038/s41598-017-00189-6

Zhou, X., and Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* 44, 821–824. doi: 10.1038/ng.2310

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 03 March 2022
doi: 10.3389/fpls.2022.843033

Maize Seed Germination Under Low-Temperature Stress Impacts Seedling Growth Under Normal Temperature by Modulating Photosynthesis and Antioxidant Metabolism

*Aiju Meng†, Daxing Wen† and Chunqing Zhang**

State Key Laboratory of Crop Biology, Agronomy College, Shandong Agricultural University, Tai'an, China

OPEN ACCESS

Edited by:
Prasanta Kumar Subudhi, Louisiana State University, United States

Reviewed by:
Alma Balestrazzi, University of Pavia, Italy
Tushar Suhas Khare, Savitribai Phule Pune University, India

***Correspondence:**
Chunqing Zhang
cqzhangsdau@163.com;
cqzhang@sdau.edu.cn

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

Received: *24 December 2021*
Accepted: *10 February 2022*
Published: *03 March 2022*

Citation:
Meng A, Wen D and Zhang C (2022) Maize Seed Germination Under Low-Temperature Stress Impacts Seedling Growth Under Normal Temperature by Modulating Photosynthesis and Antioxidant Metabolism. Front. Plant Sci. 13:843033. doi: 10.3389/fpls.2022.843033

Spring maize is usually subjected to low-temperature stress during seed germination, which retards seedling growth later even under a suitable temperature. However, the mechanism underlying maize seed germination under low-temperature stress impacting seedling growth is still ambiguous. In this study, we used one low-temperature sensitive maize (SM) and one low-temperature resistance maize (RM) to investigate the mechanism. The results showed that the SM line had higher malondialdehyde content and lower total antioxidant capacity (TAC) and germination percentage than the RM line under low-temperature stress, indicating the vulnerability of SM line to low-temperature stress. Further transcriptome analysis revealed that seed germination under low-temperature stress caused the down-regulation of photosynthesis-related gene ontology terms in two lines. Moreover, the SM line displayed down-regulation of ribosome and superoxide dismutase (SOD) related genes, whereas genes involved in SOD and vitamin B6 were up-regulated in the RM line. Kyoto Encyclopedia of Genes and Genomes enrichment analysis revealed that photosynthesis and antioxidant metabolism-related pathways played essential roles in response to low-temperature stress during seed germination. The photosynthetic system displayed a higher degree of damage in the SM line. Both qRT-PCR and physiological characteristics experiments showed similar results with transcriptome data. Taken together, we propose a model for maize seed germination in response to low-temperature stress.

Keywords: seed germination, seedling growth, low-temperature stress, maize, transcriptome

INTRODUCTION

Maize (*Zea mays* L.) originated in tropical and subtropical areas and is naturally sensitive to low-temperature stress, especially during seed germination. Low-temperature limits the spread and production of maize all over the world (Zhang et al., 2020). As spring maize, seed germination and seedling growth at an early stage are usually subjected to low-temperature stress. Despite increased temperature after exposure to low-temperature stress, seedling growth

is affected due to the inability of plants to respond quickly to favorable environmental changes (Sowiński et al., 2005). However, the mechanism underlying maize seed germination under low-temperature stress impacting seedling growth is still ambiguous.

Numerous studies have shown that cold stress at the seedling stage affects photosynthesis by reducing the activity of photosystem II (PSII; Savitch et al., 2011). Moreover, cold stress also affects energy collection and preservation at different points, and the reduction of photosynthetic activity varies among different genotypes (Ensminger et al., 2006). The chloroplast ultrastructure of seedlings developed under low-temperature stress is disordered and cannot be repaired after experiencing favorable conditions (Grzybowski et al., 2019). However, the mechanism of low-temperature stress at the germination stage affecting photosynthesis remains unknown.

Low-temperature stress can increase the accumulation of reactive oxygen species (ROS; Li et al., 2019). ROS can cause lipid peroxidation, DNA damage, protein denaturation, carbohydrate oxidation, pigment decomposition, and enzyme activity damage (Bose et al., 2014). Low-temperature tolerance is varied among different genotypes, which may be related to the antioxidant system. ROS scavenging enzymes mainly include superoxide dismutase (SOD), peroxidase (POD), catalase (CAT), glutathione peroxidase (GPX), and ascorbate peroxidase (APX) in plants. Together with antioxidants glutathione and ascorbic acid, these enzymes provide efficient approaches for cells to detoxify O_2^- and H_2O_2 (Romero-Puertas et al., 2006). There is a close relationship between glutathione and cold resistance of maize (Prasad, 1997). Reduced glutathione (GSH), a non-enzymatic antioxidant, is essential for maintaining the redox state of cells (Noctor et al., 2012).

Sugars are essential for plant growth and are involved in response to stress (Zheng et al., 2010). Sucrose is a common substance for energy storage and osmotic regulation in plant cells. Sucrose can also form a sugar layer around the cells, which has a higher membrane phase transition temperature to prevent cell dehydration (Zhang and Bartels, 2018). The increase of sucrose content was the initial reaction of plants exposed to cold conditions (Nägele and Heyer, 2013).

Previous studies have reported the mechanism underlying seedling growth in response to low-temperature stress at the seedling stage (Ma et al., 2015; Zeng et al., 2021). However, spring maize is more susceptible to low-temperature stress at the germination stage than the seedling stage. In this study, two maize inbred lines with different low-temperature resistance were used to investigate the effects of seed germination under low-temperature stress on seedling growth. Maize seed germination at 25°C for 4 days (Normal temperature, NT) was used as control, and maize seed germination at 13°C for 4 days followed by 25°C for 2 days (low-temperature stress followed by normal temperature, LNT) was considered as low-temperature treatment. The physiological experiments showed that the RM line had significantly higher TAC and sucrose content than the SM line under low-temperature stress, partially explaining the different low-temperature resistance phenotypes. Subsequently, we further explore the differences in low-temperature resistance at the transcriptome level. Taken together, the results provide new insights into maize seed germination in response to low-temperature stress.

Abbreviations: SM, Low-temperature sensitive maize; RM, Low-temperature resistance maize; TAC, Total antioxidant capacity; SOD, Superoxide dismutase; PSII, Photosystem II; ROS, Reactive oxygen species; POD, Peroxidase; CAT, Catalase; GPX, Glutathione peroxidase; APX, Ascorbate peroxidase; GSH, Glutathione; NT, 25°C for 4 days; LNT, 13°C for 4 days and then 25°C for 2 days; ABTS, 2,2′-Azino-bis (3-ethylbenzothiazoline-6-sulfonic acid) diammonium salt; FW, Fresh weight; SPAD, Soil and plant analyzer development; DEGs, Differentially expressed genes.

MATERIALS AND METHODS

Materials

The low-temperature sensitive maize (SM) B283-1 and low-temperature resistance maize (RM) 04Qun0522-1-1maize inbred lines used in this study were bred in our laboratory. They were grown at the experimental station (36°90′N, 117°90′E) of Shandong Agricultural University, Shandong, China. Seeds were sown on 12 June 2019. The plant density was 67,500 plants/ha. Seeds used in this study were harvested 50 days after pollination.

Evaluation of Seed Germination

Seed germination was evaluated according to a previous study with some modifications (Wen et al., 2018). The bottom of a sprouting bed consisted of 4 cm height silica sand (diameter of 0.05–0.8 mm) with 60% saturation moisture content in a germination box. Randomly selected 30 maize seeds were sown on the surface of the sprouting bed, and then they were covered with 2 cm height silica sand with 60% saturation moisture content. Subsequently, the germination boxes were placed in different germination conditions for various treatments. The germination boxes were placed in a growth chamber at 13°C for 4 days to detect the percentage of seeds showing radicle protrusion. The germination boxes were placed in a growth chamber at 25°C for 7 days or at 13°C for 7 days to measure germination percentage. A seed was considered as germinating seed when the radicle was similar to seed length and the germ was similar with half of the seed length. Moreover, some germination boxes were placed in a growth chamber at 25°C for 4 days (NT) as control, and some germination boxes were placed in a growth chamber at 13°C for 4 days followed by 25°C for 2 days (LNT) as low-temperature treatment. All tissues of the two inbred lines under NT and LNT were used for later experiments. Each treatment included three replicates.

Measurement of Total Antioxidant Capacity

TAC was measured by a 2,2′-azino-bis (3-ethylbenzothiazoline-6-sulfonic acid) diammonium salt (ABTS) assay kit (Comin, Suzhou, China) following the manufacturer's protocols. ABTS, a colorless substance, can be oxidized to a stable blue–green

cationic radical form $ABTS^+$ (Scalzo et al., 2005). The absorption peak of $ABTS^+$ is at 734 nm. When the tested substance is added to the $ABTS^+$ solution, the antioxidant components can react with $ABTS^+$ and fade the reaction system. The change of absorbance at 734 nm was detected. The antioxidant capacity quantified with Trolox was used as a control. The results were expressed in μmol Ttoloox/g fresh weight (FW). Three biological replicates were used for each treatment.

Measurement of Lipid Peroxidation

The content of thiobarbituric acid (TBA) reactive substances (TBARS) can be used to assess lipid peroxidation (Checchio et al., 2021). In this study, lipid peroxidation was detected using a commercial kit (Comin, Suzhou, China). The results were expressed in nmol/g FW. Three biological replicates were used for each treatment.

Measurement of Sucrose Content

The determination principle of sucrose content is that alkali is used to heat the sample together to destroy the reducing sugar (Handel, 1968). Then, sucrose was hydrolyzed to glucose and fructose under acidic conditions, and fructose reacted with resorcinol to form a colored substance with a characteristic absorption peak at 480 nm. In this study, sucrose content was detected using a sucrose content determination kit (Comin, Suzhou, China) according to the manufacturer's protocols. The results were expressed in mg/g FW. Three biological replicates were used for each treatment.

Measurement of SOD Activity

SOD activity was determined by the photoinhibition of nitro blue tetrazole (NBT) according to a previous study (Giannopolitis and Ries, 1977). In this study, SOD activity was measured using a SOD determination kit (Comin, Suzhou, China) according to the manufacturer's protocols. The results were expressed in U/g FW. Three biological replicates were used for each treatment.

Measurement of POD Activity

POD activity was determined by the method of guaiacol oxidation (Zhang et al., 2015a). In this study, POD activity was detected using a POD determination kit (Comin, Suzhou, China) according to the manufacturer's protocols. The results were expressed in U/kg FW. Three biological replicates were used for each treatment.

Measurement of Chlorophyll Content

Chlorophyll contents of the two inbred lines germinated at 25°C for 6 days or germinated at 13°C for 4 days followed by 25°C for 4 days were measured by Soil and Plant Analyzer Development (SPAD) 502 chlorophyll meter (Konica Minolta Inc., Japan). Three biological replicates were used for each treatment.

RNA-seq and Transcriptome Analysis

All tissues of the two inbred lines under NT and LNT conditions were used for RNA extraction. After quick freezing with liquid nitrogen, samples from each replication were pooled and stored at −80°C. Three biological replicates were used for each treatment. Total RNA was extracted using a biospin plant total RNA extraction kit (Bioflux, Beijing, China) according to the manufacturer's protocols. Libraries were generated using NEBNext® UltraTM RNA Library Prep kit for Illumina® (NEB, United States) following the manufacturer's protocols. Illumina sequencing was performed as described previously (Yu et al., 2021). Hisat2 (v2.0.5) was used to build an index of the reference genome (B73 v4),[1] and then paired-end clean reads were aligned to the reference genome (Yu et al., 2021). Differential expression analyses of two groups (three biological replicates per treatment) were performed using the DESeq2 R package (v1.16.1). The resulting values of p were adjusted (padj) using Benjamini and Hochberg's approach for controlling the false discovery rate. Differentially expressed genes (DEGs) with padj <0.05 and $|\log_2\text{Fold change}| \geq 1$were used for further analyses.

To understand the main biological functions of DEGs in maize, we carried out an enrichment analysis of DEGs. ClusterProfiler R package (v3.4.4) was used to achieve Gene Ontology (GO) enrichment analysis of DEGs (Yu et al., 2012). GO terms with corrected $p < 0.05$ were considered as significantly enriched GO terms. To understand the primary metabolic pathway of DEGs in maize, we carried out the Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of DEGs. ClusterProfiler R package (v3.4.4) was used to test the statistical enrichment of DEGs in the KEGG pathway (Yu et al., 2012). Similarly, the KEGG pathways with corrected $p < 0.05$ were assigned as significantly enriched pathways.

qRT-PCR

To validate the DEGs identified from transcriptome analysis, six genes were randomly selected for a qRT-PCR assay: *Zm00001d002611* (SOD 13), *Zm00001d025103* (Amine oxidase1), *Zm00001d027557* (Glutathione transferase 31), *Zm00001d027422* (PsbP-like protein 1), *Zm00001d021906* (Light harvesting complex A2), and *Zm00001d046170* (phosphoenolpyruvate carboxylase1). Primers were designed using the primer-premier software (v6.0; **Supplementary Table S6**). The maize *Actin* gene (*Zm00001d010159*) was used as an internal control (Zhang et al., 2015b). An ABI StepOne Plus Real-time PCR System (Applied Biosystems, CA, United States) was used to perform qRT-PCR according to the instructions of SYBR® Green Real-time PCR Master Mix (Takara, Dalian, China). The $2^{-\Delta\Delta CT}$ method was used to calculate the relative expression of genes (Livak and Schmittgen, 2001).

Statistical Analysis

Multiple comparisons were performed using Duncan's test at the 0.05 significance level. All the tests were conducted using SPSS Version 21.0 for Windows (SPSS, Chicago, IL, United States).

[1]http://www.maizesequence.org/index.html

RESULTS

The SM Line Is More Vulnerable to Low-Temperature Stress

To investigate the effects of maize seed germination under low-temperature stress on subsequent seedling growth under normal temperature, we first measured some traits of two maize inbred lines with different low-temperature resistance. When maize seeds germinated at 13°C for 4 days, the percentage of seeds showing radicle protrusion in the RM line was higher compared to the SM line (**Figure 1A**). In seed testing, maize seeds germinating at 25°C for 7 days are usually used for evaluating seed germination percentage. The germination percentage of the SM line decreased about 60% at 13°C for 7 days, while there was no significant difference in germination percentage between seeds germinated at 13°C for 7 days and seeds germinated at 25°C for 7 days in the RM line (**Figures 1B,C**). Maize seeds germinating at 25°C for 4 days (NT) are usually used for evaluating seed germination energy in seed testing. In this study, the two inbred lines had similar germination energy and seedling length (**Figure 1D**). To investigate the effects of maize seed germination under low-temperature stress impacting seedling growth under normal temperature, we set a similar accumulated temperature between NT and low-temperature stress followed by normal temperature (LNT) treatment. Given the same accumulated temperature, maize seeds first germinated at 13°C for 4 days followed by 25°C for 2 days under LNT treatment. At this time, the RM line had notably higher seedlings than the SM line (**Figure 1E**). Therefore, these results suggested that the SM line was more

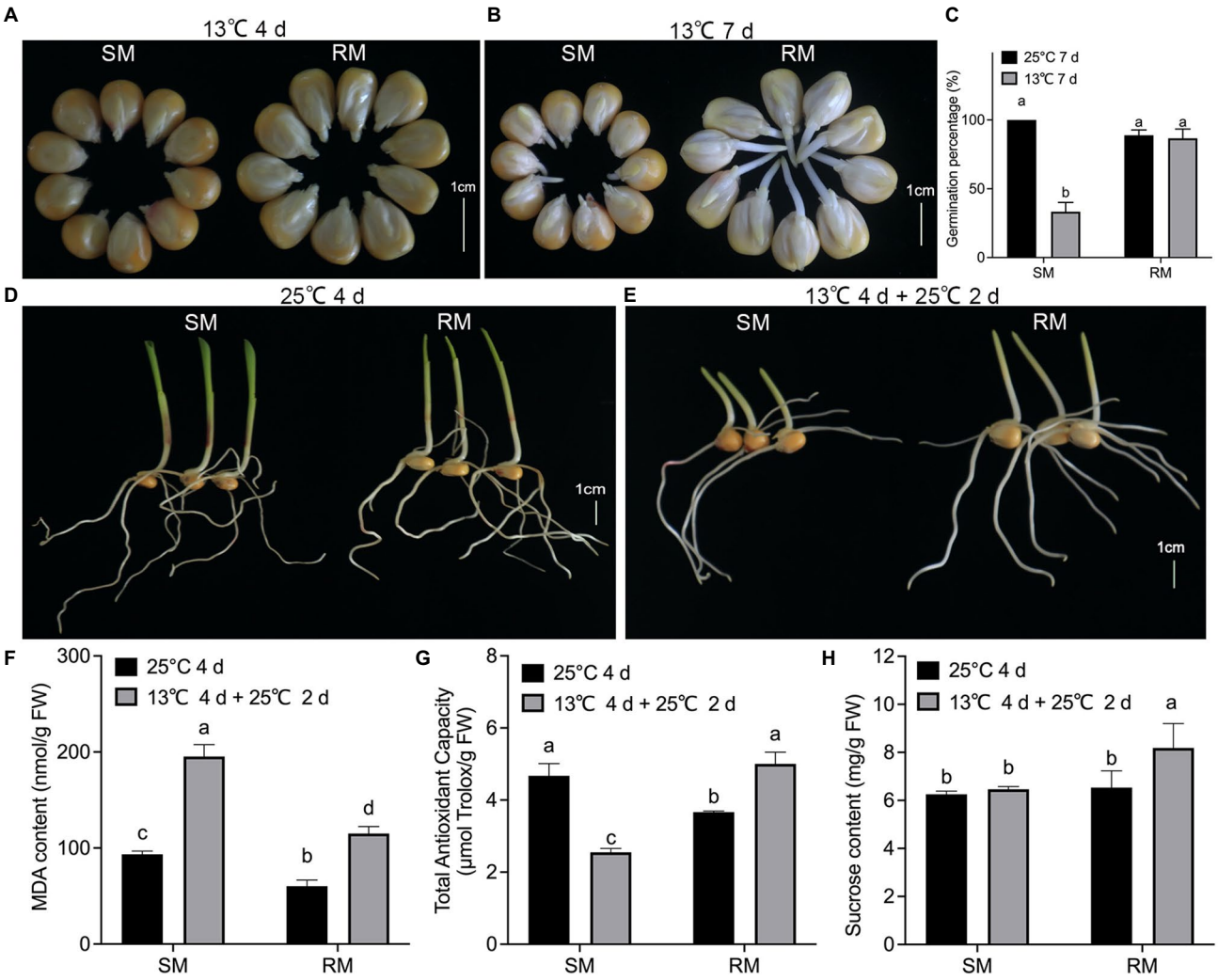


FIGURE 1 | Effects of maize seed germination under low-temperature stress on subsequent seedling growth under normal temperature. **(A)** Phenotypes of seed germination at 13°C for 4 days. **(B)** Phenotypes of seed germination at 13°C for 7 days. **(C)** Germination percentage. **(D)** Phenotypes of seed germination at 25°C for 4 days. **(E)** Phenotypes of seed germination at 13°C for 4 days followed by 25°C for 2 days. Scale bar, 1 cm. **(F)** Malondialdehyde (MDA) content. **(G)** Total antioxidant capacity. **(H)** Sucrose content. Data are means ± SD ($n = 3$ replications of thirty plants). Different letters indicate significant differences among means under different treatments ($p < 0.05$). FW: fresh weight.

vulnerable to low-temperature stress. Subsequently, we detected the changes in lipid peroxidation, expressed as malondialdehyde (MDA) content, TAC, and sucrose content. Although MDA contents of the two lines under LNT were about twice as high as those under NT, the RM line under LNT had a lower MDA content (115 nmol/g FW) than the SM line under LNT (195 nmol/g FW; **Figure 1F**). TAC decreased in the SM line but increased in the RM line under LNT treatment compared with NT treatment (**Figure 1G**). TAC under LNT (2.6 μmol Trolox/g FW) decreased to about half of that under NT (5.0 μmol Trolox/g FW) in the SM line. Moreover, the RM line had notably higher sucrose content under LNT than under NT (**Figure 1H**). Taken together, seeds germinated under low-temperature stress increased TAC and sucrose content in the RM line, which might be related to low-temperature resistance.

Transcriptome Analysis of Maize Seed Germination in Response to Low-Temperature Stress

To explore genes and metabolic pathways that control maize seed germination in response to low-temperature stress, we selected samples of RM and SM under NT and LNT for transcriptome analysis. After removing adapters and sequences with low-quality regions, there remained approximately 40–50 million clean reads (**Supplementary Table S1**). Then, about 35–46 million clean reads were mapped to the maize genome. These clean reads included 83–88% uniquely mapped reads and 2.5–3.0% multiple mapped reads. DESeq2 R package was used to identify DEGs using padj <0.05 and $|\log_2\text{Fold change}| \geq 1$ as the cutoff. The results displayed that 3,186 genes were significantly up-regulated and 4,281 genes were significantly downregulated in the SM line under LNT compared with those under NT (SM_LNTvsNT). Moreover, 2,797 genes were significantly up-regulated and 3,918 genes were significantly downregulated in the RM line under LNT compared with those under NT (RM_LNTvsNT; **Figure 2A**). Venn diagram showed that common up-regulated genes (54) were fewer than common down-regulated genes (608), most DEGs were specific in SM_LNTvsNT and RM_LNTvsNT (**Figures 2B,C**).

Maize Seed Germination Under Low-Temperature Stress Cause Down-Regulation of Photosynthesis-Related Go Terms

GO enrichment analysis displayed that there was no significantly enriched GO term for common up-regulated genes, which might be due to the low number of DEGs. For these SM-specific up-regulated DEGs, there were only two significantly enriched GO terms, i.e., myosin complex (GO: 0016459, $p = 4.82 \times 10^{-4}$) in the cellular component group and ADP binding (GO: 0043531, $p = 2.32 \times 10^{-5}$) in the molecular function group (**Figure 3A**). For these RM-specific up-regulated DEGs, there were nine significantly enriched GO terms. Among them, the most significantly enriched GO terms were superoxide metabolic (GO: 0006801, $p = 8.55 \times 10^{-6}$) in the biological process group, vitamin binding (GO: 0019842, $p = 5.35 \times 10^{-7}$) in the molecular function group, which might play essential roles in RM resistant to low-temperature stress (**Figure 3B**).

Compared with the up-regulated GO terms, there were more down-regulated GO terms. For these common down-regulated DEGs, the most significantly enriched GO terms were photosynthesis (GO: 0015979, $p = 2.16 \times 10^{-44}$) in the biological process group, thylakoid (GO: 0009579, $p = 5.33 \times 10^{-54}$) in the cellular component group, iron–sulfur cluster binding (GO: 0051536, $p = 4.38 \times 10^{-8}$) in the molecular function group (**Figure 3C**). Moreover, many photosynthesis-related GO terms were also enriched. The results indicated that maize seed germination under low-temperature stress caused the down-regulation of photosynthesis-related GO terms in the two inbred lines. Translation (GO: 0006412, $p = 6.79 \times 10^{-16}$) in the biological process group, ribosome (GO: 0005840, $p = 7.64 \times 10^{-18}$) in the cellular component group, structural constituent of ribosome (GO: 0003735, $p = 6.47 \times 10^{-18}$) in the molecular function group, represented the most markedly enriched GO terms in SM-specific

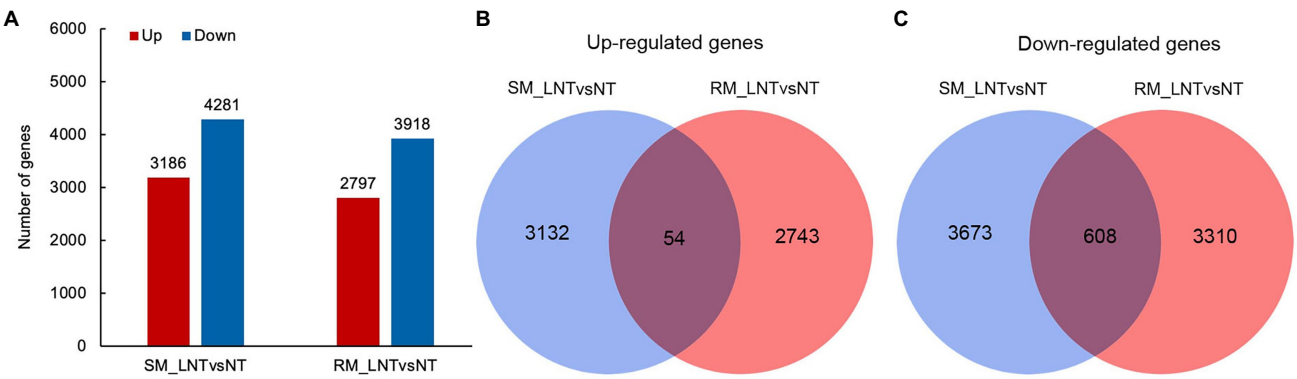


FIGURE 2 | Differentially expressed genes in two maize inbred lines. **(A)** The number of up-regulated and down-regulated differentially expressed genes (DEGs) between low-temperature sensitive maize (SM) and low-temperature resistance maize (RM) lines. **(B)** Venn diagram of up-regulated DEGs in SM_LNTvsNT and RM_LNTvsNT. **(C)** Venn diagram of down-regulated DEGs in SM_LNTvsNT and RM_LNTvsNT. SM: low-temperature sensitive maize inbred line; RM: low-temperature resistant maize inbred line. NT treatment: seeds germinated at 25°C for 4 days. LNT treatment: maize seeds germinated at 13°C for 4 days followed by 25°C for 2 days. SM_LNTvsNT: SM line samples under LNT compared with those under NT. RM_LNTvsNT: RM line samples under LNT compared with those under NT.

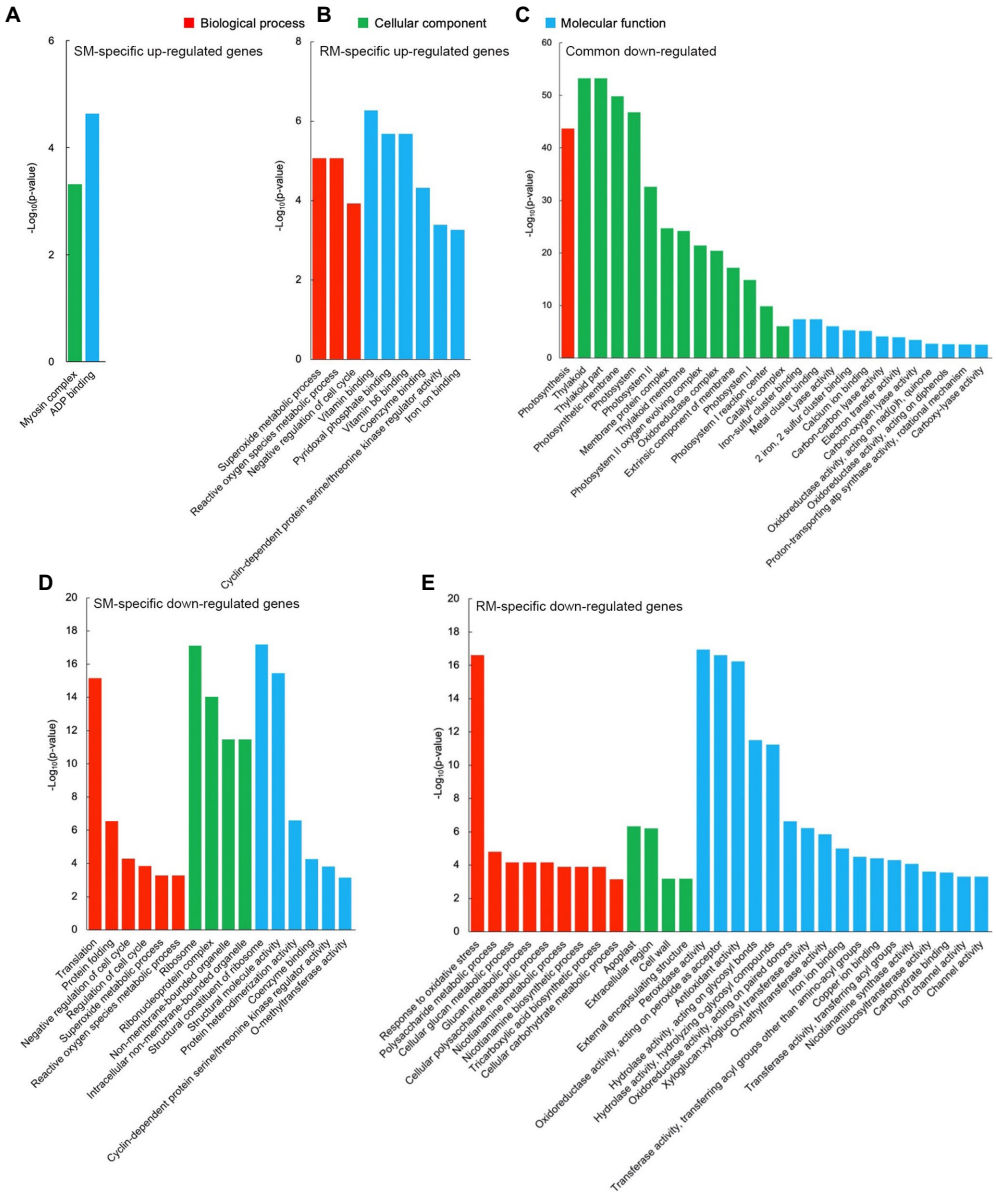


FIGURE 3 | Significantly enriched gene ontology (GO) terms in **(A)** SM-specific up-regulated genes. **(B)** RM-specific up-regulated genes. **(C)** Common down-regulated genes. **(D)** SM-specific down-regulated genes. **(E)** RM-specific down-regulated genes in SM_LNTvsNT and RM_LNTvsNT. SM: low-temperature sensitive maize inbred line; RM: low-temperature resistant maize inbred line. NT treatment: seeds germinated at 25°C for 4 days. LNT treatment: maize seeds germinated at 13°C for 4 days followed by 25°C for 2 days. SM_LNTvsNT: SM line samples under LNT compared with those under NT. RM_LNTvsNT: RM line samples under LNT compared with those under NT.

down-regulated DEGs, suggesting that ribosomes of SM line might be damaged by low-temperature stress (**Figure 3D**). For RM-specific down-regulated DEGs, the most prominently enriched GO terms were response to oxidative stress (GO: 0006979, $p = 2.41 \times 10^{-17}$) in the biological process group, apoplast (GO: 0048046, $p = 4.58 \times 10^{-7}$) in the cellular component group, peroxidase activity (GO: 0004601, $p = 1.14 \times 10^{-17}$) in the molecular function group (**Figure 3E**). The results indicated that peroxidase activity might not be involved in the enhanced antioxidant capacity of the RM line in response to low-temperature stress.

Photosynthesis and Antioxidant Metabolism Pathways Are Involved in Response to Low-Temperature Stress at the Germination Stage

KEGG enrichment analysis showed photosynthesis-antenna proteins, photosynthesis, phenylpropanoid biosynthesis, flavonoid biosynthesis, glutathione metabolism, porphyrin and chlorophyll metabolism, stilbenoid, diarylheptanoid, and gingerol biosynthesis were markedly enriched KEGG pathways for common DEGs (**Figure 4**). Therefore, photosynthesis and antioxidant metabolism

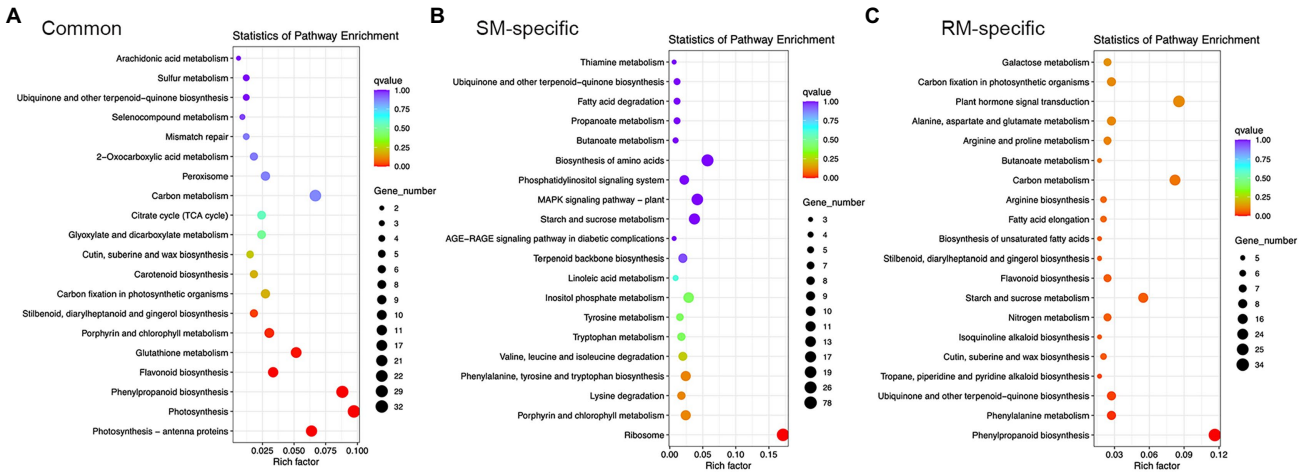


FIGURE 4 | Top 20 KEGG pathways in **(A)** Common DEGs. **(B)** SM-specific DEGs. **(C)** RM-specific DEGs in SM_LNTvsNT, and RM_LNTvsNT. SM: low-temperature sensitive maize inbred line; RM: low-temperature resistant maize inbred line. NT treatment: seeds germinated at 25°C for 4 days. LNT treatment: maize seeds germinated at 13°C for 4 days followed by 25°C for 2 days. SM_LNTvsNT: SM line samples under LNT compared with those under NT. RM_LNTvsNT: RM line samples under LNT compared with those under NT.

pathways might play essential roles in response to low-temperature stress at the germination stage.

Only the ribosome pathway was significantly enriched for SM-specific DEGs, which was also enriched in GO enrichment analysis. In the ribosome pathway, there were 78 DEGs in SM_LNTvsNT (**Supplementary Table S2**). Of the 78 DEGs, most genes were down-regulated except for three genes (*Zm00001d022111*, *Zm00001d022197*, and *Zm00001d002462*), indicating that low-temperature stress might have a significant influence on the ribosomal pathway of low-temperature sensitive inbred lines. For RM-specific DEGs, the significantly enriched KEGG pathways were related to phenylpropanoid biosynthesis, phenylalanine metabolism, ubiquinone, and other terpenoid-quinone biosynthesis, tropane, piperidine and pyridine alkaloid biosynthesis (**Supplementary Table S3**). Only one gene was down-regulated among five DEGs involved in tropane, piperidine and pyridine alkaloid biosynthesis.

The Photosynthetic System of the SM Line Is More Vulnerable to Low-Temperature Stress

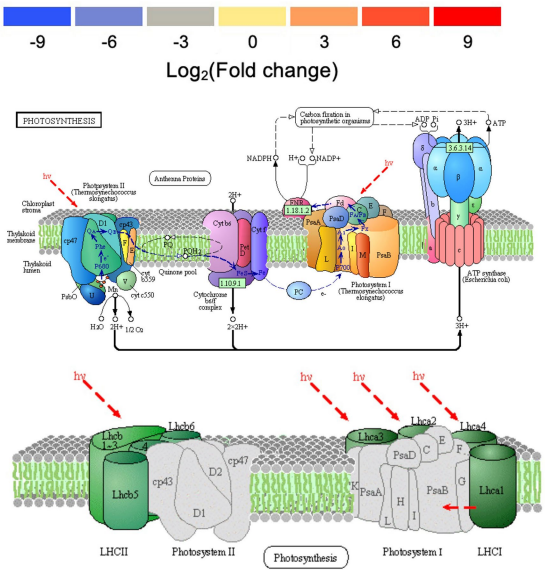
The DEGs enriched in photosynthesis-related pathways were mainly located in the chloroplast and annotated to function as antenna proteins, photosystems I and II components, porphyrin, and chlorophyll metabolism-related proteins (**Figure 5**; **Supplementary Table S4**). Although DEGs involved in photosynthetic-antenna proteins were all down-regulated in both SM and RM lines, the degree of down-regulation of the DEGs was lower in the RM line than in the SM line. Moreover, most of the DEGs involved in photosynthesis and porphyrin and chlorophyll metabolism pathway were also down-regulated in both SM and RM lines, and the RM line also had a lower degree of down-regulation of the DEGs than the SM line (**Figure 5**; **Supplementary Table S4**). Taken together, the photosynthetic system of the SM line was even more damaged when seed germinated under low-temperature stress.

Validation of Transcriptome Data by qRT-PCR and Physiological Characteristics

All the DEGs showed similar expression patterns in the qRT-PCR assays as their changes of relative expression level identified by RNA-seq, suggesting the transcriptome data were credible (**Figure 6**). By extending the seedling growth at 25°C for 2 days, we observed that maize seed germination under low-temperature markedly inhibited subsequent seedling growth under normal temperature, especially in the SM line (**Figures 7A,B**). Subsequently, we further detected the changes in SPAD value, SOD, and POD activities. Both SM and RM lines displayed a significant decrease of SPAD value under LNT, and the level of reduction in the SM line was larger than that in the RM line, which was consistent with the changes of the photosynthetic system from transcriptome analysis (**Figure 7C**). The SM line had lower SOD activity and higher POD activity under LNT than NT treatment, while the RM line showed the opposite trend of SOD and POD activities, which were consistent with the results of GO enrichment analysis (**Figures 7D,E**). Moreover, the changes in SOD activities in both SM and RM lines were similar to TAC trends (**Figures 1F**, **7D**). Therefore, SOD activity might play a key role in TAC in seed germination under low-temperature stress.

DISCUSSION

Low-temperature stress often occurs at the germination stage, which retards the seedling emergence of spring maize.



Photosynthesis

Gene ID	SM_LNTvsNT	RM_LNTvsNT
delta-1(*Zm00001d018069*)		
delta-2(*Zm00001d052242*)		
gamma(*Zm00001d021620*)		
PetC-1(*Zm00001d053432*)		
PetC-2(*Zm00001d016134*)		
PetF-1(*Zm00001d035001*)		
PetF-2(*Zm00001d035003*)		
PetF-3(*Zm00001d012293*)		
PetF-4(*Zm00001d035002*)		
PetF-5(*Zm00001d049732*)		
PetH-1(*Zm00001d042049*)		
PetH-2(*Zm00001d011833*)		
PetH-3(*Zm00001d025874*)		
PsaD(*Zm00001d034543*)		
PsaE(*Zm00001d005446*)		
PsaF-1(*Zm00001d013146*)		
PsaF-2(*Zm00001d034283*)		
PsaG-1(*Zm00001d020877*)		
PsaG-2(*Zm00001d005996*)		
PsaH(*Zm00001d038984*)		
PsaK(*Zm00001d018797*)		
PsaL(*Zm00001d039687*)		
PsaN(*Zm00001d023713*)		
PsaO(*Zm00001d003767*)		
Psb27(*Zm00001d029049*)		
Psb28(*Zm00001d042178*)		
PsbP-1(*Zm00001d018779*)		
PsbP-2(*Zm00001d007857*)		
PsbQ-1(*Zm00001d021703*)		
PsbQ-2(*Zm00001d050889*)		
PsbS-1(*Zm00001d042697*)		
PsbS-2(*Zm00001d035135*)		

Photosynthesis-antenna proteins

Gene ID	SM_LNTvsNT	RM_LNTvsNT
Lhca1(*Zm00001d015385*)		
Lhca2-1(*Zm00001d006663*)		
Lhca2-2(*Zm00001d021906*)		
Lhca2-3(*Zm00001d005814*)		
Lhca3(*Zm00001d046786*)		
Lhca4(*Zm00001d032197*)		
Lhca5(*Zm00001d018157*)		
Lhcb1-1(*Zm00001d039040*)		
Lhcb1-2(*Zm00001d021435*)		
Lhcb1-3(*Zm00001d044396*)		
Lhcb1-4(*Zm00001d044402*)		
Lhcb1-5(*Zm00001d009589*)		
Lhcb1-5(*Zm00001d050403*)		
Lhcb1-6(*Zm00001d011285*)		
Lhcb2(*Zm00001d033136*)		
Lhcb4-1(*Zm00001d006587*)		
Lhcb4-2(*Zm00001d021763*)		
Lhcb5-1(*Zm00001d048998*)		
Lhcb5-2(*Zm00001d007267*)		
Lhcb6-1(*Zm00001d026599*)		
Lhcb6-2(*Zm00001d001857*)		

Porphyrin and chlorophyll metabolism

Gene ID	SM_LNTvsNT	RM_LNTvsNT
Zm00001d026405		
Zm00001d002951		
Zm00001d008230		
Zm00001d011819		
Zm00001d018034		
Zm00001d029150		
Zm00001d023536		
Zm00001d003935		
Zm00001d029074		
Zm00001d013013		

FIGURE 5 | Heat map of photosynthesis-related genes enriched in Kyoto encyclopedia of genes and genomes (KEGG) pathways in SM_LNTvsNT and RM_LNTvsNT. SM: low-temperature sensitive maize inbred line; RM: low-temperature resistant maize inbred line. NT treatment: seeds germinated at 25°C for 4 days. LNT treatment: maize seeds germinated at 13°C for 4 days followed by 25°C for 2 days. SM_LNTvsNT: SM line samples under LNT compared with those under NT. RM_LNTvsNT: RM line samples under LNT compared with those under NT. Detailed lists of the DEGs are shown in **Supplementary Table S3**. The color code from blue to red suggests the expression level of the DEGs normalized as the $\log_2$ (Fold change).

Maize is vulnerable to low-temperature stress at the germination stage and the early stage of seedling establishment (Zhang et al., 2020). Deciphering the mechanism underlying maize seed germination in response to low-temperature stress can help improve low-temperature resistance.

ROS generation in plant cells can be induced by some environmental factors, such as cold (Li et al., 2019), drought (Zheng et al., 2020), heat (Zhao et al., 2018), and cadmium stress (Gu et al., 2019). ROS, as signal molecules, trigger signal transduction pathways in response to these abiotic stresses. In addition, ROS can cause irreversible cell damage through its strong oxidation characteristics, thereby promoting the change of plant morphology and structure and enhancing resistance (Bose et al., 2014; Ohama et al., 2017). ROS can cause lipid peroxidation, DNA damage, protein denaturation, carbohydrate oxidation, pigment decomposition, and enzyme activity damage (Bose et al., 2014). In the present study, both SM and RM lines showed enhanced MDA content in seed germination under low-temperature stress, which was consistent with previous studies.

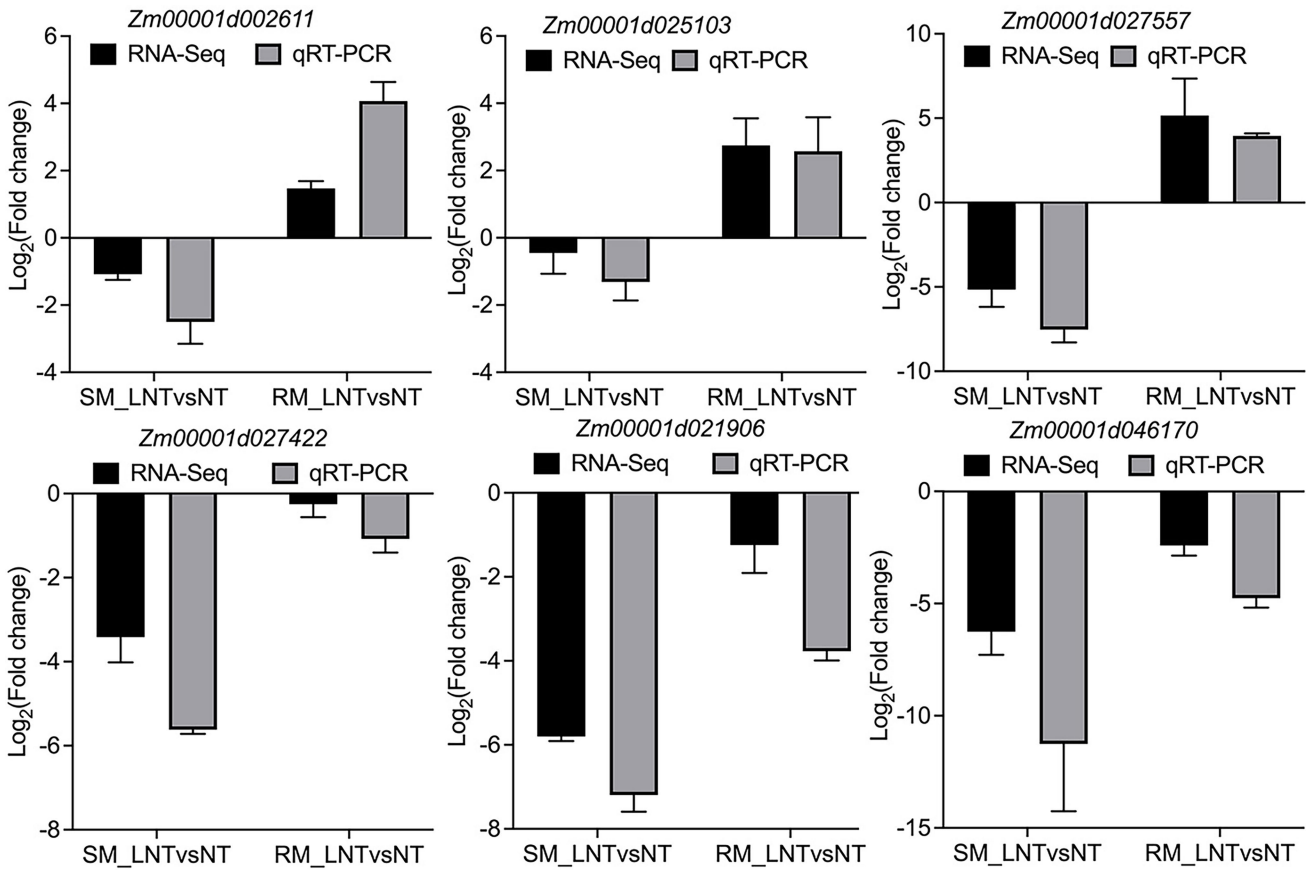


FIGURE 6 | Validation of DEGs by qRT-PCR. *Zm00001d002611* (Superoxide dismutase 13); *Zm00001d025103* (Amine oxidase1); *Zm00001d027557* (Glutathione transferase 31); *Zm00001d027422* (PsbP-like protein 1); *Zm00001d021906* (Light harvesting complex A2); *Zm00001d046170* (phosphoenolpyruvate carboxylase1). The black and gray bars represent the relative expression level from the RNA-seq and qRT-PCR data. The maize *Actin* gene, as an internal control, was used to normalize the expression levels of the target genes. The error bars represent the standard deviations of three replicates. SM: low-temperature sensitive maize inbred line; RM: low-temperature resistant maize inbred line. NT treatment: seeds germinated at 25°C for 4 days. LNT treatment: maize seeds germinated at 13°C for 4 days followed by 25°C for 2 days. SM_LNTvsNT: SM line samples under LNT compared with those under NT. RM_LNTvsNT: RM line samples under LNT compared with those under NT.

The detoxification mechanism of ROS plays a vital role in the normal metabolism of plants, especially under stress. The main ROS scavenging enzymes in plants include SOD, POD, CAT, GPX, and APX. Previous studies have shown that CAT and monodehydroascorbate reductase activities are effective screening tools for maize hybrids with cold resistance (Hodges et al., 1997). The activities of antioxidant enzymes significantly increase when maize seeds germinate under low-temperature (Cao et al., 2019). In the present study, the RM line under LNT displayed increased SOD activity and decreased POD activity compared with NT treatment, while the SM line showed opposite trends in both SOD and POD activities (**Figures 7D,E**). Moreover, GO enrichment analysis showed similar trends with the activities of SOD and POD (**Figures 3B,E**). Interestingly, the changes in TAC were consistent with SOD activity (**Figures 1F**, **7D**). Therefore, SOD activity might play a key role in the TAC of maize seed germination under low-temperature stress. A previous study has shown that the intrinsic high level of SOD in halophytes is necessary to trigger a series of adaptive responses, and the role of other enzymatic antioxidants might reduce the basic level of H_2O_2 (Bose et al., 2014). Whether SOD activity also triggers a series of adaptive responses under low-temperature stress needs further study.

Glutathione metabolism regulates redox-sensitive signal transduction of plant tissues and maintains antioxidant properties (Cnubben et al., 2001; Noctor et al., 2012). Compared with NT treatment, glutathione metabolism-related genes encoding GPX (*Zm00001d026154* and *Zm00001d002704*), glutathione transferase (*Zm00001d018220*, *Zm00001d027557, Zm00001d042102*, *Zm00001d029706* and *Zm00001d043344*), isocitrate dehydrogenase (*Zm00001d044021*) were all up-regulated only in the RM line under LNT (**Supplementary Table S5**). Of these genes, *Zm00001d026154* and *Zm00001d029706* have been reported to help maize resist drought stress (Zheng et al., 2020). Therefore, *Zm00001d026154* and *Zm00001d029706* might be essential for resistance to various abiotic stresses in maize.

Vitamin B6 contains six forms, pyridoxal, pyridoxamine, pyridoxine, pyridoxal 5′-phosphate (PLP), pyridoxamine

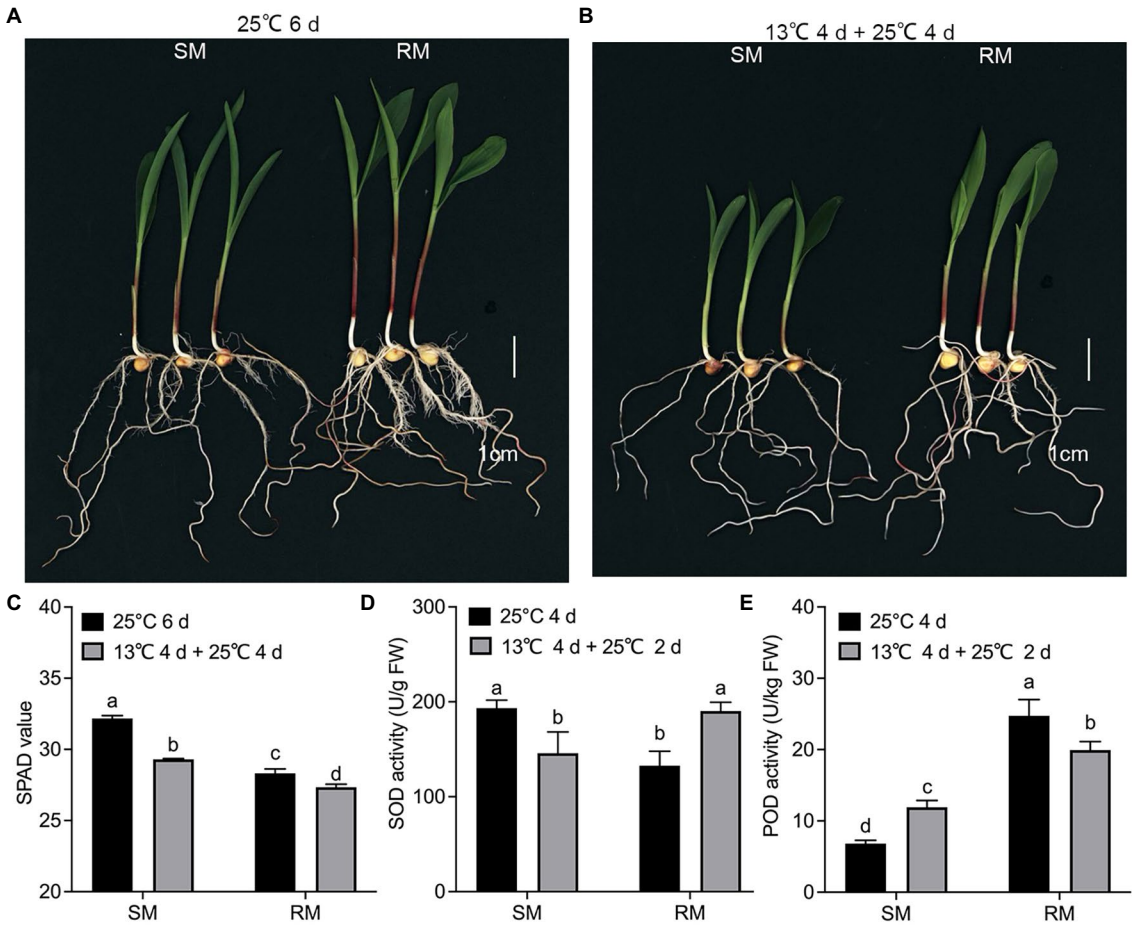


FIGURE 7 | Validation of transcriptome data by physiological characteristics. **(A)** Phenotypes of seed germination at 25°C for 6 days. **(B)** Phenotypes of seed germination at 13°C for 4 days followed by 25°C for 4 days. **(C)** SPAD value. **(D)** Superoxide dismutase. **(E)** Peroxidase. Data are means ± SD ($n = 3$ replications of 30 plants). Different letters indicate significant differences among means under different treatments ($p < 0.05$). FW: fresh weight. SM: low-temperature sensitive maize inbred line; RM: low-temperature resistant maize inbred line.

5′-phosphate, and pyridoxine 5′-phosphate, of which PLP is the active form (Denslow et al., 2007). PLP is essential for many biochemical reactions, including decarboxylation, transamination, deamination, racemization, and trans sulfur reactions, which are mainly related to amino acid synthesis (Drewke and Leistner, 2001). Vitamin B6 as a cofactor has been fully confirmed. Moreover, vitamin B6 as an effective antioxidant and a factor that can increase resistance to biotic and abiotic stress has been proved (Huang et al., 2013). Previous studies have shown that VB6 is an effective singlet oxygen quencher, and its quenching rate is equivalent to or higher than that of vitamin C and E, which are known as the two most effective biological antioxidants (Ehrenshaft et al., 1998, 1999; Denslow et al., 2007). In the present study, GO enrichment analysis of RM-specific up-regulated DEGs showed pyridoxal phosphate binding (GO: 0030170, $p = 2.08 \times 10^{-6}$) and vitamin B6 binding (GO: 0070279, $p = 3.85 \times 10^{-8}$) were significantly enriched in the molecular function group (**Figure 3B**). Therefore, vitamin B6 might be involved in enhancing the low-temperature resistance of maize at the germination stage.

Photoinhibition occurs when the harvested light energy exceeds the available energy of chloroplasts or low-temperature sensitive plants exposed to low-temperature stress (Li et al., 2019). Moreover, low-temperature stress can regulate PSII activity, leading to the loss of photosynthetic capacity (Savitch et al., 2011). The down-regulation of light-harvesting complex protein will affect the downstream energy-related processes and ultimately affect plant growth and development (Savitch et al., 2011). Most genes related to photosynthetic apparatus were down-regulated in both RM and SM lines, but the level of reduction of the SM line was greater than that of the RM line (**Figure 5**). Compared with NT treatment, the RM line showed better recovery ability in photosynthesis than the SM line under LNT. Cold affects photosynthesis through overexcitation of PSII reaction centers and the production of oxygen free radicals (Nie et al., 1992). ROS has harmful effects on photosynthetic devices (Di Fenza et al., 2017). The RM line has higher TAC than the SM line, which might be related to the smaller decrease in photosynthesis in the RM line.

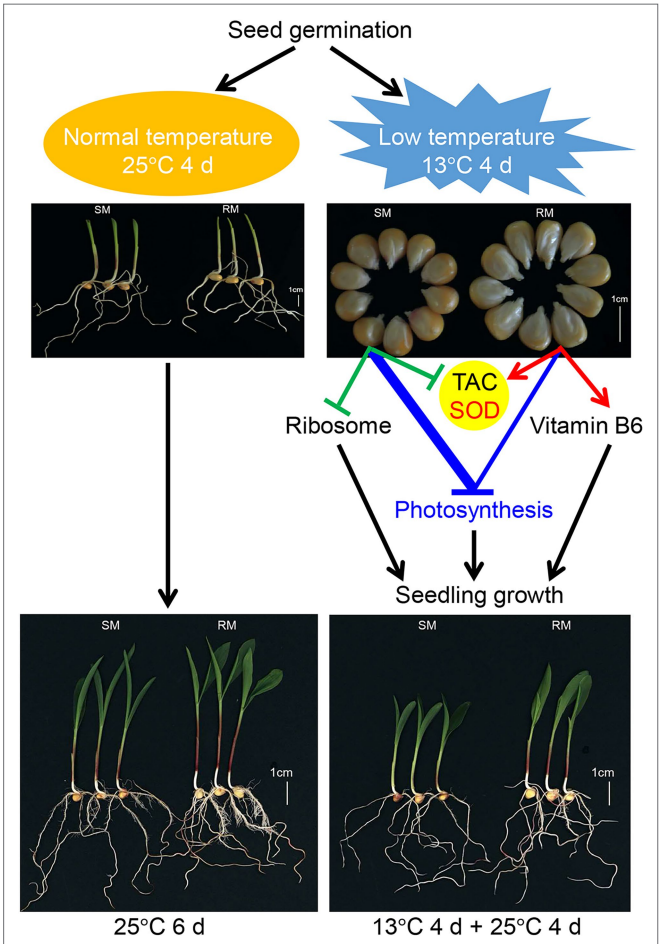


FIGURE 8 | A possible network of maize seed germination under low temperature affecting subsequent seedling growth. The thickened blue line indicates stronger changes induced by low-temperature stress (according to **Figure 5**).

After low-temperature treatment, the SPAD values of the two lines decreased significantly, but the level of reduction of the SM line was greater than that of the RM line, which was consistent with the transcriptome data. Therefore, seed germination under low-temperature stress reduced subsequent seedling photosynthesis, but the level of reduction of photosynthesis was different between maize inbred lines with various low-temperature resistance.

Ribosomes are implicated in resistance to various adverse conditions (Garcia-Molina et al., 2020). Many down-regulated DEGs were enriched in the ribosome (GO: 0005840, $p = 7.64 \times 10^{-18}$) in the SM line (**Supplementary Table S2**). In mammalian cells, nucleolus, especially ribosome, is considered to be the hub of integrating cell response to adverse conditions (Yang et al., 2018; Pfister, 2019). Some plant species have similar regulatory roles of ribosomes under abiotic and biotic stresses (Garcia-Molina et al., 2020). So far, many ribosomal proteins (RPS) mutants with defects in chloroplast ribosomes have been reported in plants (Schultes et al., 2000; Xu et al., 2013; Zhang et al., 2016). In the present study, gene (*Zm00001d012353*) encoding 30S ribosomal protein S17 chloroplastic was significantly down-regulated in SM_LNTvsNT. The first plant plastid ribosomal protein mutant (high chlorophyll fluorescence 60) in maize displays an unstable light green effect on seedling growth due to the lack of plastid ribosomal small subunit protein 17 (Schultes et al., 2000). The transcription level of RPS is up-regulated after low-temperature acclimation (Garcia-Molina et al., 2020). The up-regulation of RPS under low-temperature stress is considered to maintain the rate of protein synthesis under adverse thermodynamic conditions (Garcia-Molina et al., 2020). The genes encoding 30S ribosomal protein S1 chloroplastic (*Zm00001d038835*), 30S ribosomal protein S10 chloroplastic (*Zm00001d028153*), 30S ribosomal protein S4 chloroplastic (*Zm00001d047186*), 30S ribosomal protein S6 alpha chloroplastic (*Zm00001d034808*), 50S ribosomal protein L1 chloroplastic (*Zm00001d038084*), 50S ribosomal protein L11 chloroplastic (*Zm00001d027421*), 50S ribosomal protein L17 chloroplastic (*Zm00001d012998*), 50S ribosomal protein L21 chloroplastic (*Zm00001d053377*), 50S ribosomal protein L6 chloroplastic (*Zm00001d047462*) were all down-regulated only in the SM line. The role of translation and ribosome in adaptation to abiotic environment changes has been found in a recent analysis of the corresponding *Arabidopsis* mutants (Reiter et al., 2020). In particular, various examples of impaired cold tolerance due to the inactivation of chloroplast proteins involved in translation have been described, including subunits (Wang et al., 2017), biogenesis factors (Reiter et al., 2020) of the plastid ribosome-associated proteins (Pulido et al., 2018), translation initiation or elongation factors (Liu et al., 2010) and RNA-binding proteins (Kupsch et al., 2012). Therefore, down-regulated chloroplastic ribosomal protein-related genes in the SM line under LNT might cause the down-regulation of genes involved in the photosystem, thereby affecting photosynthesis and decreasing SPAD value.

Overall, we propose a possible network of maize seed germination under low-temperature stress affecting subsequent seedling growth (**Figure 8**). Maize seed germination under low-temperature stress caused the down-regulation of photosynthesis-related genes in both SM and RM lines, but the degree of down-regulation of the genes was lower in the RM line than in the SM line. Moreover, the SM line displayed the down-regulation of the ribosome and SOD-related genes, whereas genes involved in SOD and vitamin B6 were up-regulated in the RM line. SOD activity might play a key role in the TAC of maize seed germination under low-temperature stress because the changes in TAC were consistent with SOD activity. The inhibition of maize seed germination under low-temperature on seedling growth might be mainly due to impaired photosynthesis. The differences of TAC (especially SOD activity) among various lines might affect low-temperature resistance at the germination stage.

CONCLUSION

In summary, maize seed germination under low-temperature stress displayed an increase of lipid peroxidation and inhibited

subsequent seedling growth under normal temperature. Transcriptome analysis revealed that photosynthesis and antioxidant metabolism-related pathways played essential roles in seed germination in response to low-temperature stress at the germination stage, and the photosynthetic system of the SM line was more vulnerable to low-temperature stress. Moreover, SOD activity might play a key role in TAC in seed germination under low-temperature stress. Therefore, this study provides new insights into maize seed germination in response to low-temperature stress.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found at: NCBI, BioProject, PRJNA795360.

AUTHOR CONTRIBUTIONS

AM, DW, and CZ designed research. AM performed experiments. DW and CZ experimentally guided. AM and DW analyzed the results and wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Maize Industry Technology System in Shandong Province (SDAIT-02-02).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.843033/full#supplementary-material

REFERENCES

Bose, J., Rodrigo-Moreno, A., and Shabala, S. (2014). ROS homeostasis in halophytes in the context of salinity stress tolerance. *J. Exp. Bot.* 65, 1241–1257. doi: 10.1093/jxb/ert430

Cao, Q., Li, G., Cui, Z., Yang, F., Jiang, X., Diallo, L., et al. (2019). Seed priming with melatonin improves the seed germination of waxy maize under chilling stress via promoting the antioxidant system and starch metabolism. *Sci. Rep.* 9, 15044–15012. doi: 10.1038/s41598-019-51122-y

Checchio, M. V., de Cássia Alves, R., de Oliveira, K. R., Moro, G. V., dos Santos, D. M. M., and Gratão, P. L. (2021). Enhancement of salt tolerance in corn using *Azospirillum brasilense*: an approach on antioxidant systems. *J. Plant Res.* 134, 1279–1289. doi: 10.1007/s10265-021-01332-1

Cnubben, N. H. P., Rietjens, I. M. C. M., Wortelboer, H., Zanden, J.Van, and Van Bladeren, P. J. (2001). The interplay of glutathione-related processes in antioxidant defense. *Environ. Toxicol. Pharmacol.* 10, 141–152. doi:10.1016/S1382-6689(01)00077-1.

Denslow, S. A., Rueschhoff, E. E., and Daub, M. E. (2007). Regulation of the *Arabidopsis thaliana* vitamin B6 biosynthesis genes by abiotic stress. *Plant Physiol. Biochem.* 45, 152–161. doi: 10.1016/j.plaphy.2007.01.007

Fenza, M.Di, Hogg, B., Grant, J., and Barth, S. (2017). Transcriptomic response of maize primary roots to low temperatures at seedling emergence. *PeerJ* 5, e2839–e2817. doi:10.7717/peerj.2839.

Drewke, C., and Leistner, E. (2001). Biosynthesis of vitamin B6 and structurally related derivatives. *Vitam. Horm.* 61, 121–155. doi: 10.1016/s0083-6729(01)61004-5

Ehrenshaft, M., Chung, K. R., Jenns, A. E., and Daub, M. E. (1999). Functional characterization of *SOR1*, a gene required for resistance to photosensitizing toxins in the fungus *Cercospora nicotianae*. *Curr. Genet.* 34, 478–485. doi: 10.1007/s002940050423

Ehrenshaft, M., Jenns, A. E., Chung, K. R., and Daub, M. E. (1998). *SOR1*, a gene required for photosensitizer and singlet oxygen resistance in Cercospora fungi, is highly conserved in divergent organisms. *Mol. Cell* 1, 603–609. doi: 10.1016/S1097-2765(00)80060-X

Ensminger, I., Busch, F., and Huner, N. P. A. (2006). Photostasis and cold acclimation: sensing low temperature through photosynthesis. *Physiol. Plant.* 126, 28–44. doi: 10.1111/j.1399-3054.2006.00627.x

Garcia-Molina, A., Kleine, T., Schneider, K., Mühlhaus, T., Lehmann, M., and Leister, D. (2020). Translational components contribute to acclimation responses to high light. *Heat Cold Arabidopsis* 23, 1–21. doi: 10.1016/j.isci.2020.101331

Giannopolitis, C. N., and Ries, S. K. (1977). Superoxide dismutases. *Plant Physiol.* 59, 309–314. doi: 10.1104/pp.59.2.309

Grzybowski, M., Adamczyk, J., Jończyk, M., Sobkowiak, A., Szczepanik, J., Frankiewicz, K., et al. (2019). Increased photosensitivity at early growth as a possible mechanism of maize adaptation to cold springs. *J. Exp. Bot.* 70, 2887–2904. doi: 10.1093/jxb/erz096

Gu, L., Zhao, M., Ge, M., Zhu, S., Cheng, B., and Li, X. (2019). Transcriptome analysis reveals comprehensive responses to cadmium stress in maize inoculated with arbuscular mycorrhizal fungi. *Ecotoxicol. Environ. Saf.* 186, 109744–109749. doi: 10.1016/j.ecoenv.2019.109744

Handel, E. V. (1968). Direct microdetermination of sucrose. *Anal. Biochem.* 22, 280–283. doi: 10.1016/0003-2697(68)90317-5

Hodges, D. M., Andrews, C. J., Johnson, D. A., and Hamilton, R. I. (1997). Antioxidant enzyme and compound responses to chilling stress and their combining abilities in differentially sensitive maize hybrids. *Crop Sci.* 37, 857–863. doi: 10.2135/cropsci1997.0011183X003700030027x

Huang, S. H., Zhang, J. Y., Wang, L. H., and Huang, L. Q. (2013). Effect of abiotic stress on the abundance of different vitamin B6 vitamers in tobacco plants. *Plant Physiol. Biochem.* 66, 63–67. doi: 10.1016/j.plaphy.2013.02.010

Kupsch, C., Ruwe, H., Gusewski, S., Tillich, M., Small, I., and Schmitz-Linneweber, C. (2012). *Arabidopsis* chloroplast RNA binding proteins CP31A and CP29A associate with large transcript pools and confer cold stress tolerance by influencing multiple chloroplast RNA processing steps. *Plant Cell* 24, 4266–4280. doi: 10.1105/tpc.112.103002

Li, Y., Wang, X., Ban, Q., Zhu, X., Jiang, C., Wei, C., et al. (2019). Comparative transcriptomic analysis reveals gene expression associated with cold adaptation in the tea plant *Camellia sinensis*. *BMC Genomics* 20, 624–617. doi: 10.1186/s12864-019-5988-3

Liu, X., Rodermel, S. R., and Yu, F. (2010). A var2 leaf variegation suppressor locus, suppressor of variegation3, encodes a putative chloroplast translation elongation factor that is important for chloroplast development in the cold. *BMC Plant Biol.* 10, 1–18. doi: 10.1186/1471-2229-10-287

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Ma, Y., Dai, X., Xu, Y., Luo, W., Zheng, X., Zeng, D., et al. (2015). *COLD1* confers chilling tolerance in rice. *Cell* 160, 1209–1221. doi: 10.1016/j.cell.2015.01.046

Nägele, T., and Heyer, A. G. (2013). Approximating subcellular organisation of carbohydrate metabolism during cold acclimation in different natural accessions of *Arabidopsis thaliana*. *New Phytol.* 198, 777–787. doi: 10.1111/nph.12201

Nie, G. -Y., Long, S. P., and Baker, N. R. (1992). The effects of development at sub-optimal growth temperatures on photosynthetic capacity and susceptibility to chilling-dependent photoinhibition in *Zea mays*. *Physiol. Plant.* 85, 554–560. doi: 10.1111/j.1399-3054.1992.tb05826.x

Noctor, G., Mhamdi, A., Chaouch, S., Han, Y., Neukermans, J., Marquez-Garcia, B., et al. (2012). Glutathione in plants: an integrated overview. *Plant Cell Environ.* 35, 454–484. doi: 10.1111/j.1365-3040.2011.02400.x

Ohama, N., Sato, H., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2017). Transcriptional regulatory network of plant heat stress response. *Trends Plant Sci.* 22, 53–65. doi: 10.1016/j.tplants.2016.08.015

Pfister, A. S. (2019). Emerging role of the nucleolar stress response in autophagy. *Front. Cell. Neurosci.* 13, 1–18. doi: 10.3389/fncel.2019.00156

Prasad, T. K. (1997). Role of catalase in lnducing chilling tolerance in pre-emergent maize seedlings. *Plant Physiol.* 114, 1369–1376. doi: 10.1104/pp.114.4.1369

Pulido, P., Zagari, N., Manavski, N., Gawroński, P., Matthes, A., Scharff, L. B., et al. (2018). Chloroplast ribosome associated supports translation under stress and interacts with the ribosomal 30S subunit. *Plant Physiol.* 177, 1539–1554. doi: 10.1104/pp.18.00602

Reiter, B., Vamvaka, E., Marino, G., Kleine, T., Jahns, P., Bolle, C., et al. (2020). The Arabidopsis protein CGL20 is required for plastid 50S ribosome biogenesis. *Plant Physiol.* 182, 1222–1238. doi: 10.1104/pp.19.01502

Romero-Puertas, M. C., Corpas, F. J., Sandalio, L. M., Leterrier, M., Rodríguez-Serrano, M., Del Río, L. A., et al. (2006). Glutathione reductase from pea leaves: response to abiotic stress and characterization of the peroxisomal isozyme. *New Phytol.* 170, 43–52. doi: 10.1111/j.1469-8137.2006.01643.x

Savitch, L. V., Ivanov, A. G., Gudynaite-Savitch, L., Huner, N. P. A., and Simmonds, J. (2011). Cold stress effects on PSI photochemistry in *Zea mays*: differential increase of FQR-dependent cyclic electron flow and functional implications. *Plant Cell Physiol.* 52, 1042–1054. doi: 10.1093/pcp/pcr056

Scalzo, J., Politi, A., Pellegrini, N., Mezzetti, B., and Battino, M. (2005). Plant genotype affects total antioxidant capacity and phenolic contents in fruit. *Nutrition* 21, 207–213. doi: 10.1016/j.nut.2004.03.025

Schultes, N. P., Sawers, R. J. H., Brutnell, T. P., and Krueger, R. W. (2000). Maize high chlorophyll fluorescent 60 mutation is caused by an ac disruption of the gene encoding the chloroplast ribosomal small subunit protein 17. *Plant J.* 21, 317–327. doi: 10.1046/j.1365-313x.2000.00676.x

Sowiński, P., Rudzińska-Langwald, A., Adamczyk, J., Kubica, I., and Fronk, J. (2005). Recovery of maize seedling growth, development and photosynthetic efficiency after initial growth at low temperature. *J. Plant Physiol.* 162, 67–80. doi: 10.1016/j.jplph.2004.03.006

Wang, W., Zheng, K., Gong, X., Xu, J., Huang, J., Lin, D., et al. (2017). The rice *TCD11* encoding plastid ribosomal protein S6 is essential for chloroplast development at low temperature. *Plant Sci.* 259, 1–11. doi: 10.1016/j.plantsci.2017.02.007

Wen, D., Hou, H., Meng, A., Meng, J., Xie, L., and Zhang, C. (2018). Rapid evaluation of seed vigor by the absolute content of protein in seed within the same crop. *Sci. Rep.* 8, 5569–5568. doi: 10.1038/s41598-018-23909-y

Xu, T., Lee, K., Gu, L., Kim, J. I., and Kang, H. (2013). Functional characterization of a plastid-specific ribosomal protein PSRP2 in *Arabidopsis thaliana* under abiotic stress conditions. *Plant Physiol. Biochem.* 73, 405–411. doi: 10.1016/j.plaphy.2013.10.027

Yang, K., Yang, J., and Yi, J. (2018). Nucleolar stress: hallmarks, sensing mechanism and diseases. *Cell Stress* 2, 125–140. doi: 10.15698/cst2018.06.139

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). ClusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS J. Integr. Biol.* 16, 284–287. doi: 10.1089/omi.2011.0118

Yu, T., Zhang, J., Cao, J., Cai, Q., Li, X., Sun, Y., et al. (2021). Leaf transcriptomic response mediated by cold stress in two maize inbred lines with contrasting tolerance levels. *Genomics* 113, 782–794. doi: 10.1016/j.ygeno.2021.01.018

Zeng, R., Li, Z., Shi, Y., Fu, D., Yin, P., Cheng, J., et al. (2021). Natural variation in a type-A response regulator confers maize chilling tolerance. *Nat. Commun.* 12, 4713–4713. doi: 10.1038/s41467-021-25001-y

Zhang, Q., and Bartels, D. (2018). Molecular responses to dehydration and desiccation in desiccation-tolerant angiosperm plants. *J. Exp. Bot.* 69, 3211–3222. doi: 10.1093/jxb/erx489

Zhang, L., Pei, Y., Wang, H., Jin, Z., Liu, Z., Qiao, Z., et al. (2015a). Hydrogen sulfide alleviates cadmium-induced cell death through restraining ROS accumulation in roots of *Brassica rapa* L. *ssp. pekinensis*. *Oxidative Med. Cell. Longev.* 2015, 1–11. doi: 10.1155/2015/804603

Zhang, Z., Yang, J., and Wu, Y. (2015b). Transcriptional regulation of zein gene expression in maize through the additive and synergistic action of opaque2, prolamine-box binding factor, and O2 heterodimerizing proteins. *Plant Cell* 27, 1162–1172. doi: 10.1105/tpc.15.00035

Zhang, J., Yuan, H., Yang, Y., Fish, T., Lyi, S. M., Thannhauser, T. W., et al. (2016). Plastid ribosomal protein S5 is involved in photosynthesis, plant development, and cold stress tolerance in Arabidopsis. *J. Exp. Bot.* 67, 2731–2744. doi: 10.1093/jxb/erw106

Zhang, H., Zhang, J., Xu, Q., Wang, D., Di, H., Huang, J., et al. (2020). Identification of candidate tolerance genes to low-temperature during maize germination by GWAS and RNA-seq approaches. *BMC Plant Biol.* 20, 1–17. doi: 10.1186/s12870-020-02543-9

Zhao, Q., Zhou, L., Liu, J., Du, X., Asad, M. A. U., Huang, F., et al. (2018). Relationship of ROS accumulation and superoxide dismutase isozymes in developing anther with floret fertility of rice under heat stress. *Plant Physiol. Biochem.* 122, 90–101. doi: 10.1016/j.plaphy.2017.11.009

Zheng, Z., Xu, X., Crosley, R. A., Greenwalt, S. A., Sun, Y., Blakeslee, B., et al. (2010). The protein kinase SnRK2.6 mediates the regulation of sucrose metabolism and plant growth in arabidopsis. *Plant Physiol.* 153, 99–113. doi: 10.1104/pp.109.150789

Zheng, H., Yang, Z., Wang, W., Guo, S., Li, Z., Liu, K., et al. (2020). Transcriptome analysis of maize inbred lines differing in drought tolerance provides novel insights into the molecular mechanisms of drought responses in roots. *Plant Physiol. Biochem.* 149, 11–26. doi: 10.1016/j.plaphy.2020.01.027

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 04 March 2022
doi: 10.3389/fpls.2022.842726

Comparative Transcriptome Analysis Unravels Defense Pathways of *Fraxinus velutina* Torr Against Salt Stress

Xinmei Ma[1†], Jian Ning Liu[1†], Liping Yan[2†], Qiang Liang[1,3,4], Hongcheng Fang[1,3,4], Changxi Wang[1], Yuhui Dong[1,3,4], Zejia Chai[1], Rui Zhou[1], Yan Bao[1], Wenrui Hou[1], Ke Qiang Yang[1,3,4] and Dejun Wu[2]**

[1] College of Forestry, Shandong Agricultural University, Tai'an, China, [2] Shandong Provincial Academy of Forestry, Jinan, China, [3] State Forestry and Grassland Administration Key Laboratory of Silviculture in the Downstream Areas of the Yellow River, Shandong Agricultural University, Tai'an, China, [4] Shandong Taishan Forest Ecosystem Research Station, Shandong Agricultural University, Tai'an, China

OPEN ACCESS

Edited by:
Mukesh Jain,
Jawaharlal Nehru University, India

Reviewed by:
Parviz Heidari,
Shahrood University of Technology, Iran
Tushar Suhas Khare,
Savitribai Phule Pune University, India

***Correspondence:**
Ke Qiang Yang
yangwere@126.com
Dejun Wu
sdlky412x@163.com

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

Received: 24 December 2021
Accepted: 17 January 2022
Published: 04 March 2022

Citation:
Ma X, Liu JN, Yan L, Liang Q, Fang H, Wang C, Dong Y, Chai Z, Zhou R, Bao Y, Hou W, Yang KQ and Wu D (2022) Comparative Transcriptome Analysis Unravels Defense Pathways of Fraxinus velutina Torr Against Salt Stress. Front. Plant Sci. 13:842726. doi: 10.3389/fpls.2022.842726

Fraxinus velutina Torr with high salt tolerance has been widely grown in saline lands in the Yellow River Delta, China. However, the salt-tolerant mechanisms of *F. velutina* remain largely elusive. Here, we identified two contrasting cutting clones of *F. velutina*, R7 (salt-tolerant), and S4 (salt-sensitive) by measuring chlorophyll fluorescence characteristics (Fv/Fm ratio) in the excised leaves and physiological indexes in roots or leaves under salt treatment. To further explore the salt resistance mechanisms, we compared the transcriptomes of R7 and S4 from leaf and root tissues exposed to salt stress. The results showed that when the excised leaves of S4 and R7 were, respectively, exposed to 250 mM NaCl for 48 h, *Fv/Fm* ratio decreased significantly in S4 compared with R7, confirming that R7 is more tolerant to salt stress. Comparative transcriptome analysis showed that salt stress induced the significant upregulation of stress-responsive genes in R7, making important contributions to the high salt tolerance. Specifically, in the R7 leaves, salt stress markedly upregulated key genes involved in plant hormone signaling and mitogen-activated protein kinase signaling pathways; in the R7 roots, salt stress induced the upregulation of main genes involved in proline biosynthesis and starch and sucrose metabolism. In addition, 12 genes encoding antioxidant enzyme peroxidase were all significantly upregulated in both leaves and roots. Collectively, our findings revealed the crucial defense pathways underlying high salt tolerance of R7 through significant upregulation of some key genes involving metabolism and hub signaling pathways, thus providing novel insights into salt-tolerant *F. velutina* breeding.

Keywords: ***Fraxinus velutina*** **Torr, salt stress, comparative transcriptome, stress-responsive gene, defense response**

INTRODUCTION

Soil salinization has become a global major challenge. Saline soils are a major contributor to the regional fragile ecosystems, thereby posing a heavy burden on the sustainable advancement of local economies (Egamberdieva et al., 2019; Wani et al., 2020). As one kind of the most important tree species for afforestation, the *Fraxinus* species play crucial roles in ecological restoration (Yildiz et al., 2018; He et al., 2021). Thus, it is of importance to breeding high salt-tolerant *Fraxinus* species suitable for afforestation, especially in the areas with heavy saline-alkaline pollution. *Fraxinus velutina* Torr is a deciduous tree native to southwestern North America. Due to the rapid growth rate and superior salinity tolerance, the species has been intentionally introduced and widely planted in saline land in the Yellow River Delta, China (Mao et al., 2017). However, the molecular mechanisms underlying the high salt tolerance of *F. velutina* remain largely unknown.

Saline stress imposes primary stresses (ionic and osmotic) and secondary stresses (oxidative stress) on plants (Zhu, 2002; Yang and Guo, 2018a). Thus, plants mainly rely on the reestablishment of cellular homeostasis, namely, ionic, osmotic, and reactive oxygen species (ROS) to cope with salt stress. The process is complicated and involved in multiple signals and pathways, such as Ca^{2+}, plant hormone, and ROS signaling pathways (Yang and Guo, 2018b). After salt exposure, the excessive extracellular Na^{+} induces the generation of several second messengers (e.g., cytosolic Ca^{2+}), which are sensed by their sensors/receptors and protein kinases including mitogen-activated protein kinase (MAPK) and then transduce the stimuli signals into downstream components that switch on transcriptional cascades to defense against salt stress (Chinnusamy et al., 2004; Huang et al., 2012; Van Zelm et al., 2020; Chen et al., 2021). Plant hormones, namely, ethylene, salicylic acid, and abscisic acid (ABA) play irreplaceable roles in the defense response of plants to salt stress (Bari and Jones, 2009; Yu et al., 2020). For instance, salt stress can enhance ABA signaling and activate an ABA-dependent responsive complex to cope with saline stress in plants (Umezawa et al., 2009; Cai et al., 2017; Soma et al., 2017; Lin et al., 2021). In addition, salt stress induces the rapid production of ROS, which is sensed by ROS sensor/receptor and then transduced to regulate the defense response of plants under salt stress (Ashraf, 2009; Yang and Guo, 2018b). However, different plants or even different accessions of the same plant species respond differently to salt stress (Seki et al., 2003; Ji et al., 2013). Therefore, an in-depth understanding of the defense response of *F. velutina* against salt stress is an essential aid to breeding work on salt resistance.

Transcriptome profiling has been widely used to analyze salt-induced gene expression in various plants (Zhu et al., 2019; Zhang et al., 2020; Li et al., 2021; Ma et al., 2021). With the aid of RNA sequencing, significant pathways of differentially expressed genes (DEGs) are identified that would otherwise have been overlooked. In this study, two contrasting materials, *F. velutina* R7 (salt-tolerant) and S4 (salt-sensitive) were identified by measuring chlorophyll fluorescence characteristics (*Fv*/*Fm* ratio) in the excised leaves and physiological indexes in roots or leaves under salt treatment. To further explore the mechanisms underlying salt tolerance, a comparative transcriptome analysis was performed on the leaf and root of R7 and S4 clone exposed to salt stress. Our findings revealed the crucial defense response genes underlying high salt tolerance, thus providing insights into the salt-tolerant *F. velutina* breeding.

MATERIALS AND METHODS

Plant Materials

The *F. velutina* materials were obtained from the Experimental Base of Afforestation on Saline—Alkali Soil of Shandong Provincial Academy of Forestry, Shouguang City, Shandong Province, China (118°42′9.18″ E, 37°9′38.94″ N), China. The salt-tolerant *F. velutina* R7 accession and salt-sensitive *F. velutina* S4 accession were identified from 189 *F. velutina* accessions (unpublished data) by measuring *Fv*/*Fm* ratio on the excised leaves exposed to 250 mM NaCl for 48 h according to the previously described method (Smethurst et al., 2009). Due to easy availability and appropriate size of 1-year-old cuttings of *F. velutina* (height: 26.40 ± 0.50 cm) for the experiment. In addition, 1-year-old cuttings are frequently used as plant materials for salt resistance research (Gucci et al., 1997; Tsabarducas et al., 2015; Ran et al., 2021). So, 1-year-old cuttings of *F. velutina* R7 and S4 were used in this study. After gently removing the soil around the root, the cuttings were firstly precultivated in distilled water for 2 weeks and then transferred to a plastic container containing 6 L of half-strength Hoagland's solution for 4 weeks. The solutions were refreshed every 7 days. The cuttings were grown in a growth incubator (LICHEN, Shanghai, China) under 25/20°C (day/night temperature), 65% relative humidity, 16 h light (1,200 μmol m^{-2} s^{-1})/8 h dark.

Chlorophyll Fluorescence

As previously described (Gabilly et al., 2019), chlorophyll fluorescence parameters (*Fv*/*Fm* ratio) were determined using a pulse-amplitude modulated chlorophyll fluorometer (FMS2, Hansatech Instruments, Pentney King's Lynn, United Kingdom). The images reflecting chlorophyll fluorescence parameters were acquired with FluorCam imaging fluorimeters (Photon Systems Instruments, Brno, Czech Republic) according to the instructions of the manufacturer.

Salt Treatment

After 6 weeks of acclimatization, the healthy cuttings with appropriate size (height: 26.40 ± 0.50 cm) for this experiment were selected and exposed to the solution containing 250 mM NaCl for 12 h, and the clone without NaCl treatment were considered as control. The NaCl concentration and exposure time were selected based on our preliminary tests and previous studies (Li et al., 2012; Yan et al., 2019; Catalá et al., 2021). The leaves and roots of each clone from three independent biological replicates were harvested and stored in -80°C until being used.

For easy understanding, the letters "S" and "C" were used to represent the salt-treated samples and control, respectively, while "L" and "R" represented the leaves and roots samples, respectively.

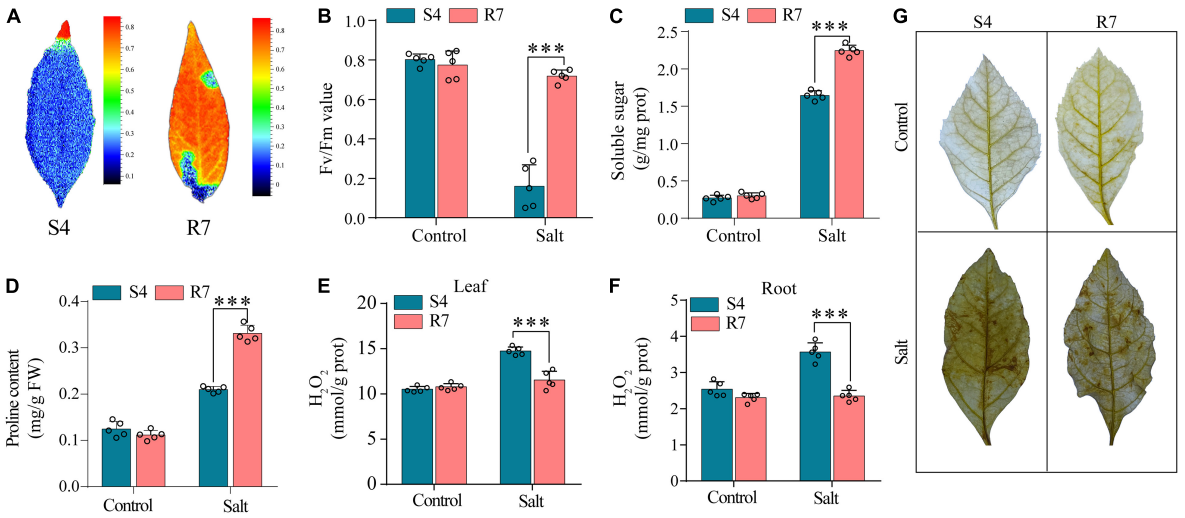


FIGURE 1 | Determination of selected biochemical parameters in *F. velutina* R7 and S4. **(A)** The images reflecting chlorophyll fluorescence parameters. **(B)** Chlorophyll fluorescence parameters (*Fv/Fm* ratio) of the excised leaves of *F. velutina* R7 and S4 exposed to 250 mM NaCl for 48 h. Soluble sugar **(C,D)** proline content in R7 and S4 roots with or without salt treatment. **(E,F)** H_2O_2 content in leaves **(E)** and roots **(F)** of R7 and S4 before and after salt stress. **(G)** *In situ* visualization of H_2O_2 accumulation by DAB staining in R7 and S4 leaves before and after salt stress. Error bars represented SD ($n = 5$). ***$P < 0.001$ represented significant differences.

For instance, the sample R7SL represented the leaves of R7 clone treated with NaCl solution.

Physiological Parameter Measurements

The contents of proline, soluble sugar, and H_2O_2 were determined using the commercial kits purchased from Nanjing Jiangcheng Bioengineering Institute (Nanjing, China) following the instructions of the manufacturer. Histochemical location of H_2O_2 was conducted by staining with 3,3′-diaminobenzidine (DAB) according to the previously described (Daudi and O'Brien, 2012). Cell death was determined by trypan blue staining as previously described (Pogány et al., 2009). Each sample group contained five biological replicates.

Transcriptome Sequencing

Total RNA was extracted using the GeneJET Plant RNA Purification Mini Kit (Thermo Fisher Scientific, Waltham, Massachusetts, United States). The concentration and integrity of RNA were evaluated using a NanoDrop ND-2000 (Thermo Fisher Scientific, Waltham, Massachusetts, United States) and Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, California, United States), respectively. About 1 μg of qualified total RNA was used to construct cDNA libraries with an insert size of 350 bp using the TruSeq RNA Sample Preparation Kit v2 (Illumina, San Diego, California, United States). The qualified libraries were sequenced on an Illumina NovaSeq 6000 platform (KeGene Science and Technology Corporation Ltd., Shandong, China) with a paired-end 150 mode.

Transcriptome Data Analysis

The raw reads obtained were trimmed using Trimmomatic v0.39 (Bolger et al., 2014), and the high-quality reads obtained were mapped to velvet ash reference genome[1] (Kelly et al., 2020) using HISAT2 version 2.2.1 software (Kim et al., 2015). The transcripts were assembled and quantified using StringTie v2.1.5 (Pertea et al., 2015), and the gene expression levels were measured using fragments per kilobase of transcript per million fragments mapped (FPKM). The genes with FPKM ≥ 5 in at least one sample were selected to perform differential gene expression analysis using DESeq2 v4.1 (Love et al., 2014) based on $|\log2$ (fold change)$| \geq 2$ and false discovery rate < 0.01.

Gene Set Enrichment Analysis

The Gene Ontology (GO) enrichment analysis of the DEGs was performed using the Cytoscape v3.9.0 plug-in ClueGO v2.5.8 (Bindea et al., 2009). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of the DEGs was carried out using clusterProfiler v4.2.0 (Yu et al., 2012). The transcription factors were identified by aligning all the transcripts obtained against the plant TF database PlantTFDB v5.0 (Tian et al., 2020).

Quantitative Real-Time Reverse Transcription PCR Analysis

To validate the reliability of transcriptome sequencing data, 16 DEGs were randomly selected to perform quantitative real-time (qRT)-PCR analysis. The CFX Connect Real-Time System (Bio-Rad, Hercules, CA, United States) was used for qRT-PCR. The PCR assays were conducted as previously described (Fang et al., 2021). The actin genes were chosen as internal reference (Li et al., 2013, 2019). The genes were quantified using the $2^{-\Delta\Delta CT}$ method (Arocho et al., 2006). Each sample group

[1]http://ashgenome.org/

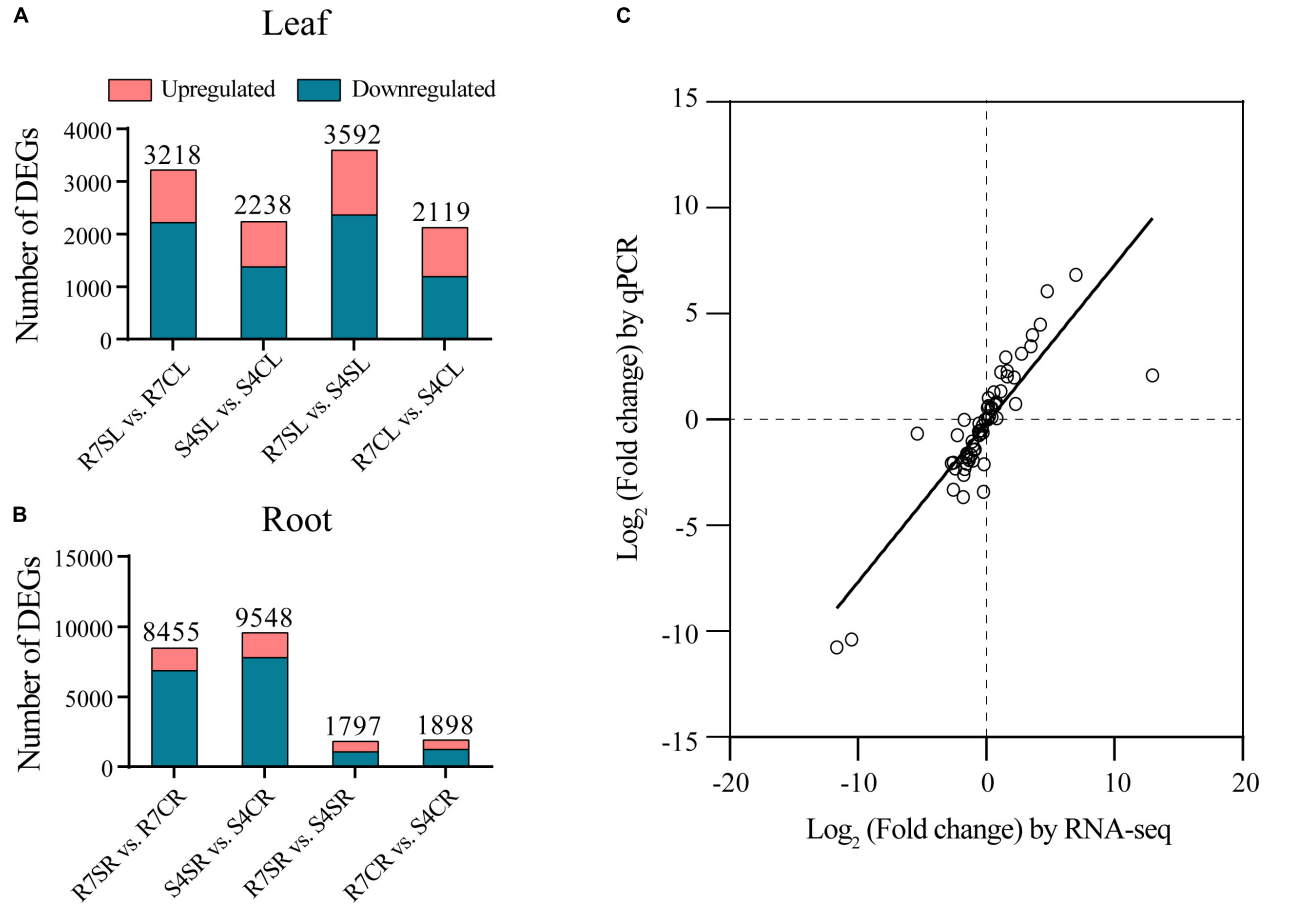


FIGURE 2 | Transcriptome data and the differentially expressed genes (DEGs) in *F. velutina* R7 and S4 plants with or without salt treatment. **(A)** The DEG number of each pairwise comparison in leaf samples. **(B)** The DEG number of each pairwise comparison in root samples. **(C)** Correlation between qPCR (*X*-axis) and RNA-seq (*Y*-axis) was calculated based on log2 (fold change).

contained three biological replicates. The primers are listed in **Supplementary Table 1**.

Statistical Analysis

Statistical data were represented as mean ± SD. The Student's *t*-test was used to determine the differences between the two groups. The statistical analyses were performed with GraphPad Prism v9.0 (GraphPad Software Inc., La Jolla, United States). $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$ represented statistical significance.

RESULTS

Fraxinus velutina R7 Accession Is More Salt Tolerant Than S4

To evaluate the salt tolerance between R7 and S4, the *Fv*/*Fm* parameters in the salt-treated excised leaves were determined. The results showed that when the excised leaves of S4 and R7 were, respectively, exposed to 250 mM NaCl for 48 h, *Fv*/*Fm* ratio decreased significantly in S4 than in R7 (**Figures 1A,B**), suggesting that R7 is more salt tolerant.

Several selected biochemical parameters, such as soluble sugars and proline content in roots, and H_2O_2 content (both leaves and roots) of R7 and S4 clone exposed to salt stress were determined. The results showed that the levels of soluble sugars and proline in R7 roots were significantly higher than in S4 (**Figures 1C,D**). Moreover, R7 leaves and roots presented lower levels of H_2O_2 than S4 under salt stress (**Figures 1E,F**), which was further confirmed by the histochemical staining (**Figure 1G**). The above results further demonstrated that R7 is more tolerant to salt stress.

Transcriptome Sequencing of R7 and S4 Accessions

To explore the differences between R7 and S4 accessions regarding gene expressions, a comparative transcriptome analysis was performed on the leaf and root of R7 and S4 clone exposed to salt stress. The results showed that 1.01 billion clean reads were obtained, with an average Q30 value was 98%. Then the clean reads were mapped to the *F. velutina* genome assembly and the

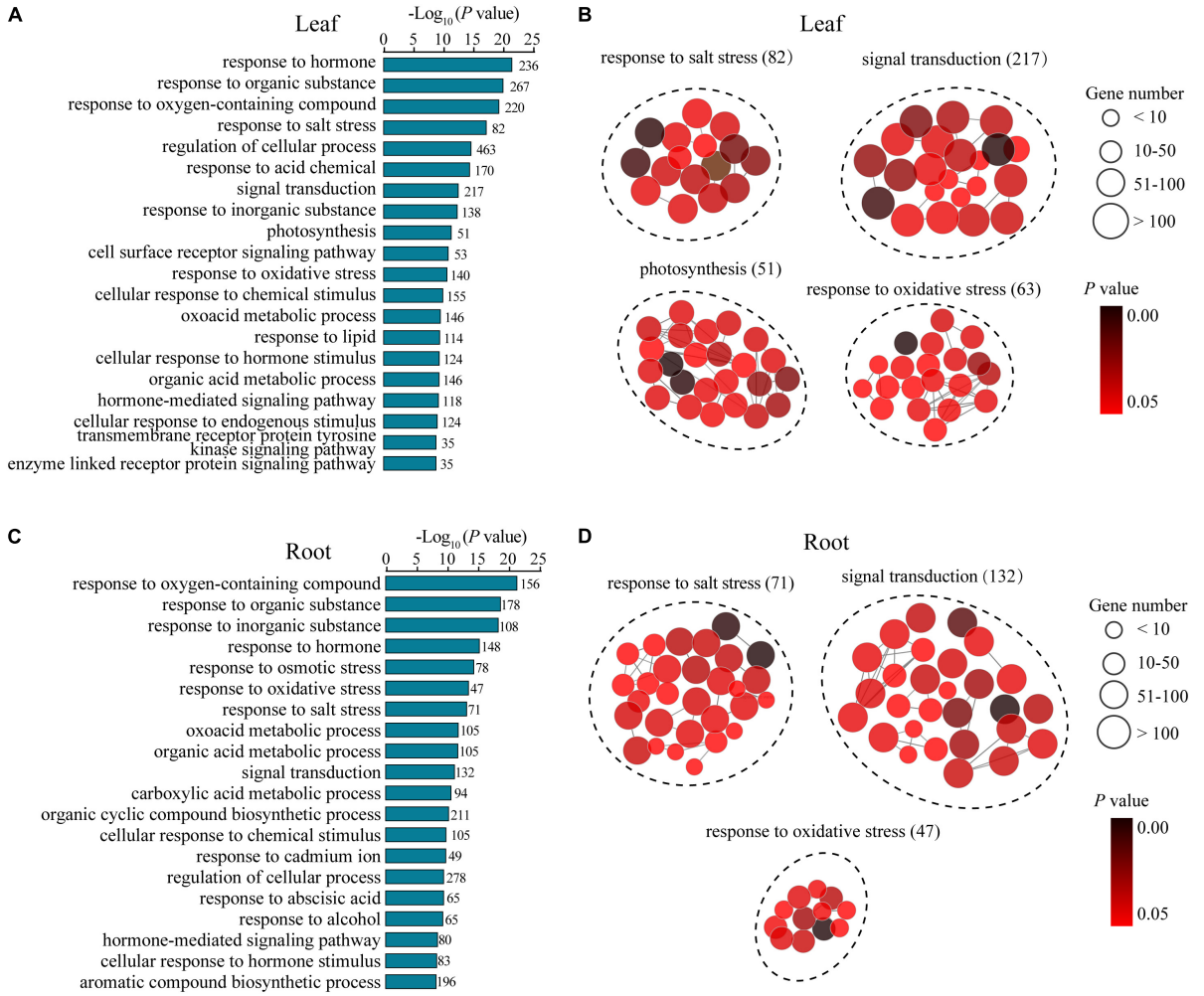


FIGURE 3 | The Gene Ontology (GO) enrichment analysis of the differentially expressed genes (DEGs) in leaves and roots. **(A,B)** the GO terms of the DEGs generated from comparison R7SL vs. S4SL in leaves. **(C,D)** the GO terms of the DEGs generated from comparison R7SR vs. S4SR in leaves.

results showed that 73.96–94.85% of total reads were mapped to the genome. After transcript assembly, 32,887 genes were finally produced by RNA sequencing (**Supplementary Table 3**).

To evaluate the consistency among the biological replicates, hierarchical clustering of all samples based on the correlation coefficient γ^2 between each sample was performed. The results showed that leaf and root samples were clustered individually whereas the three biological replicates of each group were clustered together (**Supplementary Figure 1**), indicating that the biological replicates in each group are highly consistent.

To identify salt-responsive genes, the differential gene expression analysis of pairwise comparisons R7SL vs. R7CL, S4SL vs. S4CL, R7SL vs. S4SL, and R7CL vs. S4CL in leaves and R7SR vs. R7CR, S4SR vs. S4CR, R7SR vs. S4SR, and R7CR vs. S4CR in roots were performed, respectively (**Supplementary Table 4**). The results showed that 3,218, 2,238, 3,592, and 2,119 DEGs were identified in comparisons R7SL vs. R7CL, S4SL vs. S4CL, R7SL vs. S4SL, and R7CL vs. S4CL in leaves, respectively (**Figure 2A** and **Supplementary Figure 2**). There were 8,455, 9,548, 1,797, and 1,898 DEGs in comparisons R7SR vs. R7CR, S4SR vs. S4CR, R7SR vs. S4SR, and R7CR vs. S4CR in roots, respectively (**Figure 2B** and **Supplementary Figure 3**).

To validate the reliability of RNA sequencing data, 16 DEGs were randomly selected to perform qRT-PCR analysis. The results showed a high correlation coefficient (R^2 = 0.7266) between RNA sequencing data and qRT-PCR results, indicating that the RNA sequencing data are reliable (**Figure 2C**). We, therefore, concluded that the global transcriptome in R7 is altered than in S4.

Fraxinus velutina R7 Enriched Stress-Responsive Genes Under Salt Stress

To understand the potential mechanisms underlying the distinguishing R7 and S4 in response to salt stress, a GO enrichment analysis of the DEGs in the comparisons R7SL vs. S4SL and R7SR vs. S4SR was performed, respectively.

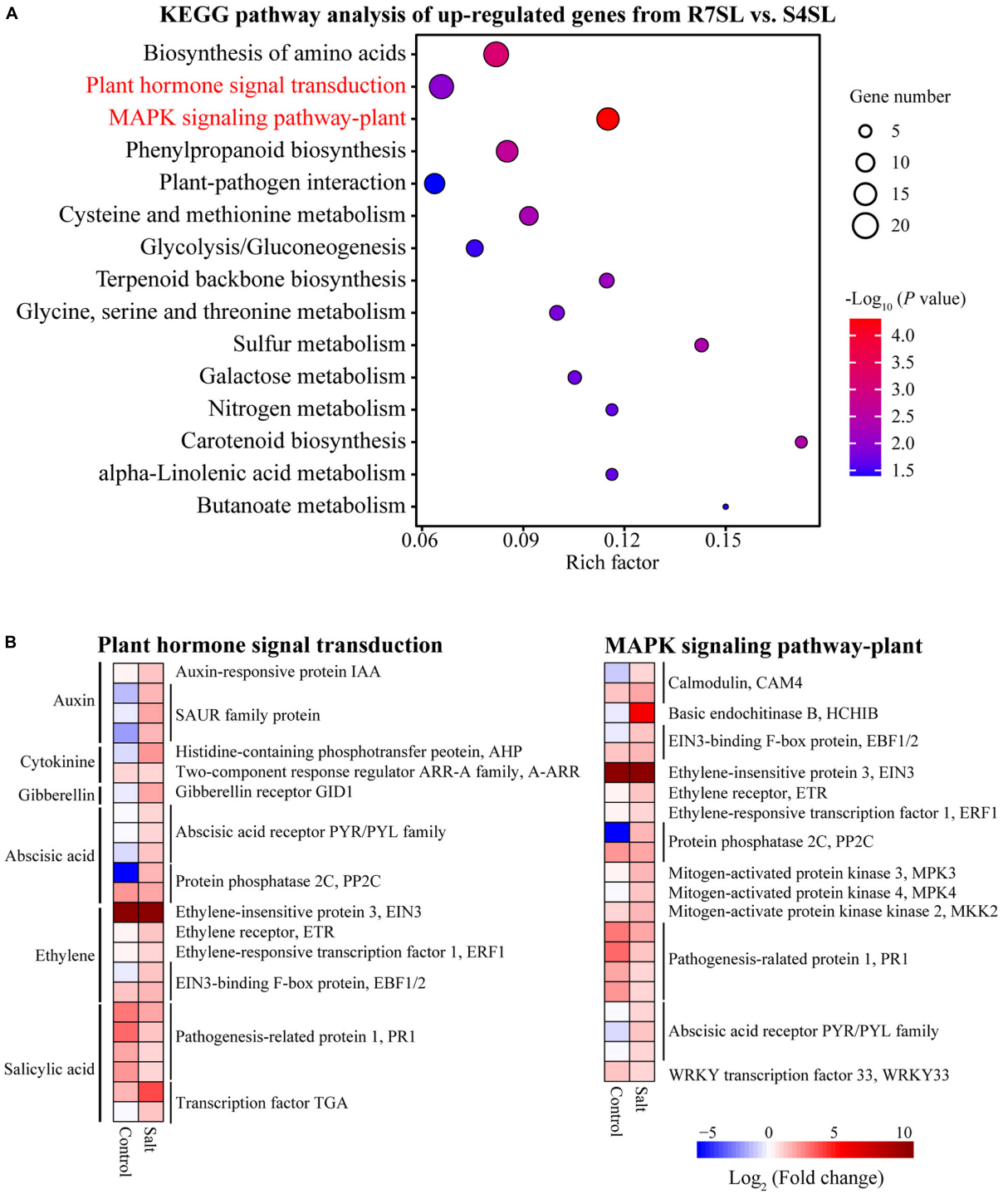


FIGURE 4 | The Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of the differentially upregulated genes from leaves. **(A)** The top 15 enriched pathways (ranked by *P*-value) of upregulated genes in comparison R7SL vs. S4SL were showed. The size of dot represented the number of DEGs. From blue to red represented the *P*-value from low to high. **(B)** Heatmap showed the expression patterns of DEGs related to plant hormone signal transduction and MAPK signaling pathway pant. From blue to red represented the fold change in tolerant leaves compared to sensitive leaves from low to high.

The results showed that multiple stress-associated GO terms were significantly enriched in both leaves and roots, such as response to an inorganic substance, response to salt stress, response to an organic substance, and response to the hormone (**Figures 3A,C**). Several GO terms related to oxidative stress, signal transduction, and metabolic process were also

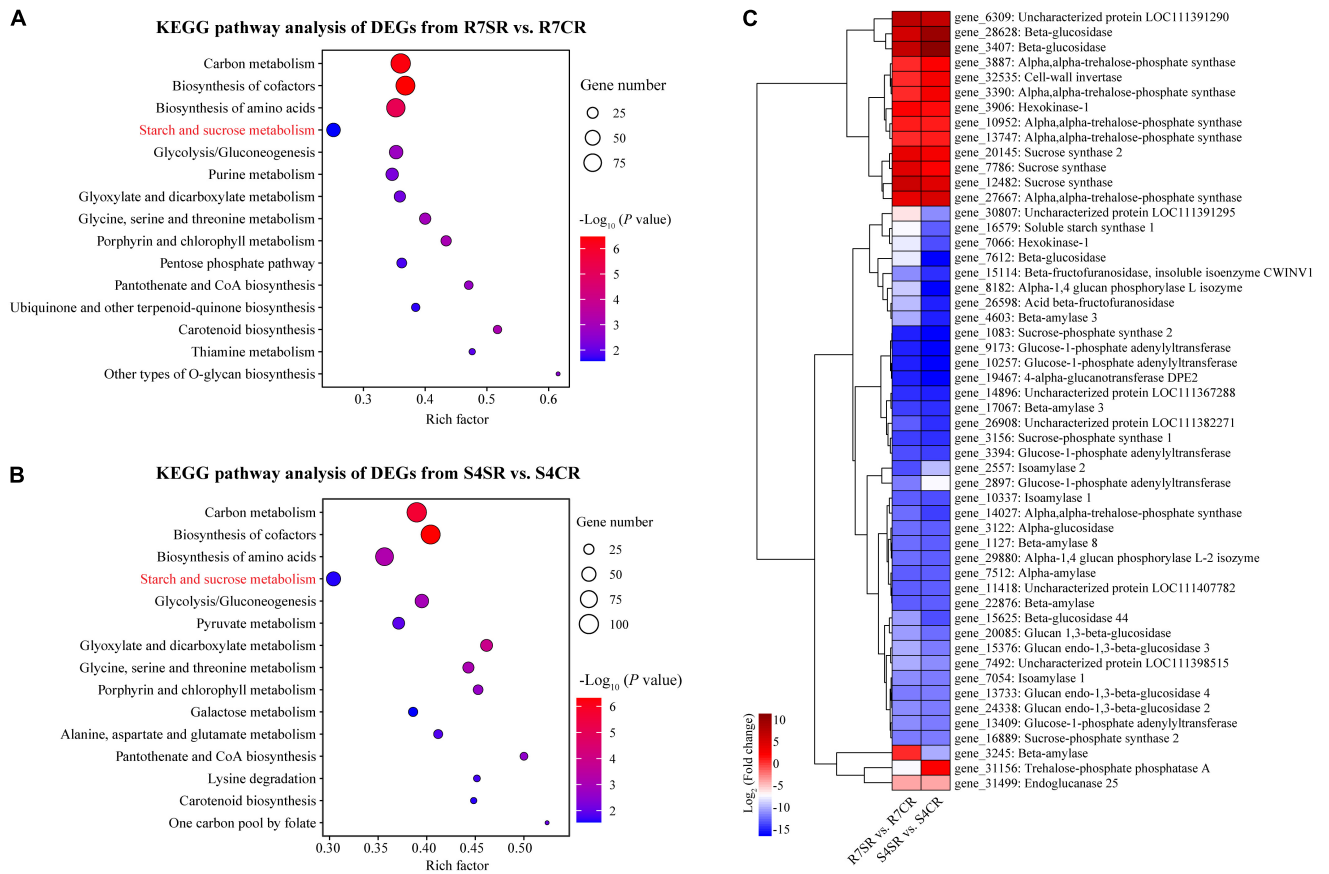


FIGURE 5 | Functional enrichment analysis of the DEGs and measuring soluble sugar contents in *F. velutina* R7 and S4 roots. **(A,B)** The top 15 enriched KEGG pathways (ranked by *P*-value) for DEGs in response to salt stress in tolerant **(A)** and sensitive roots **(B)** were showed. **(C)** The expression patterns of DEGs associated with "starch and sucrose metabolism" pathway in R7 and S4 roots underlying salt conditions. Heatmap colors represented the fold change and from blue to red represented the value from low to high.

enriched in both leaves and roots. In addition, several GO terms, namely, photosynthesis, transmembrane receptor protein tyrosine kinase signaling pathway, and cell surface receptor signaling pathway were specifically enriched in leaves, whereas the terms including response to cadmium ion and response to osmotic stress were peculiarly enriched in roots (**Figures 3B,D** and **Supplementary Table 5**). These results suggested that salt stress induces the higher enrichment of specific stress-responsive genes in R7 than in S4.

Fraxinus velutina R7 Leaves Enriched Hormonal and MAPK Signaling Pathways Under Salt Stress

To identify the signaling pathways involved in enhancing the salt tolerance of R7, the KEGG pathway enrichment analysis was performed using upregulated and downregulated DEGs from the comparison R7SL vs. S4SL. The results showed that the upregulated DEGs were mainly enriched in "plant hormone signal transduction" and "MAPK signaling pathway-plant" (**Figure 4A** and **Supplementary Table 6**), whereas the downregulated DEGs were enriched in "photosynthesis" and "photosynthesis-antenna proteins" (**Supplementary Figure 4**).

There were 23 DEGs involved in "plant hormone signal transduction," namely, the signaling network of auxin (IAA and SAUR), cytokinin (AHP and A-ARR), gibberellin (GID1), abscisic acid (PYR/PYL and PP2C), ethylene (EIN3, ETR, ERF1, and EBF1/2), and salicylic acid (PR1and TGA). In addition, there were 21 DEGs involved in the "MAPK signaling pathway plant." All these genes exhibited upregulation in the leaves of *F. velutina* R7 in comparison to the one in S4 under salt treatment (**Figure 4B**). The findings indicated that hormonal and MAPK signaling pathways are altered in the leaves of R7 after salt exposure.

Fraxinus velutina R7 Roots Accumulated More Soluble Sugar and Proline Than S4 Under Salt Stress

To explore the differences between R7 and S4 roots in response to salt stress, KEGG pathway enrichment analysis was performed based on the DEGs from comparisons R7SR vs. R7CR and S4SR vs. S4CR, respectively (**Supplementary Table 7**). The results

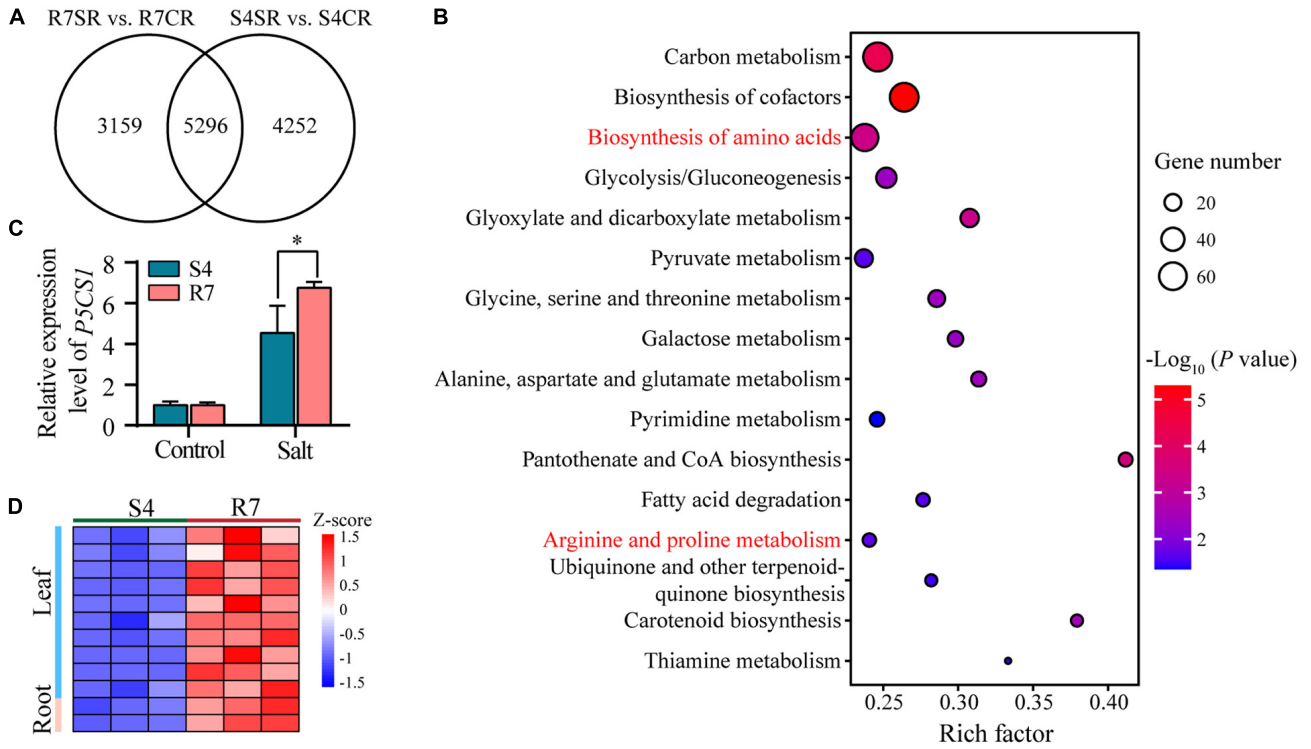


FIGURE 6 | The KEGG pathway enrichment analysis of the shared DEGs between comparisons R7SR vs. R7CR and S4SR vs. S4CR and further analysis of proline metabolism. **(A)** Venn diagram of DEGs in response to salt stress in *F. velutina* R7 and S4 roots. **(B)** The significantly enriched pathways ($P < 0.05$). Size of dot represented the number of DEGs and from blue to red represented the *P*-value from low to high. **(C)** Detecting the relative expression levels of *P5CS1* in *F. velutina* R7 and S4 roots with or without salt treatment. **(D)** Relative expression levels of a peroxidase (POD) gene in R7 and S4 roots with or without salt treatment. Heatmap color represented z-score and from blue to red represented the value from low to high. *$P < 0.05$ represented significant differences.

showed that three pathways were overlapped between two comparisons included "carbon metabolism," "biosynthesis of amino acids," and "starch and sucrose metabolism" (**Figures 5A,B**).

Starch and sucrose metabolism can regulate osmotic adjustment *via* determining the contents of soluble sugars, playing critical roles in salt tolerance (Dahro et al., 2016; Wei et al., 2019). Thus, the expression patterns of DEGs in "starch and sucrose metabolism" in R7 and S4 roots under salt stress were further analyzed. The results showed that the expression patterns of several genes were different between R7 and S4 roots under salt stress (**Figure 5C**), which was consistent with the significantly elevated levels of soluble sugars in R7 roots than in S4 (**Figure 1C**), suggesting that starch and sucrose metabolism induced by salt stress in R7 and S4 roots is different.

To explore the common mechanisms in response to salt stress between R7 and S4 roots, an intersection analysis between the comparisons R7SR vs. R7CR and S4SR vs. S4CR was performed. The results identified 5,296 shared DEGs in *F. velutina* R7 and S4 roots under salt stress (**Figure 6A**). KEGG enrichment analysis of these genes revealed 16 significantly enriched pathways, such as "biosynthesis of amino acids" and "Arginine and proline metabolism" (**Figure 6B** and **Supplementary Table 8**), suggesting that accumulation of amino acids might improve the salt tolerance in both *F. velutina* R7 and S4 roots. Proline acting as an osmoprotectant plays an important role in improving salt tolerance in the plant (Liang et al., 2018; Rady et al., 2019). As a rate-limiting enzyme involved in proline biosynthesis (Turchetto-Zolet et al., 2009), a gene encoding delta-1-pyrroline-5-carboxylate synthase 1 (P5CS1) was differentially upregulated to higher levels in R7 roots than in S4 (**Figure 6C**), which was consistent with the significantly elevated levels of proline in R7 roots than in S4 (**Figure 1D**).

Fraxinus velutina R7 Roots and Leaves Accumulated Less Reactive Oxygen Species Than S4

The increase in ROS levels leads to oxidative stress, posing detrimental effects on plant cells and tissues (Choudhury et al., 2017). Therefore, the DEGs involved in response to oxidative stress were analyzed further *via* dissecting the oxidative stress-related GO terms. Among the genes, 12 encoding peroxidases (PODs) were upregulated higher in R7 leaves (10 genes) and roots (2 genes) than in S4 (**Figure 6D** and **Supplementary Table 4**) under salt stress. Consistently, R7 leaves and roots presented lower levels of H_2O_2 than S4 under salt stress (**Figures 1E–G**). Collectively, these results suggested that

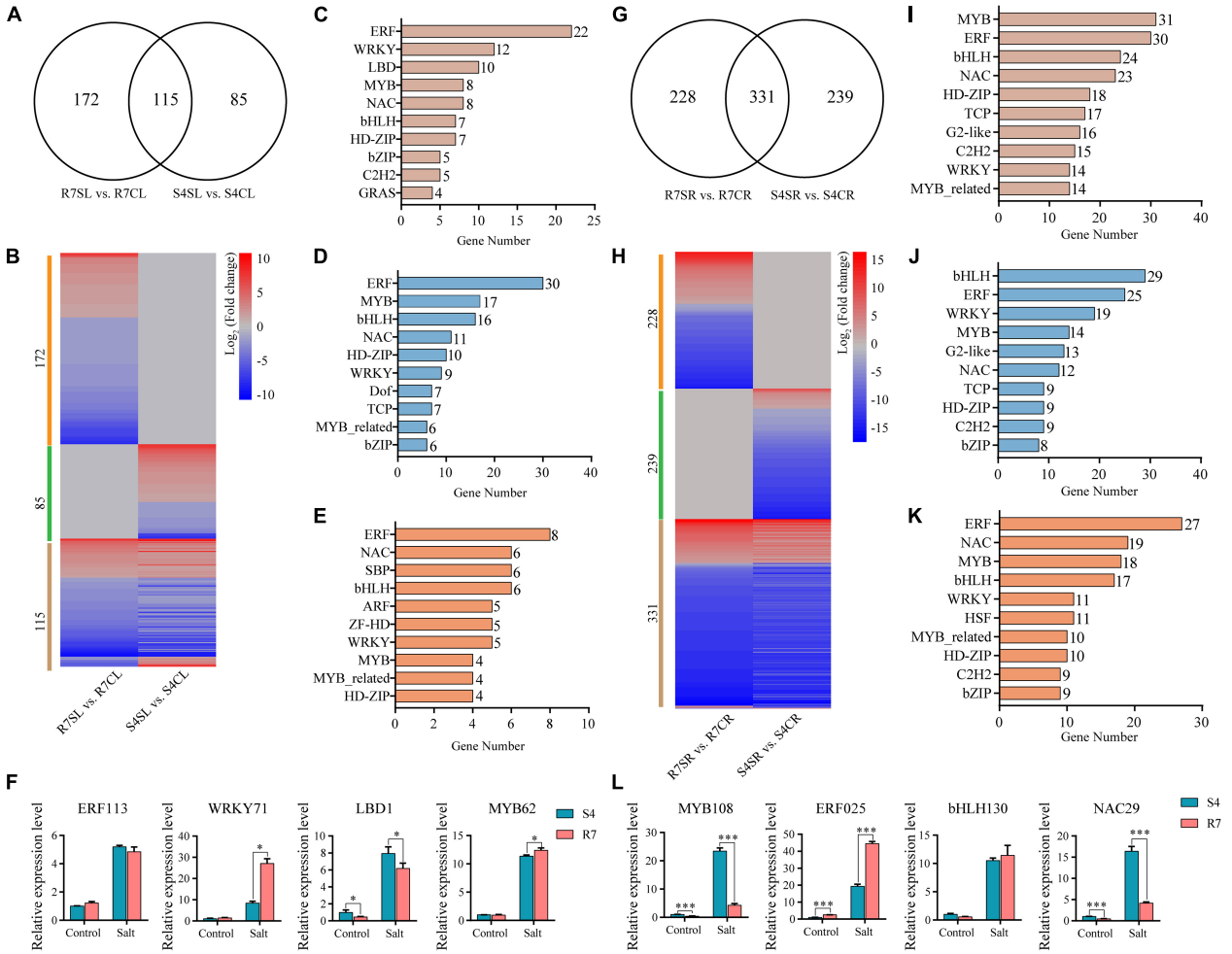


FIGURE 7 | Analysis of the differentially expressed transcription factors (TFs) in R7 and S4. **(A,G)** Venn diagrams showed the distribution of salt-induced TFs in R7 and S4 leaves and roots. **(B,H)** Heatmap showed the fold change of salt-treated R7 and S4 roots and leaves compared to their corresponding non-treated samples. **(A–E, I–K)** The top 10 TF families (ranked by gene number) in specific and common TFs among different comparisons. **(F,L)** The relative expression levels of the shared differentially expressed TFs in leaves and roots. $^{*}P < 0.05$ and $^{***}P < 0.001$ represented significant differences.

R7 exhibits higher ROS scavenging capacity than S4 under salt stress.

Several Transcription Factors Were Involved in Defense Response of *Fraxinus velutina* R7 Against Salt Stress

Transcription factors (TFs) act as the master switches in regulating multiple downstream target genes, thus playing crucial roles in various biological processes including salt stress response. By dissecting the RNA sequencing data, multiple TF genes were significantly altered between R7 and S4 under salt stress (**Figure 7** and **Supplementary Tables 9**, **10**). Among the significantly expressed TFs, 172 and 85 TFs were specifically expressed in R7SL vs. R7CL and S4SL vs. S4CL, respectively, and 115 were overlapped in both R7 and S4 (**Figures 7A,B**). Of the shared TFs, ERF, WRKY, LBD, and MYB were the most represented TF families induced by salt stress (**Figure 7C**).

In leaves under salt stress, 30 genes were encoding ERF, 17 encoding MYB, 16 encoding bHLH, and 11 encoding NAC prominently expressed in R7 (**Figure 7D**), whereas there were 8 encoding ERF, 6 encoding NAC, 6 encoding SBP, and 6 encoding bHLH prominently expressed in S4 (**Figure 7E**). In root tissues, bHLH, ERF, WRKY, and MYB TF families were prominently induced by salt stress in both R7 and S4 (**Figures 7H,I**). There were 29 genes encoding bHLH, 25 encoding ERF, and 19 encoding WRKY preferentially induced in R7 (**Figure 7J**), whereas 27 genes were encoding ERF, 19 encoding NAC, and 18 encoding MYB preferentially induced in S4 (**Figure 7K**).

To further confirm the roles of these TFs in regulating the stress response, 8 TF genes, namely, *ERF113*, *WRKY71*, *LBD1*, and *MYB62* in leaves and *MYB108*, *ERF025*, *bHLH130*, and *NAC29* in roots were chosen to perform qRT-PCR analysis. The results showed that both *WRKY71* and *MYB62* were significantly upregulated higher in R7 leaf than in S4, while *LBD1* was downregulated in R7 compared with S4 (**Figure 7F**). In addition,

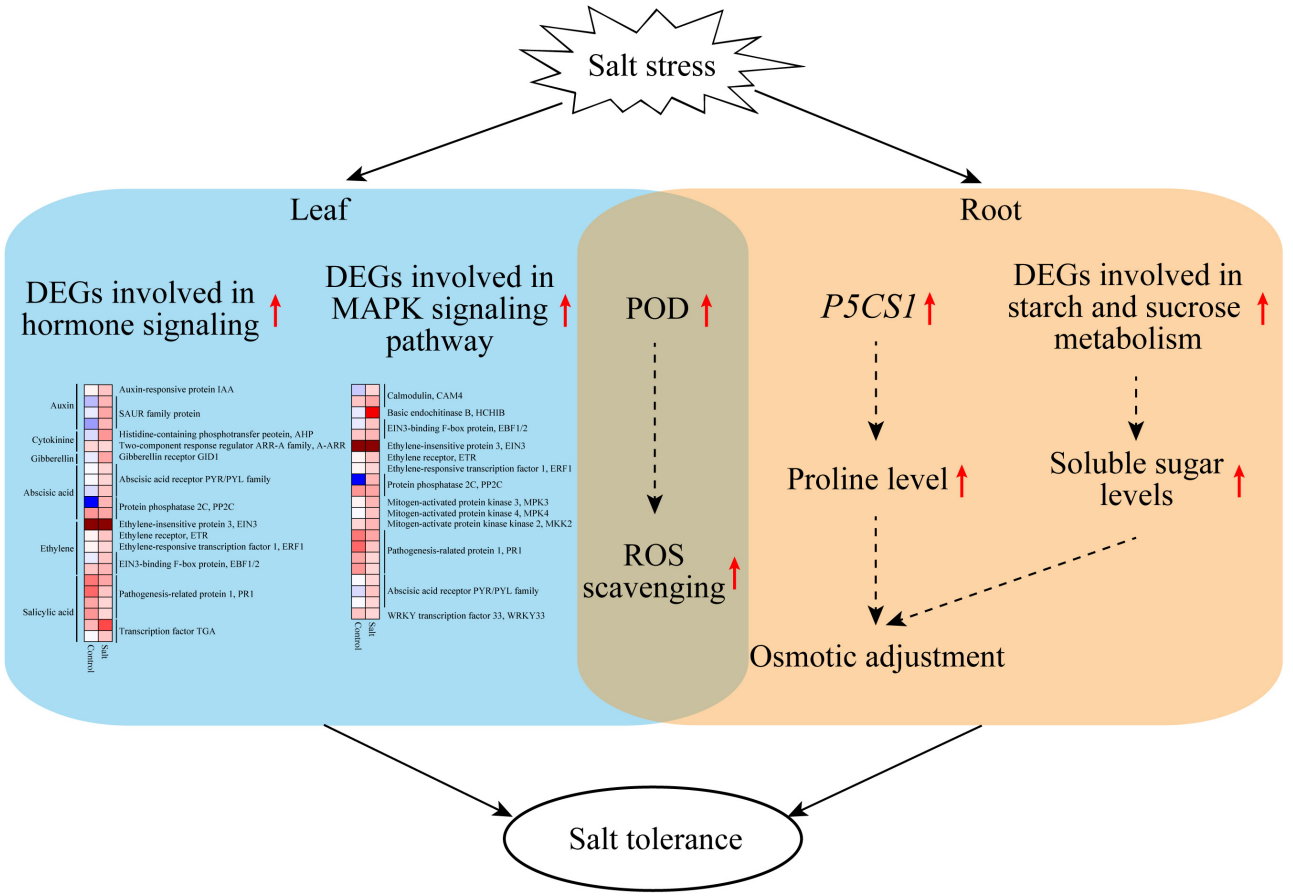


FIGURE 8 | A model proposed potentially to explain the high salt tolerance of *F. velutina* R7. Under salt stress, R7 activates multifaceted defense machines in the roots and leaves to enhance salt tolerance. In the R7 root, salt stress induces the upregulation of *P5CS1* and the key genes involved in starch and sucrose metabolism, therefore increasing osmotic adjustment. In the R7 leaf, salt stress induces the upregulation of important genes involved in plant hormone signaling (especially *PYR1/PYL/RCAR* ABA receptor genes and *PP2C*) and MAPK signaling pathways (especially *CAM4*, *MPK3*, and *WRKY33*). In addition, salt stress induces the upregulation of *POD* genes in both roots and leaves, so enhancing ROS scavenging capacity.

MYB108 and *NAC29* were found downregulated in R7 root than in S4, whereas *ERF025* was significantly upregulated in R7 in comparison to S4 (**Figure 7L**). These results suggested that the stress-responsive TFs, such as *WRKY71*, *MYB62*, and *ERF025* are the important contributors to enhancing salt tolerance in R7.

DISCUSSION

In this study, to explore the mechanisms underlying salt tolerance of *F. velutina*, a comparative transcriptome analysis was performed on the leaf and root of two contrasting materials, *F. velutina* R7 (salt-tolerant) and S4 (salt-sensitive) clone exposed to salt stress. The results showed that the high salt tolerance of R7 is mainly attributed to the enrichment of stress-responsive genes. The stress-responsive genes in the R7 leaf were associated with plant hormone signaling and MAPK signaling pathways, whereas the genes in the root were involved in proline biosynthesis and starch and sucrose metabolism. The genes encoding POD were upregulated after salt exposure, resulting in high ROS scavenging capacity in R7.

Plant hormones are the important contributors of plant growth and developmental processes and play crucial roles in biotic and abiotic stress responses (Wang et al., 2013; Xia et al., 2015; Kaleem et al., 2018). In this study, the pathway "plant hormone signal transduction" was significantly enriched in the R7 leaf, which included 23 DEGs related to auxin, cytokinin, gibberellin, abscisic acid, ethylene, and salicylic acid. Our results were consistent with the previous studies showed that these hormones play irreplaceable roles in the response to salt stress in plants (Zheng et al., 2018; Feng et al., 2019; Gao et al., 2021; Huang et al., 2021; Saini et al., 2021; Zhu et al., 2021). For instance, salt stress can induce ABA signaling and activate an ABA-dependent responsive complex to cope with saline stress in plants (Umezawa et al., 2009; Cai et al., 2017; Soma et al., 2017; Lin et al., 2021). In this study, 5 DEGs related to ABA signal transduction were significantly affected under salt stress, such the significantly upregulated genes encoding protein phosphatase (PP2C), which has been confirmed that the upregulation of *PP2C*

can lead to enhanced salt tolerance in *Arabidopsis* and maize (Liu et al., 2012). These findings suggested that the activation of hormone signaling improves salt tolerance in plants. In addition, we found that *PYR1/PYL/RCAR* ABA receptor genes and *PP2C* were the top salt responsive genes, suggesting that these genes are the candidate genes for developing salinity resilience through genetic engineering in the future.

The stress-responsive genes in the R7 leaf were also involved in the MAPK signaling pathway, which is documented to modulate plant tolerance to various abiotic stress, such as salt stress (De Zelicourt et al., 2016). The pathway included multiple genes that respond to salt stress, such as calmodulin (*CAM4*), mitogen-activated protein kinase 3 (*MPK3*), and WRKY transcription factor 33 (*WRKY33*). The overexpression of *MsCML46* in tobacco can lead to enhanced tolerance to multiple stresses, namely, drought, freezing, and salt stress (Du et al., 2021). In potatoes, it is confirmed that the overexpression of *MPK3* can enhance osmosis and salinity tolerances by modulating the antioxidant system and proline biosynthesis (Zhu et al., 2020). In *Arabidopsis*, *WRKY33* leads to enhanced salt tolerance by modulating *CYP94B1* expression (Krishnamurthy et al., 2020). These results indicated that the genes involved in the MAPK signaling pathway are important contributors to high salt tolerance in plants, especially *CAM4*, *MPK3*, and *WRKY33* which are the important candidate genes for future salt-tolerant plants breeding through genetic engineering.

Under salt stress conditions, the compatible osmolytes soluble sugars and proline are critical for adjusting osmotic potential induced by excessive salt stress (Zhu, 2016; Liang et al., 2018). Increasing evidence indicated that the accumulation of soluble sugars and proline is associated with enhanced stress tolerance in plant cells (Mansour and Ali, 2017; Dai et al., 2018; Rady et al., 2019). Proline content, antioxidant activities, and potassium content are influenced in response to abiotic stresses for controlling membrane stability and mitigating the toxicity of sodium (Heidari et al., 2021; Musavizadeh et al., 2021). In our study, a gene encoding P5CS1, a rate-limiting enzyme involved in proline biosynthesis (Turchetto-Zolet et al., 2009), was differentially upregulated to higher levels in R7 roots than in S4. Meanwhile, significantly elevated levels of proline were also found in R7 roots. In addition, the contents of soluble sugars were determined and we found R7 roots exhibited higher levels of soluble sugars than S4. These results suggested that the increased levels in soluble sugars and proline lead to enhanced salt tolerance and that the salt stress-induced *P5CS1* is a crucial candidate for salt-tolerant engineering in *F. velutina*.

Salt stress causes oxidative stress by promoting the generation of ROS in the plant (Zhu, 2016). It is documented that plants can alleviate ROS-caused damages by activating several enzymatic and non-enzymatic pathways involved in the antioxidant system (Choudhury et al., 2017). Previous studies have revealed that the improved ROS scavenging capacity can enhance plant tolerance to several stresses (Gill and Tuteja, 2010; Zhang et al., 2015; Wei et al., 2019). In our study, 12 genes encoding POD were upregulated to higher in R7 compared with S4 under salt stress. Consistently, lower levels of H_2O_2 and fewer cell damages were found in R7 than in S4.

CONCLUSION

This study revealed that *F. velutina* R7 presents higher salt tolerance than S4. The findings reported here allow us to propose a potential mechanism to explain the high salt tolerance of R7 (**Figure 8**). Under salt stress, R7 activates multifaceted defense machines in the roots and leaves to enhance salt tolerance. In the R7 root, salt stress induces the upregulation of *P5CS1* and the key genes involved in starch and sucrose metabolism, therefore increasing osmotic adjustment. In the R7 leaf, salt stress induces the upregulation of important genes involved in plant hormone signaling (especially *PYR1/PYL/RCAR* ABA receptor genes and *PP2C*) and MAPK signaling pathways (especially *CAM4*, *MPK3*, and *WRKY33*). In addition, salt stress induces the upregulation of *POD* genes in both roots and leaves, enhancing ROS scavenging capacity. Collectively, our findings revealed the crucial defense pathways underlying the high salt tolerance of R7, thus providing insights into the salt-tolerant *F. velutina* breeding.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**. The RNA-Seq data presented in this study can be available at the Sequence Read Archive under accession numbers PRJNA791967.

AUTHOR CONTRIBUTIONS

KY and DW conceptualized the research program. JL, XM, and LY finished the analysis of this study and wrote the manuscript. QL, HF, and CW conducted the RNA sequencing data analysis. YD and ZC designed the qRT-PCR experiment and finished the operation. RZ, YB, and WH planted the material and finished the physiology analysis. JL, KY, and DW revised the manuscript. All authors discussed the results, commented on the manuscript, and approved the submitted version.

FUNDING

This study was supported by the Improved Variety Program of Shandong Province of China (2019LZGC009).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.842726/full#supplementary-material

Supplementary Figure 1 | All the samples were hierarchical clustered by calculating correlation coefficient (γ^2) between each sample. From blue to red represented the value of γ^2 from low to high.

Supplementary Figure 2 | Identification of the differentially expressed genes (DEGs) in leaves. **(A)** Heatmap showed the expression patterns of DEGs in

comparisons S4SL vs. S4CL, R7CL vs. S4CL, R7SL vs. S4SL, and R7SL vs. R7CL based on Z-score. The color represented the Z-score. From blue to red represented the Z-score from low to high. **(B)** Venn diagram showed the distribution of upregulated and downregulated DEGs among different comparisons, respectively.

Supplementary Figure 3 | Identification of the differentially expressed genes (DEGs) in roots. **(A)** Heatmap showed the expression patterns of DEGs in comparisons S4SR vs. S4CR, R7CR vs. S4CR, R7SR vs. S4SR, and R7SR vs. R7CR based on Z-score. The color represented the Z-score. From blue to red represented the Z-score from low to high. **(B)** Venn diagram showed the distribution of upregulated and downregulated DEGs among different comparisons, respectively.

Supplementary Figure 4 | The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of downregulated genes in comparison R7SL vs. S4SL.

REFERENCES

Arocho, A., Beiyun, C., Marc, L., and Qiulu, P. (2006). Validation of the 2-DeltaDeltaCt calculation as an alternate method of data analysis for quantitative PCR of BCR-ABL P210 transcripts. *Diagn. Mol. Pathol*. 15, 56–61. doi: 10.1097/00019606-200603000-00009

Ashraf, M. (2009). Biotechnological approach of improving plant salt tolerance using antioxidants as markers. *Biotechnol. Adv*. 27, 84–93. doi: 10.1016/j.biotechadv.2008.09.003

Bari, R., and Jones, J. D. (2009). Role of plant hormones in plant defence responses. *Plant Mol. Biol*. 69, 473–488. doi: 10.1007/s11103-008-9435-0

Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., et al. (2009). ClueGO: a cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* 25, 1091–1093. doi: 10.1093/bioinformatics/btp101

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Cai, S., Chen, G., Wang, Y., Huang, Y., Marchant, D. B., Wang, Y., et al. (2017). Evolutionary conservation of ABA signaling for stomatal closure. *Plant Physiol*. 174, 732–747. doi: 10.1104/pp.16.01848

Catalá, R., López-Cobollo, R., Berbís, M. Á, Jiménez-Barbero, J., and Salinas, J. (2021). Trimethylamine N-oxide is a new plant molecule that promotes abiotic stress tolerance. *Sci. Adv*. 7:eabd9296. doi: 10.1126/sciadv.abd9296

Chen, X., Ding, Y., Yang, Y., Song, C., Wang, B., Yang, S., et al. (2021). Protein kinases in plant responses to drought, salt, and cold stress. *J. Integr. Plant. Biol*. 63, 53–78. doi: 10.1111/JIPB.13061

Chinnusamy, V., Schumaker, K., and Zhu, J. K. (2004). Molecular genetic perspectives on cross-talk and specificity in abiotic stress signalling in plants. *J. Exp. Bot*. 55, 225–236. doi: 10.1093/jxb/erh005

Choudhury, F. K., Rivero, R. M., Blumwald, E., and Mittler, R. (2017). Reactive oxygen species, abiotic stress and stress combination. *Plant J*. 90, 856–867. doi: 10.1111/tpj.13299

Dahro, B., Wang, F., Peng, T., and Liu, J. H. (2016). PtrA/NINV, an alkaline/neutral invertase gene of *Poncirus trifoliata*, confers enhanced tolerance to multiple abiotic stresses by modulating ROS levels and maintaining photosynthetic efficiency. *BMC Plant Biol*. 16:76. doi: 10.1186/s12870-016-0761-0

Dai, W., Wang, M., Gong, X., and Liu, J. H. (2018). The transcription factor FcWRKY40 of *Fortunella crassifolia* functions positively in salt tolerance through modulation of ion homeostasis and proline biosynthesis by directly regulating SOS2 and P5CS1 homologs. *New Phytol*. 219, 972–989. doi: 10.1111/nph.15240

Daudi, A., and O'Brien, J. A. (2012). Detection of hydrogen peroxide by DAB staining in *Arabidopsis* leaves. *Bio Protoc*. 2:e263. doi: 10.21769/BioProtoc.263

De Zelicourt, A., Jean, C., and Heribert, H. (2016). The role of MAPK modules and ABA during abiotic stress signaling. *Trends Plant Sci*. 21, 677–685. doi: 10.1016/j.tplants.2016.04.004

Du, B., Chen, N., Song, L., Wang, D., Cai, H., and Yao, L. (2021). Alfalfa (*Medicago sativa* L.) MsCML46 gene encoding calmodulin-like protein confers tolerance to abiotic stress in tobacco. *Plant Cell Rep*. 40, 1907–1922. doi: 10.1007/S00299-021-02757-7

Egamberdieva, D., Wirth, S., Bellingrath-Kimura, S. D., Mishra, J., and Arora, N. K. (2019). Salt-tolerant plant growth promoting rhizobacteria for enhancing crop productivity of saline soils. *Front Microbiol*. 10, 2791–2791. doi: 10.3389/fmicb.2019.02791

Fang, H., Liu, X., Dong, Y., Feng, S., Zhou, R., and Wang, C. (2021). Transcriptome and proteome analysis of walnut (*Juglans regia* L.) fruit in response to infection by colletotrichum gloeosporioides. *BMC Plant Biol*. 21:249. doi: 10.1186/S12870-021-03042-1

Feng, Y., Liu, J., Zhai, L., Gan, Z., Zhang, G., Yang, S., et al. (2019). Natural variation in cytokinin maintenance improves salt tolerance in apple rootstocks. *Plant Cell Environ*. 42, 424–436. doi: 10.1111/pce.13403

Gabilly, S. T., Baker, C. R., Wakao, S., Crisanto, T., Guan, K., Bi, K., et al. (2019). Regulation of photoprotection gene expression in *Chlamydomonas* by a putative E3 ubiquitin ligase complex and a homolog of CONSTANS. *Proc. Natl. Acad. Sci. U.S.A*. 116, 17556–17562. doi: 10.1073/pnas.1821689116

Gao, H., Huang, H., Lu, K., Wang, C., Liu, X., Song, Z., et al. (2021). OsCYP714D1 improves plant growth and salt tolerance through regulating gibberellin and ion homeostasis in transgenic poplar. *Plant Physiol. Biochem*. 168, 447–456. doi: 10.1016/J.PLAPHY.2021.10.023

Gill, S. S., and Tuteja, N. (2010). Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. *Plant Physiol. Biochem*. 48, 909–930. doi: 10.1016/j.plaphy.2010.08.016

Gucci, R., Lombardini, L., and Tattini, M. (1997). Analysis of leaf water relations in leaves of two olive (Olea europaea) cultivars differing in tolerance to salinity. *Tree Physiol*. 17, 13–21. doi: 10.1093/treephys/17.1.13

He, L., Xu, Y., Zeng, F., Tian, H., Xiao, Y., Liu, H., et al. (2021). Establishment of a micropropagation supporting technology for the *Fraxinus mandshurica* × *Fraxinus sogdiana*. *Vitro Cell Dev. Biol*. 57, 307–318. doi: 10.1007/S11627-021-10157-5

Heidari, P., Entazari, M., Ebrahimi, A., Ahmadizadeh, M., Vannozzi, A., Palumbo, F., et al. (2021). Exogenous EBR ameliorates endogenous hormone contents in tomato species under low-temperature stress. *Horticulturae* 7:84. doi: 10.3390/horticulturae7040084

Huang, G. T., Ma, S. L., Bai, L. P., Zhang, L., Ma, H., Jia, P., et al. (2012). Signal transduction during cold, salt, and drought stresses in plants. *Mol. Biol. Rep*. 39, 969–987. doi: 10.1007/s11033-011-0823-1

Huang, Y., Zhou, J., Li, Y., Quan, R., Wang, J., Huang, R., et al. (2021). Salt stress promotes abscisic acid accumulation to affect cell proliferation and expansion of primary roots in rice. *Int. J. Mol. Sci*. 22:10892. doi: 10.3390/IJMS221910892

Ji, H., Pardo, J. M., Batelli, G., Van Oosten, M. J., Bressan, R. A., Li, X., et al. (2013). The Salt Overly Sensitive (SOS) pathway: established and emerging roles. *Mol. Plant*. 6, 275–286. doi: 10.1093/mp/sst017

Kaleem, F., Shabir, G., Aslam, K., Rasul, S., Manzoor, H., Shah, S. M., et al. (2018). An overview of the genetics of plant response to salt stress: present status and the way forward. *Appl. Biochem. Biotechnol*. 186, 306–334. doi: 10.1007/s12010-018-2738-y

Kelly, L. J., Plumb, W. J., Carey, D. W., Mason, M. E., Cooper, E. D., Crowther, W., et al. (2020). Convergent molecular evolution among ash species resistant to the emerald ash borer. *Nat. Ecol. Evol*. 4, 1116–1128. doi: 10.1038/s41559-020-1209-3

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods*. 12, 357–360. doi: 10.1038/nmeth.3317

Krishnamurthy, P., Vishal, B., Ho, W. J., Lok, F. C. J., Lee, F. S. M., Kumar, P. P., et al. (2020). Regulation of a cytochrome P450 gene CYP94B1 by WRKY33 transcription factor controls apoplastic barrier formation in roots to confer salt tolerance. *Plant Physiol*. 184, 2199–2215. doi: 10.1104/pp.20.01054

Li, C., Cheng, P., Li, Z., Xu, Y., Sun, Y., Qin, D., et al. (2021). Transcriptomic and metabolomic analyses provide insights into the enhancement of torulene and torularhodin production in rhodotorula glutinis ZHK under moderate

salt conditions. *J. Agric. Food Chem*. 69, 11523–11533. doi: 10.1021/ACS.JAFC.1C04028

Li, M., Li, Y., Li, H., and Wu, G. (2012). Improvement of paper mulberry tolerance to abiotic stresses by ectopic expression of tall fescue FaDREB1. *Tree Physiol.* 32, 104–113. doi: 10.1093/treephys/tpr124

Li, T., Peng, Z. Y., Bi, Y. P., and Fan, Z. X. (2013). Molecular cloning and expression analysis of FvMYB1 from *Fraxinus velutina* Torr. *Turk J. Agric. For*. 37, 517–526. doi: 10.3906/tar-1209-4

Li, T., Sun, J., Li, C., Lu, Z., and Xia, J. (2019). Cloning and expression analysis of the FvNCED3 gene and its promoter from ash (*Fraxinus velutina*). *J. For. Res*. 30, 471–482. doi: 10.1007/s11676-018-0632-7

Liang, W., Ma, X., Wan, P., and Liu, L. (2018). Plant salt-tolerance mechanism: a review. *Biochem. Biophys. Res. Commun*. 495, 286–291. doi: 10.1016/j.bbrc.2017.11.043

Lin, Z., Li, Y., Wang, Y., Liu, X., Ma, L., Zhang, Z., et al. (2021). Initiation and amplification of SnRK2 activation in abscisic acid signaling. *Nat. Commun*. 12:2456. doi: 10.1038/S41467-021-22812-X

Liu, X., Zhu, Y., Zhai, H., Cai, H., Ji, W., Luo, X., et al. (2012). AtPP2CG1, a protein phosphatase 2C, positively regulates salt tolerance of *Arabidopsis* in abscisic acid-dependent manner. *Biochem. Biophs. Res. Commun*. 422, 710–715. doi: 10.1016/j.bbrc.2012.05.064

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8

Ma, L., Zhang, M., Chen, J., Qing, C., He, S., Zou, C., et al. (2021). GWAS and WGCNA uncover hub genes controlling salt tolerance in maize (*Zea mays L.*) seedlings. *Theor. Appl. Genet*. 134, 3305–3318. doi: 10.1007/S00122-021-03897-W

Mansour, M. M. F., and Ali, E. F. (2017). Evaluation of proline functions in saline conditions. *Phytochemistry* 140, 52–68. doi: 10.1016/j.phytochem.2017.04.016

Mao, P., Tang, Q., Cao, B., Liu, J., Shao, H., Cai, Z., et al. (2017). Eco-physiological adaptability in mixtures of *robinia pseudoacacia* and *Fraxinus velutina* and coastal eco-engineering. *Ecol. Eng.* 106, 109–115. doi: 10.1016/j.ecoleng.2017.05.021

Musavizadeh, Z., Najafi-Zarrini, H., Kazemitabar, S. K., Hashemi, S. H., Faraji, S., Barcaccia, G., et al. (2021). Genome-wide analysis of potassium channel genes in rice: expression of the *OsAKT* and *OsKAT* genes under salt stress. *Genes* 12:784. doi: 10.3390/genes12050784

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., Salzberg, S. L., et al. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295. doi: 10.1038/nbt.3122

Pogány, M., Von Rad, U., Grün, S., Dongó, A., Pintye, A., Simoneau, P., et al. (2009). Dual roles of reactive oxygen species and NADPH oxidase RBOHD in an *Arabidopsis*-alternaria pathosystem. *Plant Physiol*. 151, 1459–1475. doi: 10.1104/pp.109.141994

Rady, M. M., Elrys, A. S., Abo El-Maati, M. F., and Desoky, E. M. (2019). Interplaying roles of silicon and proline effectively improve salt and cadmium stress tolerance in *Phaseolus vulgaris* plant. *Plant Physiol. Biochem*. 139, 558–568. doi: 10.1016/j.plaphy.2019.04.025

Ran, X., Wang, X., Gao, X., Liang, H., Liu, B., and Huang, X. (2021). Effects of salt stress on the photosynthetic physiology and mineral ion absorption and distribution in white willow (*Salix alba* L.). *PLoS One* 16:e0260086. doi: 10.1371/journal.pone.0260086

Saini, S., Kaur, N., Marothia, D., Singh, B., Singh, V., Gantet, P., et al. (2021). Morphological analysis, protein profiling and expression analysis of auxin homeostasis genes of roots of two contrasting cultivars of rice provide inputs on mechanisms involved in rice adaptation towards salinity stress. *Plants (Basel)* 10:1544. doi: 10.3390/PLANTS10081544

Seki, M., Ayako, K., Kazuko, Y. S., and Kazuo, S. (2003). Molecular responses to drought, salinity and frost: common and different paths for plant protection. *Curr. Opin. Biotechnol*. 14, 194–199. doi: 10.1016/S0958-1669(03)00030-2

Smethurst, C. F., Warwick, M. G., and Sergey, S. (2009). Using excised leaves to screen lucerne for salt tolerance: physiological and cytological evidence. *Plant Signal. Behav*. 4, 39–41. doi: 10.4161/psb.4.1.7269

Soma, F., Mogami, J., Yoshida, T., Abekura, M., Takahashi, F., Kidokoro, S., et al. (2017). ABA-unresponsive SnRK2 protein kinases regulate mRNA decay under osmotic stress in plants. *Nat. Plants* 3:16204. doi: 10.1038/nplants.2016.204

Tian, F., Yang, D. C., Meng, Y. Q., Jin, J., and Gao, G. (2020). PlantRegMap: charting functional regulatory maps in plants. *Nucleic Acids Res*. 48, D1104–D1113. doi: 10.1093/nar/gkz1020

Tsabarducas, V., Chatzistathis, T., Therios, I., Koukourikou-Petridou, M., and Tananaki, C. (2015). Differential tolerance of 3 self-rooted Citrus limon cultivars to NaCl stress. *Plant Physiol. Biochem*. 97, 196–206. doi: 10.1016/j.plaphy.2015.10.007

Turchetto-Zolet, A. C., Margis-Pinheiro, M., and Margis, R. (2009). The evolution of pyrroline-5-carboxylate synthase in plants: a key enzyme in proline synthesis. *Mol. Genet. Genomics* 281, 87–97. doi: 10.1007/s00438-008-0396-4

Umezawa, T., Sugiyama, N., Mizoguchi, M., Hayashi, S., Myouga, F., Yamaguchi-Shinozaki, K., et al. (2009). Type 2C protein phosphatases directly regulate abscisic acid-activated protein kinases in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A*. 106, 17588–17593. doi: 10.1073/pnas.0907095106

Van Zelm, E., Zhang, Y. X., and Testerink, C. (2020). Salt tolerance mechanisms of plants. *Annu. Rev. Plant Biol*. 71, 403–433. doi: 10.1146/annurev-arplant-050718-100005

Wang, Z., Wang, M., Liu, L., and Meng, F. (2013). Physiological and proteomic responses of diploid and tetraploid black locust (*Robinia pseudoacacia* L.) subjected to salt stress. *Int. J. Mol. Sci*. 14, 20299–20325. doi: 10.3390/ijms141020299

Wani, S. H., Kumar, V., Khare, T., Guddimalli, R., Parveda, M., Solymosi, K., et al. (2020). Engineering salinity tolerance in plants: progress and prospects. *Planta* 251:76. doi: 10.1007/s00425-020-03366-6

Wei, T., Wang, Y., Xie, Z., Guo, D., Chen, C., Fan, Q., et al. (2019). Enhanced ROS scavenging and sugar accumulation contribute to drought tolerance of naturally occurring autotetraploids in *Poncirus trifoliata*. *Plant. Biotechnol. J.* 17, 1394–1407. doi: 10.1111/pbi.13064

Xia, X. J., Zhou, Y. H., Shi, K., Zhou, J., Foyer, C. H., Yu, J. Q., et al. (2015). Interplay between reactive oxygen species and hormones in the control of plant development and stress tolerance. *J. Exp. Bot*. 66, 2839–2856. doi: 10.1093/jxb/erv089

Yan, L., Liu, C., Wang, Y., Wang, K., Ren, F., Yao, J., et al. (2019). De novo transcriptome analysis of *Fraxinus velutina* Torr in response to NaCl stress. *Tree Genet. Genomes* 15:56. doi: 10.1007/s11295-019-1340-y

Yang, Y., and Guo, Y. (2018a). Elucidating the molecular mechanisms mediating plant salt-stress responses. *New Phytol.* 217, 523–539. doi: 10.1111/nph.14920

Yang, Y., and Guo, Y. (2018b). Unraveling salt stress signaling in plants. *J. Integr. Plant Biol*. 60, 796–804. doi: 10.1111/jipb.12689

Yildiz, O., Altundag, E., Cetin, B., Guner, S. T., Sarginci, M., Toprak, B., et al. (2018). Experimental arid land afforestation in central anatolia, turkey. *Environ. Monit. Assess*. 190:355. doi: 10.1007/s10661-018-6724-1

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). ClusterProfiler: an R package for comparing biological themes among gene clusters. *Omics* 16, 284–287. doi: 10.1089/omi.2011.0118

Yu, Z., Duan, X., Luo, L., Dai, S., Ding, Z., Xia, G., et al. (2020). How plant hormones mediate salt stress responses. *Trends Plant Sci*. 25, 1117–1130. doi: 10.1016/j.tplants.2020.06.008

Zhang, K., Cui, H., Li, M., Xu, Y., Cao, S., Long, R., et al. (2020). Comparative time-course transcriptome analysis in contrasting *Carex rigescens* genotypes in response to high environmental salinity. *Ecotoxicol. Environ. Saf*. 194:110435. doi: 10.1016/j.ecoenv.2020.110435

Zhang, Q., Wang, M., Hu, J., Wang, W., Fu, X., and Liu, J. H. (2015). PtrABF of *Poncirus trifoliata* functions in dehydration tolerance by reducing stomatal density and maintaining reactive oxygen species homeostasis. *J. Exp. Bot*. 66, 5911–5927. doi: 10.1093/jxb/erv301

Zheng, J., Ma, X., Zhang, X., Hu, Q., and Qian, R. (2018). Salicylic acid promotes plant growth and salt-related gene expression in *Dianthus superbus* L. (Caryophyllaceae) grown under different salt stress conditions. *Physiol. Mol. Biol. Plants* 24, 231–238. doi: 10.1007/s12298-017-0496-x

Zhu, H., He, M., Jahan, M. S., Wu, J., Gu, Q., Shu, S., et al. (2021). CsCDPK6, a CsSAMS1-Interacting protein, affects polyamine/ethylene biosynthesis in cucumber and enhances salt tolerance by overexpression in tobacco. *Int. J. Mol. Sci*. 22:11133. doi: 10.3390/IJMS222011133

Zhu, J. K. (2002). Salt and drought stress signal transduction in plants. *Annu. Rev. Plant Biol*. 53, 247–273. doi: 10.1146/annurev.arplant.53.091401.143329

Zhu, J. K. (2016). Abiotic stress signaling and responses in plants. *Cell* 167, 313–324. doi: 10.1016/j.cell.2016.08.029

Zhu, X., Zhang, N., Liu, X., Wang, S., Li, S., Yang, J., et al. (2020). StMAPK3 controls oxidase activity, photosynthesis and stomatal aperture under salinity and osmosis stress in potato. *Plant Physiol. Biochem*. 156, 167–177. doi: 10.1016/j.plaphy.2020.09.012

Zhu, Y., Zhu, Y., Yin, J., Liang, Y., Liu, J., Jia, J., et al. (2019). Transcriptomic dynamics provide an insight into the mechanism for silicon-mediated alleviation of salt stress in cucumber plants. *Ecotoxicol. Environ. Saf*. 174, 245–254. doi: 10.1016/j.ecoenv.2019.02.075

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 21 March 2022
doi: 10.3389/fpls.2022.831839

Reprogramming of the *Hevea brasiliensis* Epigenome and Transcriptome in Response to Cold Stress

Xiao Tang[1,2†], Yonglei Zhang[1†], Hong-Mei Yuan[1], Jinling Zhai[1] and Xi Huang[1*]*

[1]*Hainan Yazhou Bay Seed Laboratory, Sanya Nanfan Institute of Hainan University, Sanya, China,* [2]*Hunan Rice Research Institute, Hunan Academy of Agricultural Sciences, Changsha, China*

OPEN ACCESS

Edited by:
Mukesh Jain, Jawaharlal Nehru University, India

Reviewed by:
Ming-Bo Wang, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia
Daisuke Miki, Shanghai Center for Plant Stress Biology, Shanghai Institute for Biological Sciences (CAS), China

****Correspondence:***
Jinling Zhai
jinlingzhai@126.com
Xi Huang
xihuang@hainanu.edu.cn

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

Received: *09 December 2021*
Accepted: *08 February 2022*
Published: *21 March 2022*

Citation:
Tang X, Zhang Y, Yuan H-M, Zhai J and Huang X (2022) Reprogramming of the Hevea brasiliensis Epigenome and Transcriptome in Response to Cold Stress. Front. Plant Sci. 13:831839. doi: 10.3389/fpls.2022.831839

Low temperature is a key factor limiting the rubber plantation extending to high latitude area. Previous work has shown that cold-induced DNA demethylation was coordinated with the expression of cold-responsive (*COR*) genes in *Hevea brasiliensis*. In this work, reduced representation bisulphite sequencing analysis of *H. brasiliensis* showed that cold treatment induced global genomic DNA demethylation and altered the sequence contexts of methylated cytosines, but the levels of mCG methylation in transposable elements were slightly enhanced by cold treatment. Integrated analysis of the DNA methylome and transcriptome revealed 400 genes whose expression correlated with altered DNA methylation. DNA demethylation in the upstream region of gene seems to correlate with higher gene expression, whereas demethylation in the gene body has less association. Our results suggest that cold treatment globally change the genomic DNA methylation status of the rubber tree, which might coordinate reprogramming of the transcriptome.

Keywords: *Hevea brasiliensis*, cold, DNA demethylation, transposon, RRBS

INTRODUCTION

Low temperature is a major environmental stress that seriously affects plant development and geographic distribution. Depending on the extent of the cold sensitivity of plants, cold stress can be subdivided into chilling (0–15°C) and freezing (<0°C; Wani et al., 2013). Chilling disrupts plant growth physiology by inducing photosynthesis-associated damage, chlorosis, apoptosis, and loss of membrane fluidity. Freezing causes ice formation and protoplast dehydration that directly kill cells. Plants exhibit increased freezing tolerance upon long-term exposure to non-freezing temperatures, a process known as cold acclimation (Liu et al., 2017a). Cold tolerance relies on many changes in plant processes, ranging from gene expression to physiological, biochemical, and metabolic processes (Knight and Knight, 2012; Theocharis et al., 2012). Cold acclimation activates rapid expression of the C-repeat binding factor (CBF) transcription factors and induces CBF-targeted genes, such as cold-responsive (*COR*) genes, which contribute to enhanced freezing tolerance (Jaglo-Ottosen et al., 1998; Fowler and Thomashow, 2002). Cold acclimation also increases the cytoplasmic volume and levels of organic acids (alpha-ketoglutarate, citrate, fumarate, and malate; Strand et al., 1999). During cold acclimation, the suppression of photosynthesis and photosynthesis gene expression is removed in leaves and is accompanied

by increase in the activities of enzymes in the Calvin cycle and in the sucrose biosynthesis pathway (Strand et al., 1999).

Epigenetic regulation *via* mechanisms including DNA methylation and histone modification is involved in plant stress responses (Chinnusamy and Zhu, 2009a). Cytosine methylation is an important epigenetic mechanism that regulates gene expression and transposable elements (TEs) silencing and safeguards genome stability (Law and Jacobsen, 2010). In plant genomes, cytosine can be methylated in three sequence contexts: symmetric contexts CG and CHG (H = A, T, or C), or asymmetric context CHH (Law and Jacobsen, 2010; Vanyushin and Ashapkin, 2011). Both symmetric- and asymmetric-methylations are associated with repressive chromatin in gene promoters and with repression of gene transcription. Whole-genome DNA methylome analyses revealed that heavy cytosine methylation occurs in repetitive sequences and TEs (Penterman et al., 2007; Cokus et al., 2008). Methyltransferases Domains Rearranged Methylase 1 (DRM1) and DRM2 catalyze *de novo* cytosine methylation through the RNA-directed DNA methylation (RdDM) pathway, while the maintenance of symmetric CG and CHG methylation is mediated by the DNMT1-like enzyme MET1 and the plant-specific enzyme Chromomethylase 3 (CMT3), respectively (Cao and Jacobsen, 2002; Henderson and Jacobsen, 2007).

Plant cold response involves in epigenetic regulation. Vernalisation is a well-known example of epigenetic regulation in the plant cold response (Gendall et al., 2001; Bastow et al., 2004; Sung and Amasino, 2004). For other example, cold induced expression of *ZmMI1* was correlated with a reduction in methylation in the DNA of the nucleosome core in maize roots (Steward et al., 2002). In tobacco, cold stresses induced-DNA demethylation in the coding sequence of a glycerophosphodiesterase-like protein (*NtGPDL*) gene correlated with *NtGDPL* gene expression (Choi and Sano, 2007). However, cold induced epigenetic regulation is still relatively undocumented.

Hevea brasiliensis Muell. Arg. is an economically important tree originating from the tropical Amazon rain forest. When the locations of rubber plantations are extended to high latitude areas, such as southern China and northern Vietnam, cold stress becomes a key factor limiting rubber production. Low temperature prevents the tapping of trees for 1–3 months per year in the suboptimal areas (Rao et al., 1998; Jacob et al., 1999). Much attention has been paid to breeding cold-tolerant cultivars and studying the chilling physiology of *H. brasiliensis* (Rao et al., 1998; Mai et al., 2009, 2010), but the cold response mechanisms of *H. brasiliensis* have not been well elucidated. In a previous report, we showed that cold treatment altered the expression patterns of DNA methylation-associated genes, such as *HbMET*, *HbCMT*, and *HbDRM*. Long-term cold treatment induced DNA hypomethylation in the promoters of *HbICE1*, *HbCBF2*, and *HbMET*. Under natural conditions, the *HbICE1* and *HbMET* promoters switch from hypomethylation to hypermethylation status as the seasons change from winter to summer (Tang et al., 2018). *COR* gene expression levels were correlated with the status of DNA demethylation but not with the genetic background; we therefore proposed that epigenetic modification was closely related to the cold response of *H. brasiliensis* (Tang et al., 2018). However, the key question of how genome-wide DNA methylation in *H. brasiliensis* is affected by cold stress conditions remains open. Understanding this response might help us to better understand the mechanisms of cold acclimation in *H. brasiliensis*. In this study, we use reduced representation bisulphite sequencing (RRBS) to examine the alteration of the epigenome in response to cold stress. RRBS is a cost-effective method for rapid and affordable genome-wide DNA methylation analysis (Gu et al., 2011; Tanas et al., 2017). Additionally, we used digital gene expression (DGE) analysis based on RNA sequencing (RNA-seq) to analyze the expression of *COR* genes. Combined RRBS and DGE analyses allowed us to identify the causal link between genome-wide changes in DNA methylation and gene expression.

Abbreviations: CBF, C-repeat binding factor; COR, Cold-responsive; DEGs, Differentially expressed genes; DGE, Digital gene expression; DMCs, Differentially methylated cytosines; DMP, Differentially methylated position; DMR, Differentially methylated region; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; RRBS, Reduced representation bisulphite sequencing; TEs, Transposable elements; UTR, Untranslated region.

MATERIALS AND METHODS

Plant Materials and Treatments

To minimize the variation of the genomic background, tissue-cultured self-rooting juvenile clones (JCs) from the same donor clones Reyan 7–33-97 were selected in the study (Li et al., 2016). Reyan 7–33-97 is an elite cultivar widely planted in China and its genome has been sequenced and fine assembled (Tang et al., 2016). The 60-day old JCs saplings (after transferring from sterile tube to polybag) were transferred to climate chamber for cold treatment as reported previously (Tang et al., 2018). Twenty saplings were treated under 19°C with the following growth conditions: 16-h light (100 lux)/8-h dark, 75% relative humidity for 1 month. Twenty control saplings were cultivated under 28°C with the same lighting, photoperiod, and relative humidity conditions. The cold-treated saplings were transferred to 28°C for a 1-month recovery and then further transferred to 19°C for the second and third treatments. A one-month interval of recovery between two treatments could prevent the saplings from leaf falling. After the third treatment, the treated and control leaf samples were collected, frozen immediately, ground in liquid nitrogen, and stored at −80°C until RNA and DNA extraction, respectively (Xia et al., 2011). Three independently treated leaf samples (treated or control) were mixed for pooling sample for RNA-seq and bisulphite sequencing.

Reduced Representation Bisulphite Sequencing

After genomic DNAs were extracted from the pooling leaf samples, DNA concentration and integrity were detected by NanoPhotometer® spectrophotometer (IMPLEN, CA, United States) and agarose gel electrophoresis, respectively, (Tang et al., 2018). The qualified DNA was digested by restriction

endonucleases (*Msp*I) and repaired by 3′-end addition and adaptor ligation. The 40–220 bp fragments were selected for bisulfite conversion with Methylation-Gold kit (ZYMO, CA, United States). The converted DNA was PCR amplified and measured using an Agilent 2,100 bioanalyzer instrument, then the RRBS library was constructed and subjected the Illumina X10 sequencing platform with paired-end 150 bp sequencing by Gene Denovo Biotechnology Co. (Guangzhou, China). To get high quality clean reads, raw reads were filtered according to the following rules: (1) removing reads containing more than 10% of unknown nucleotides (*N*); (2) removing low quality reads containing more than 40% of low quality (Q-value ≤ 20) bases. The obtained clean reads were mapped to the Clone 7–33-97 reference genome using BSMAP software (Xi and Li, 2009). Then a custom Perl script was used to call methylated cytosines and the methylated cytosines were tested with the correction algorithm described (Lister et al., 2009). The methylation level was calculated based on methylated cytosine percentage in the whole genome, in each scaffold and in different regions for each sequence context (CG, CHG, and CHH). To assess different methylation patterns in different genomic regions, the methylation profile at flanking 2 kb regions and genebody (or TEs) was plotted based on the average methylation levels for each window.

Differentially Methylated Regions Analysis

Differentially methylated regions for each sequence context (CG, CHG, and CHH) between two samples were identified according to the following criteria: (1) more than five methylated cytosines in at least one sample; (2) more than 10 reads coverage for each cytosine, and more than four reads for each methylated cytosine; (3) region length is between 40 bp and 10 kb; (4) the distance between adjacent methylated sites <200 bp; (5) the fold change of the average methylation level > 2; (6) Pearson's chi-square test (χ^2) value $p \leq 0.05$. The putative DMRs overlapping at adjacent 2 kb (upstream or downstream) or body regions of genes or TEs were sorted out for further study (Akalin et al., 2012).

Functional Enrichment Analysis of DMC/DMR-Related Genes

To analyze functional enrichment of genes affected by DMRs, Gene Ontology (GO) enrichment analysis and KEGG pathway enrichment analysis were conducted for DMR-related genes. GO enrichment analysis provides all GO terms that significantly enriched in genes comparing to the genome background, and filter the genes that correspond to biological functions. Firstly, all genes were mapped to GO terms in the Gene Ontology database.[1] KEGG is the major public pathway-related database.[2]

RNA Sequencing

After total RNA was extracted, mRNA was enriched by Oligo (dT) beads. Then the enriched mRNA was fragmented into short fragments and reverse transcripted into cDNA with random primers. Second-strand cDNA were synthesized, purified with QiaQuick PCR extraction kit, and then ligated to Illumina sequencing adapters. The ligation products were size selected by agarose gel electrophoresis, PCR amplified, and sequenced using Illumina HiSeqTM 2,500 by Gene Denovo Biotechnology Co. (Guangzhou, China). High quality clean reads were further filtered according to the following rules: (1) removing reads containing adapters; (2) removing reads containing more than 10% of unknown nucleotides (*N*); and (3) removing low quality reads containing more than 50% of low quality (Q-value ≤ 20) bases.

The clean reads were then mapped to reference genome of Clone 7–33-97 by TopHat2 (version 2.0.3.12; Kim et al., 2013; Tang et al., 2016). The alignment parameters were as follows: (1) maximum read mismatch is 2; (2) the distance between mate-pair reads is 50 bp; and (3) the error of distance between mate-pair reads is ±80 bp. After aligned with reference genome, unmapped reads (or mapped very poorly) were then re-aligned with Bowtie2, the enriched unmapped reads were split into smaller segments which were then used to find potential splice sites. Splice sites were built with initial unmapped reads by TopHat2 without relying on the known genes annotation (Trapnell et al., 2010). The reconstruction of transcripts was carried out with software Cufflinks (Trapnell et al., 2012), which together with TopHat2, allow to identify new genes and new splice variants of known reference genes. EdgeR software was used to normalize the data and extract digital gene expression (DGE) data (Robinson et al., 2010). Differentially expressed genes (DEGs) were selected when the false discovery rate (FDR) was <0.05. DGE data were used for clustering of DMRs. Gene Ontology (GO) term and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways analyses were visualized with Pathview software (Young et al., 2010; Wang et al., 2015).

Integrated Analysis of DNA Methylation and the Transcriptome

After differentially Methylated Regions (DMRs) and DEG analysis, the distribution of Spearman rho statistics between DNA methylation and gene expression was calculated including either all pairs of genes/methylation positions or only pairs formed by a DMP and a DEG. All DNA methylation and transcriptomics data analyzed were output in Bam format.[3] A list of genes and transposons and the corresponding DNA methylation positions was obtained, and a correlation coefficient was calculated for each pair. The average DNA methylation levels of the genes and TEs in the control and cold-treated plants were displayed in a Bam format file. All data were visualized and analyzed using box plots, heat maps, and circos plots.

Validation of Differential Gene Expression by Quantitative RT-PCR

Quantitative RT-PCR (qRT-PCR) was used to confirm differential expression of the unigenes obtained from sequencing and to further analyze the reliability of the RNA-seq data generated

[1]http://www.geneontology.org/
[2]http://www.kegg.jp/kegg/
[3]http://samtools.sourceforge.net/SAM1.pdf

in this study. Primers used in this work are listed in **Supplementary Table S6**. The *HbACT7b* gene was used as an internal control. First-strand cDNA was synthesized using a cDNA synthesis kit according to the manufacturer's instructions (Fermentas, Vilnius, Lithuania). qRT-PCR was performed according to Li's method (Li et al., 2014). All relative expression data were analyzed using GraphPad Prism 7 software. Each biological sample was represented by three independent replicates. qRT-PCR conditions were as follows: denaturation at 95°C for 30 s followed by 50 cycles for 10 s at 94°C, 30 s at 60°C, and 15 s at 72°C for amplification. Data obtained from qRT-PCR analysis were clustered in accordance with the instructions provided by Stratagene (Santa Clara, CA, United States). The expression analysis was performed from three individual reactions.

RESULTS

Cold Treatment Induced Global DNA Demethylation in *H. brasiliensis*

After cold treatment, genomic DNA was isolated from leaf samples for RRBS. The sequencing statistic was shown in **Table 1** and **Supplementary Table S1**. There were 83,531,456 and 82,345,180 clean reads generated by RRBS for control and cold treated samples, respectively (**Supplementary Table S1**). After filtering, 7,900 MB of clean data and 82,143,730 clean reads were generated from the control samples, 91.24% of which were uniquely mapped to genomic sequences of *H. brasiliensis* cultivar Reyan7-33-97 (NCBI SRA data: PRJNA741882). In total, 7,766 MB of clean data and 81,565,518 clean reads were generated from the cold-treated samples, 94.80% of which were uniquely mapped (**Table 1** and **Supplementary Table S1**).

Sequencing depth analysis showed that the coverage level in genomic functional elements ranged from 21 to 51%, in which the coverage rate of C, CG, CHG, and CHH were indicated, respectively (**Supplementary Table S3**). RRBS results also showed that the numbers of the mCG、mCHG and mCHH in the control samples were 1,610,060, 2,579,031, and 2,582,602, respectively (23.78, 38.08, and 38.14%, respectively). By contrast, the methylation numbers of the cold-treated samples were 632,899, 939,769, and 798,268, respectively (26.69, 39.64, and 33.67%, respectively). These data indicated that cold treatment strongly decreased the methylation number of mCG, mCHG, and mCHH. The total number of three types methylated cytosines decreased from 6,771,693 (control) to 2,370,936 (cold), suggesting that average DNA methylation level were down regulated by cold treatment (**Table 1** and **Supplementary Table S4**).

Further analysis of the distribution of methylated cytosines in various gene regions revealed that the level of mCG methylation was much higher than those of the other two types of Methylated cytosines. Lower methylation was observed near the transcription start site. The levels of mCG methylation located in the gene body, 2-kb upstream region, and 2-kb downstream region were evidently decreased in the cold-treated samples. mCHG methylation was lower than mCG methylation but higher than mCHH methylation. Cold treatment slightly induced demethylation of mCHG and mCHH in the 2-kb upstream and downstream regions (**Figure 1A**). When gene regions were further subdivided into the gene body, 2-kb upstream region, 2-kb downstream region, exon, intron, coding sequence (CDS), 5′-untranslated region (UTR), and 3′-UTR, all methylation levels distributed among the mCG, mCHG, and mCHH were lower in the cold-treated samples than in the control samples (**Figure 1B**), suggesting that cold treatment induced global DNA demethylation in *H. brasiliensis*.

Statistics showed that cold treatment induced significant alteration of the sequence context upstream of methylated cytosines (e.g., mCG, mCHG, and mCHH) and the sequence context downstream of mCG, whereas no difference was found in the sequence context of unmethylated cytosines after cold treatment (**Figure 1C**). These results suggest that cold treatment reprogrammed the *H. brasiliensis* epigenome.

Cold Treatment Enhanced mCG Methylation in TEs

To reveal the differential methylation patterns between functional genes and TEs, which were annotated in *Hevea brasilensis* genomic sequences (Tang et al., 2016), statuses of mCG, or mCHG, or mCHH between control and cold treatment were compared, respectively. The distribution of methylation in genes exhibited a "W" shape. Higher methylation levels showed in the gene body, 2-kb upstream region, and 2-kb downstream region but markedly lower level at the transcription start site and transcription termination site (**Figure 2**, left panel). By contrast, the distribution of methylation in TEs exhibited a "single-hump camel" profile. The DNA methylation level changed dramatically at the TE boundaries, whereas both the upstream and downstream regions had lower methylation levels (**Figure 2**, right panel). Similar phenomenon was also observed in *Brassica rapa* (Chen et al., 2015).

DNA methylation plays an important role in modulating TE silencing. The DNA methylation level is much higher in TEs than in genic regions. Cold treatment reduced average

TABLE 1 | Reduced representation bisulphite sequencing (RRBS) statistics.

Sample	Clean reads/clean data (bp)	GC content (%)	Mapping rate (%)	mCG no./ proportion	mCHG no./ proportion	mCHH no./ proportion	Total methylated
Control	82,143,730/7,900,525,750	24.58%	91.24%	1,610,060/23.78%	2,579,031/38.08%	2,582,602/38.14%	6,671,693
Cold-treated	81,565,518/7,766,520,664	24.56%	94.80%	632,899/26.69%	939,769/39.64%	798,268/33.67%	2,370,936

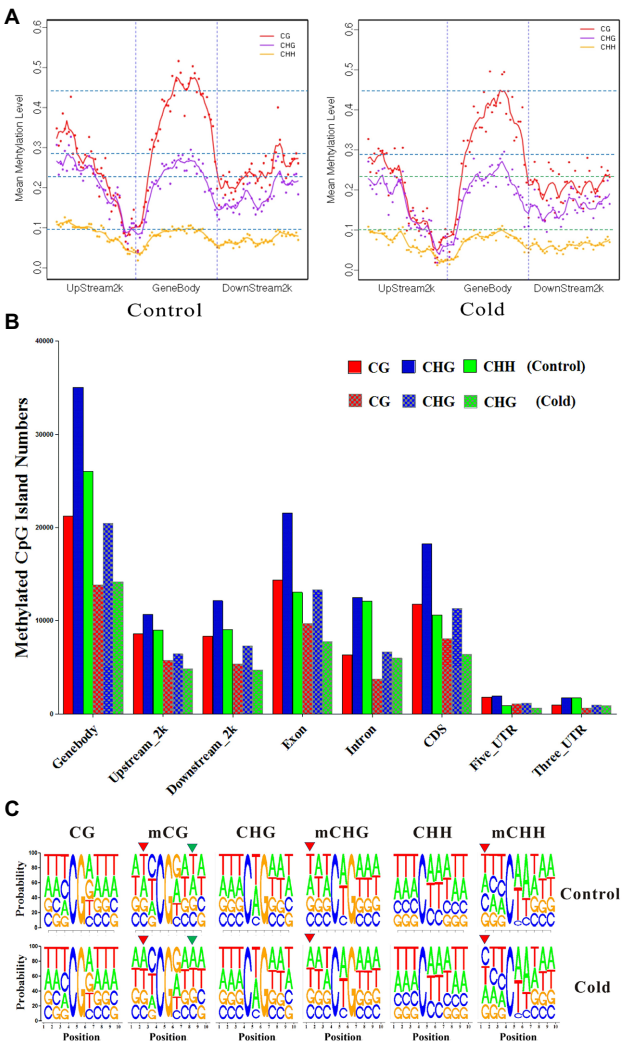


FIGURE 1 | Distribution of methylated cytosines in genic regions and sequence preference after cold treatment. **(A)** Methylation levels of CG, CHG, and CHH distributed in the gene body, upstream region, and downstream region. Upstream, 2-kb region upstream of the transcription start site; gene body; downstream, the 2-kb region downstream of the transcription termination site. **(B)** Distribution of mCG, mCHG and mCHH in the gene body, 2-kb upstream region, 2-kb downstream region, exon, intron, CDS, 5-UTR, and 3-UTR. Solid column and squares column histograms represent the control and cold-treated samples, respectively. The data were normalized by the mapping rate and mapping reads. **(C)** Logos of sequence contexts that are preferentially unmethylated/methylated CG, CHG, and CHH in control and cold-treated samples in which the cytosine is in the fifth position. The logo graphically displays the sequence enrichment at a particular position in the alignment of 10-mers in each class.

DNA methylation levels in mCG, mCHG, and mCHH in genic regions, including the gene body, 2-kb upstream region, and 2-kb downstream region (**Figure 2**, left panel). Interestingly, cold treatment slightly increased the mCG methylation levels in gene body of TEs, although mCHG and mCHH methylation levels were significantly decreased by cold treatment (**Figure 2**, right panel). These differential methylation patterns might reflect the distinct roles of genes and TEs in the *H. brasiliensis* mechanisms for coping with cold stress.

Digital Gene Expression Analysis of Cold-Treated *H. brasiliensis*

To gain a global view of the gene expression patterns induced by the cold treatment of *H. brasiliensis*, we used RNA-seq to identify differentially expressed genes. In total, 3,764,205,386 and 3,433,710,574 HQ (high quality) clean data were acquired from the control and cold-treated samples, respectively (**Supplementary Table S2**, NCBI SRA data: PRJNA741882). Genes coverage analysis showed that the RNA-seq for control sample covers 70.55% of Reyan7-33-97 reference genes at 80–100% coverage rate and 14.85% genes at 60–80% coverage rate (**Supplementary Figure S1A**); RNA-seq for cold treated sample has similar coverage rate (**Supplementary Figure S1B**), suggesting that RNA-seq data have relative higher coverage. Statistics of RNA-Seq reads showed that 29131and 30,242 genes were mapped to the genes of Reyan7-33-97 for control and cold treated samples, respectively (**Supplementary Table S5**), of which 5,736 genes showed differential expression (low FDR < 0.001 and p-value <0.05), including 3,693 upregulated genes and 2043 downregulated genes (**Supplementary Figure S1C**). Scatter plots indicated that most data points were distributed near the diagonal, suggesting that the expression levels of the majority of genes were largely unaffected by cold treatment (**Supplementary Figure S1D**).

To identify pathways in which the cold-induced differentially expressed genes (DEGs) might be involved, the DEGs were categorized based on KEGG pathways and GO term enrichment. The cold-induced DEGs were enriched in the KEGG pathways of plant hormone signal transduction, phenylpropanoid biosynthesis, and amino acid biosynthesis (**Figure 3**). The DEGs were enriched in 47 GO terms, among which "catalytic activity," "metabolic process," "cellular process," "single-organism process," and "binding" were the most enriched (**Supplementary Figure S2**).

Integrated Analysis of DEGs and Changes in DNA Methylation

To establish a correlation between RRBS and DEGs induced by cold treatment, we engineered the mutual information-based analysis using count data from the DGE analysis and overlaid the DNA methylation data by summarizing the methylation changes of the linked DMRs and DMPs. Additionally, to better understand the distribution of methylation levels among the cold-induced DEGs in DMRs, the methylation levels of the DEGs were visualized using a heat map. First, the methylation levels of mCG, mCHG, and mCHH of DEGs in control and cold-treated samples were shown by a heat map. Highest methylation level appeared in mCG, in which demethylation induced by cold could be observed in upstream, genebody and downstream regions. Although methylation level of mCHG and mCHH was much lower than mCG, cold treatment also evidently induced demethylation (**Figure 4A**). Cluster analysis using heat map to represent the methylation levels of DEGs reveals different patterns of methylation modification after cold

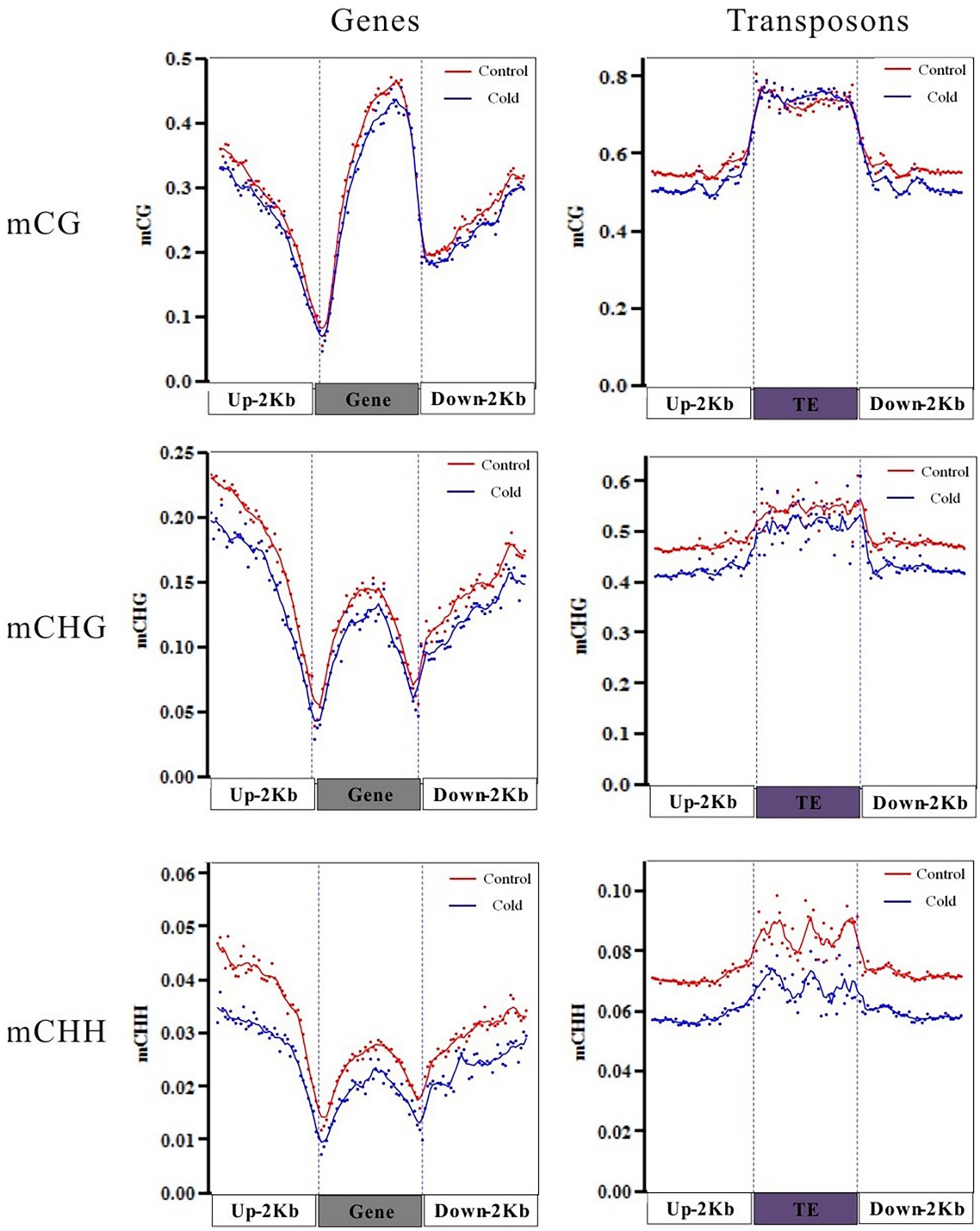


FIGURE 2 | Methylation patterns of genic and transposon element (TE) regions after cold treatment. Levels of mC, mCG, mCHG, and mCHH methylation distributed in the gene body, upstream region, and downstream region of genes and TEs. 2-kb region upstream of the transcription start site; gene body; downstream, the 2-kb region downstream of the transcription termination site.

treatment. The more genes showed demethylation under cold stress, some of which at upstream regions (marked with red rings) and some of which at genebody (marked with blue rings). Interestingly, some cluster of genes exhibited increased methylation level, e.g., at downstream regions (marked with green rings; **Figure 4B**). These results suggest that cold treatment induced different modifications of DNA methylation in different genes regions of rubber tree. To better visualize the combined

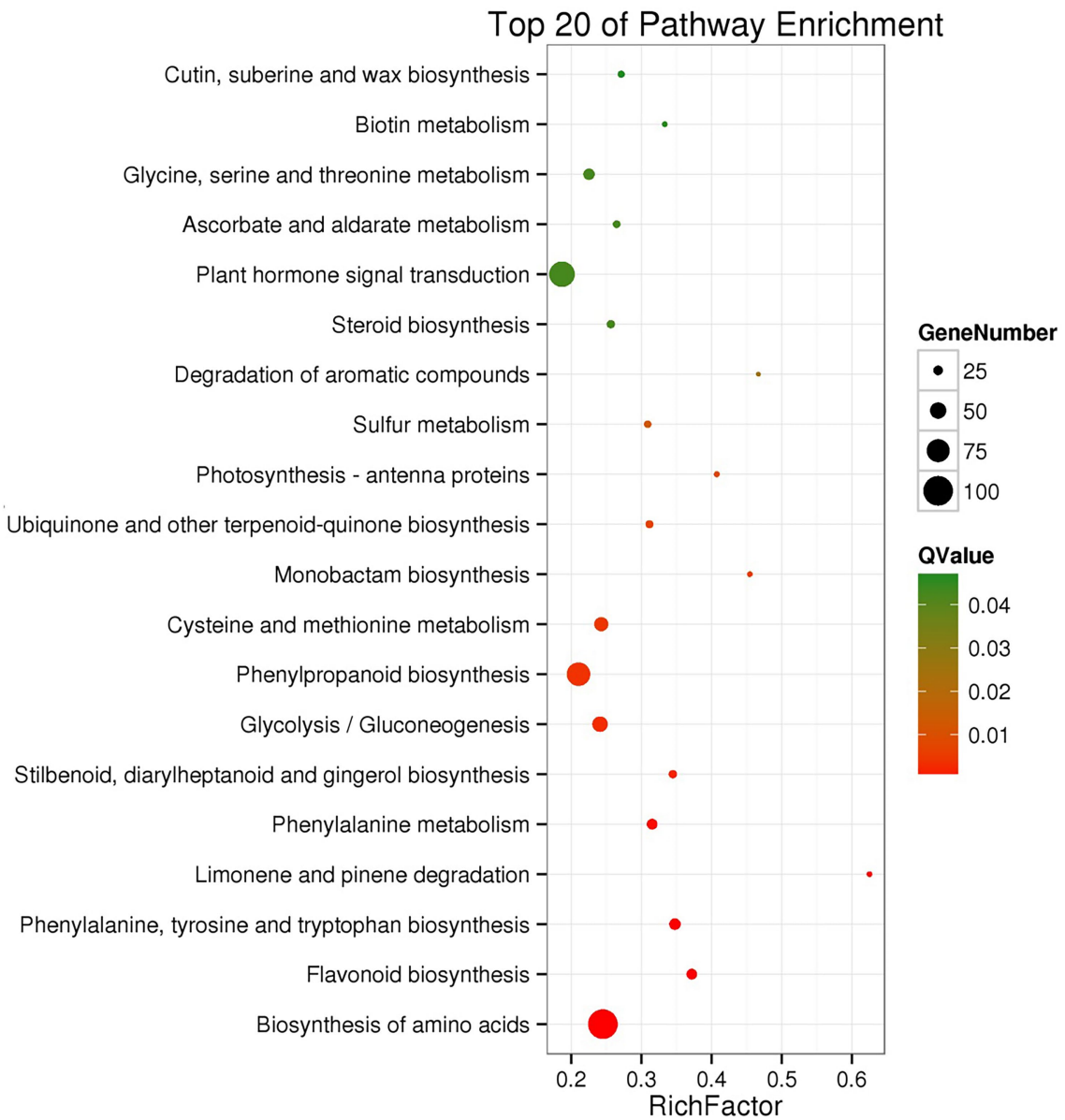


FIGURE 3 | KEGG pathway enrichment of differentially expressed genes after cold treatment. The solid circles represent differentially expressed genes located in the pathway. The larger of the solid circle represents the RichFactor enrichment level of the class. The *Q*-value is a value of *p* after a multiple hypothesis test correction from 0 to 1.

DEGs and RRBS analysis, a circos plot of epigenomic changes superimposed with DEGs was generated. The methylation sites of mCG, mCHG, and mCHH, and the expression levels of the corresponding DEGs were visualized at the genomic level. The DEGs that correlated with alterations in DNA methylation were distributed among 30 different scaffolds, which enable us to globally observe the methylation status and gene expression at the genomic width (**Figure 5**).

Integrated analysis of DEGs and modifications to DNA methylation indicated that highly expressed genes (red line) showed higher methylation levels in the gene body but low methylation levels in the upstream and downstream regions (**Figure 6**). By contrast, genes that were not expressed (pink line) showed lower methylation levels in the gene body but higher methylation levels in the downstream region at mCG, mCHG, and mCHH. Methylation status in upstream region might correlate with gene expression, there are fewer reports about the role of downstream region in gene transcription regulation, although plenty documents have shown that 3′UTR involve in mRNA-based processes (Mayr, 2017, 2019). It is reasonable to speculate that 2 kb downstream region of one gene might include part or even complete upstream sequence of the next adjacent gene. DMR-related differentially expression genes was analyzed by GO annotation. These genes involve

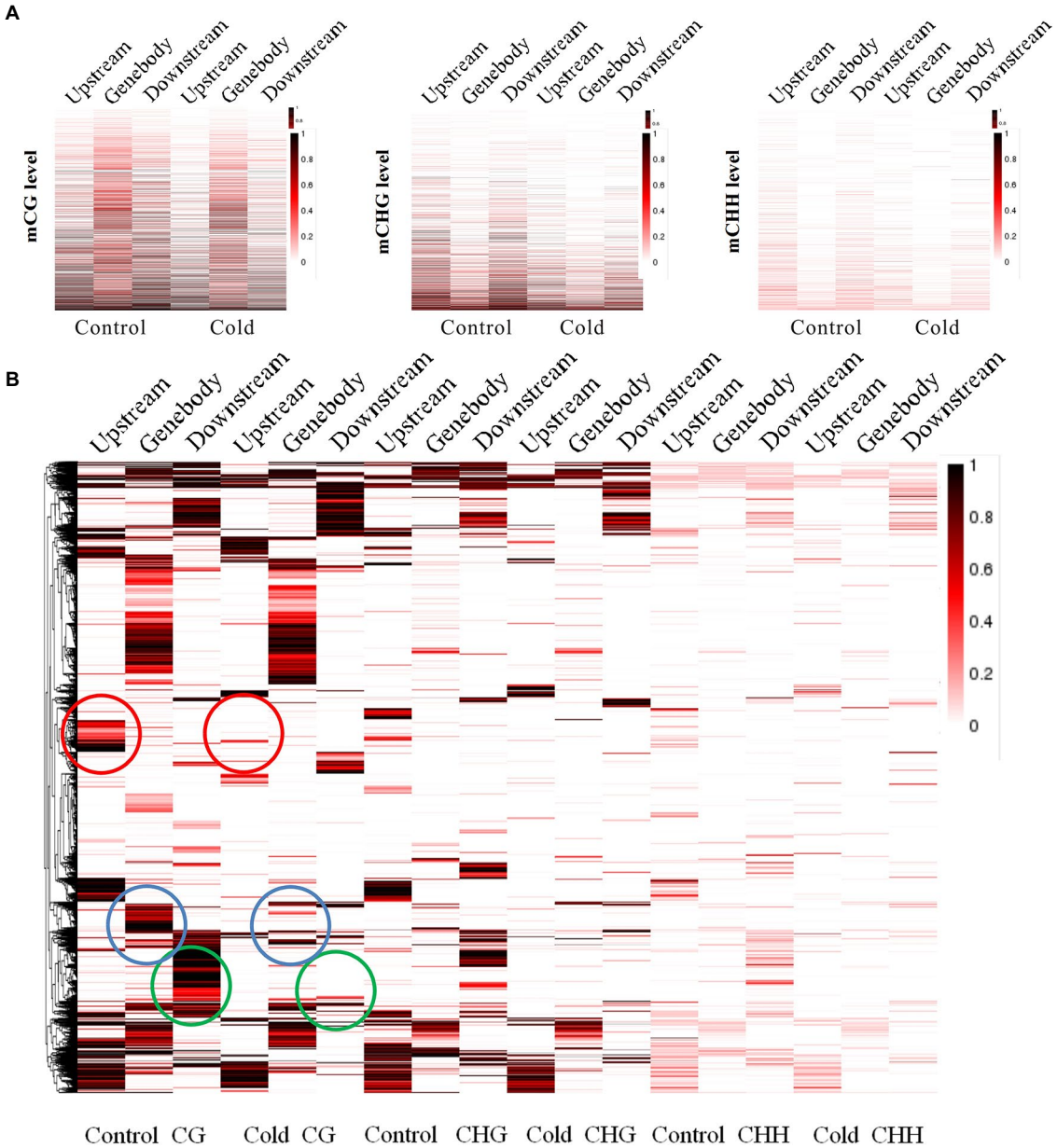


FIGURE 4 | Distribution and clustering of the methylation of differentially expressed genes. **(A)** The methylation levels of mCG, mCHG and mCHH in control and cold-treated samples are shown by a heat map (upper panel). The box plot shows CG, CHG, and CHH methylation levels at hyper-DMRs. Dark horizontal line, median; edges of boxes, 25th (bottom) and 75th (top) percentiles; whiskers, minimum and maximum percentages of DNA methylation (lower panel). **(B)** Heat map of gene clustering representing the methylation levels of differentially expressed genes at various gene regions. The methylation level is indicated by the depth of the color.

in many processes, such as protein modification, response to stimulus, metabolic process, macromolecular modification, and so on (**Supplementary Table S7**).

Validation of Cold-Induced Genes by qRT-PCR

Circle mapping and whole-genome visualisation using the Integrative Genomics Viewer (IGV) software revealed that more than 400 genes exhibited expression levels that were correlated with methylation modification (**Figure 6**). We selected 23 of these genes for validation by qRT-PCR. These genes included a cold-regulated plasma membrane protein gene (*HbCRP1*), chloroplast genes (*HbCHL1/2*), dehydratase genes (*HbDEH1/2/3*), a histone demethylase gene (*HbHDM1*), zinc finger protein genes (*HbCZF2/VZF3*), a methyltransferase gene (*HbTMET1*), a sugar transporter gene (*HbSTE3*), glycosyltransferase genes (*HbUGT1/GT1*), an early-responsive to dehydration gene (*HbERD1*), NAC domain-containing protein

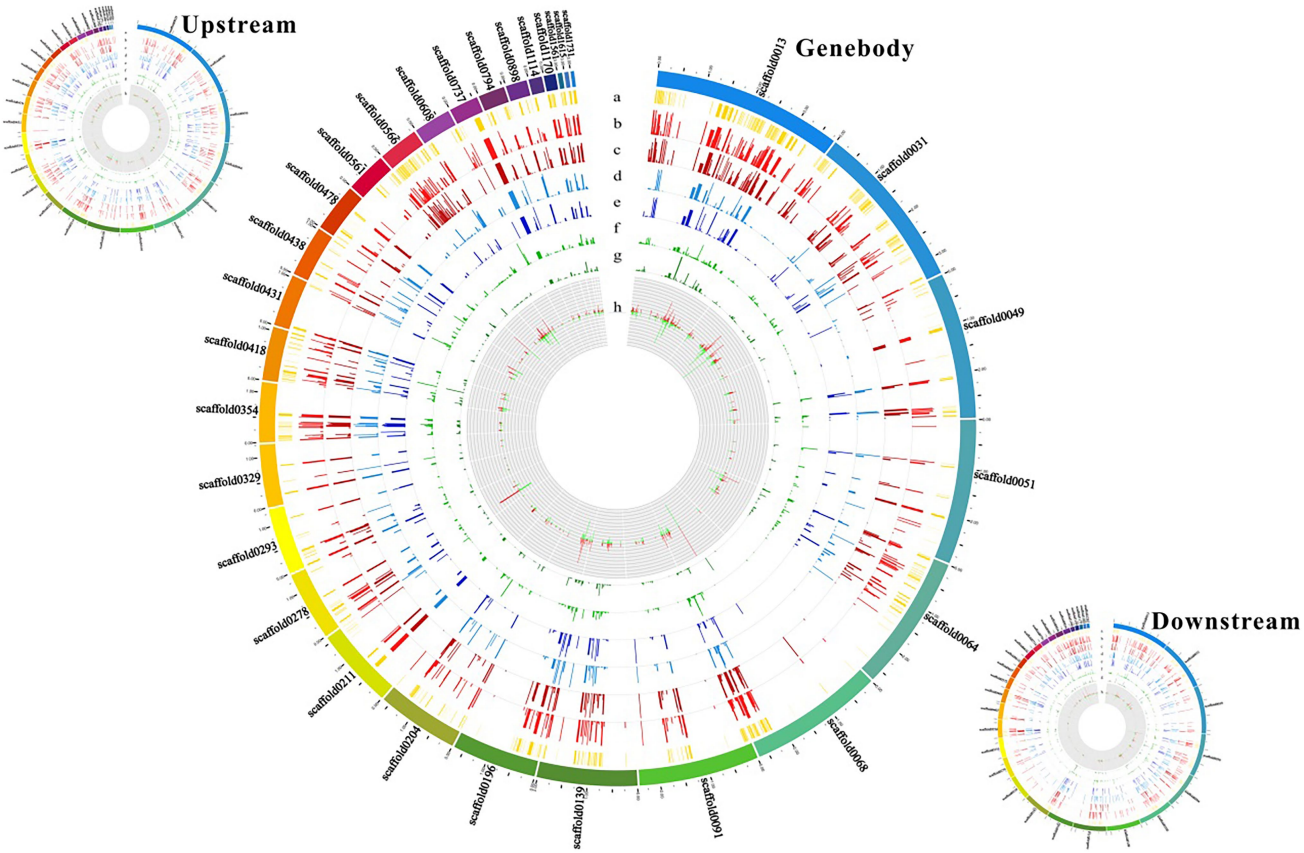


FIGURE 5 | Circos plot showing methylation sites and gene expression levels at the genomic level. Outer rings represent different scaffolds. (a) Yellow column represents the base site where the methylation takes place. (b) Red column represents the mCG methylation site in the control sample. (c) Dark red column represents the mCG methylation site in the cold-treated sample. (d) Blue column represents the mCHG methylation site in the control sample. (e) Dark blue column represents the mCHG methylation site in the cold-treated sample. (f) Green column represents the mCHH methylation site in the control sample. (g) Dark green column represents the mCHH methylation site in the cold-treated sample. (h) Up- and down−regulation of gene expression levels after cold treatment indicated by red and green colors, respectively.

genes (*HbNAC1/NAC2*), a dehydration-responsive element-binding protein gene (*HbDRE1*), transposons (*HbRTN1/2*), and ethylene-responsive transcription factor genes (*ERF*). The qRT-PCR analysis confirmed the cold treatment-induced expression of these genes shown by the DGE analysis (**Supplementary Figure S3**).

DISCUSSION

Being sessile organisms, higher plants inevitably endure more environmental stresses (e.g., cold, heat, UV irradiation, salinity, drought, and heavy metal) than animals in their life cycles (Puli and Raghavendra, 2012; Liu et al., 2017b). Of these abiotic stresses, cold stress is an important factor that limits plant range expansion and survival (Sanghera et al., 2011). When rubber tree was introduced to suboptimal planting areas, low temperature becomes the most detrimental factor influencing latex production and limits the expanding of rubber plantations to higher-latitude areas. All popular rubber tree cultivars are derived from "Wickham base," and little genetic variation was observed among the different cultivars of *H. brasiliensis* (Tang et al., 2016). Despite their similar genetic makeup, rubber tree cultivars exhibit various levels of cold tolerance (Mai et al., 2009, 2010). Epigenetic modification induced by location-specific environmental factors has been reported in rubber trees (Uthup et al., 2011). Our previous reports showed that cold treatment switched *HbICE1* promoters from hypermethylation status to hypomethylation status under both artificial and natural conditions, cold-induced DNA demethylation of *COR* gene might play a key role in cold acclimation in rubber tree (Tang et al., 2018).

In contrast to whole-genome bisulphite sequencing, which requires deep re-sequencing of the entire genome (Cokus et al., 2008; Zhong et al., 2013), RRBS enriches CG-rich parts through digestion with *Msp*I (recognized context sequence C′CGG) and ligation of genomic DNA to adapters for bisulphite conversion and Illumina sequencing (Gu et al., 2011). This work demonstrated that RRBS is a cost-effective method for genome-wide DNA methylation analysis. RRBS revealed that cold treatment induced global DNA demethylation and significant genome-wide alteration of the sequence contexts of methylated cytosines in *H. brasiliensis*.

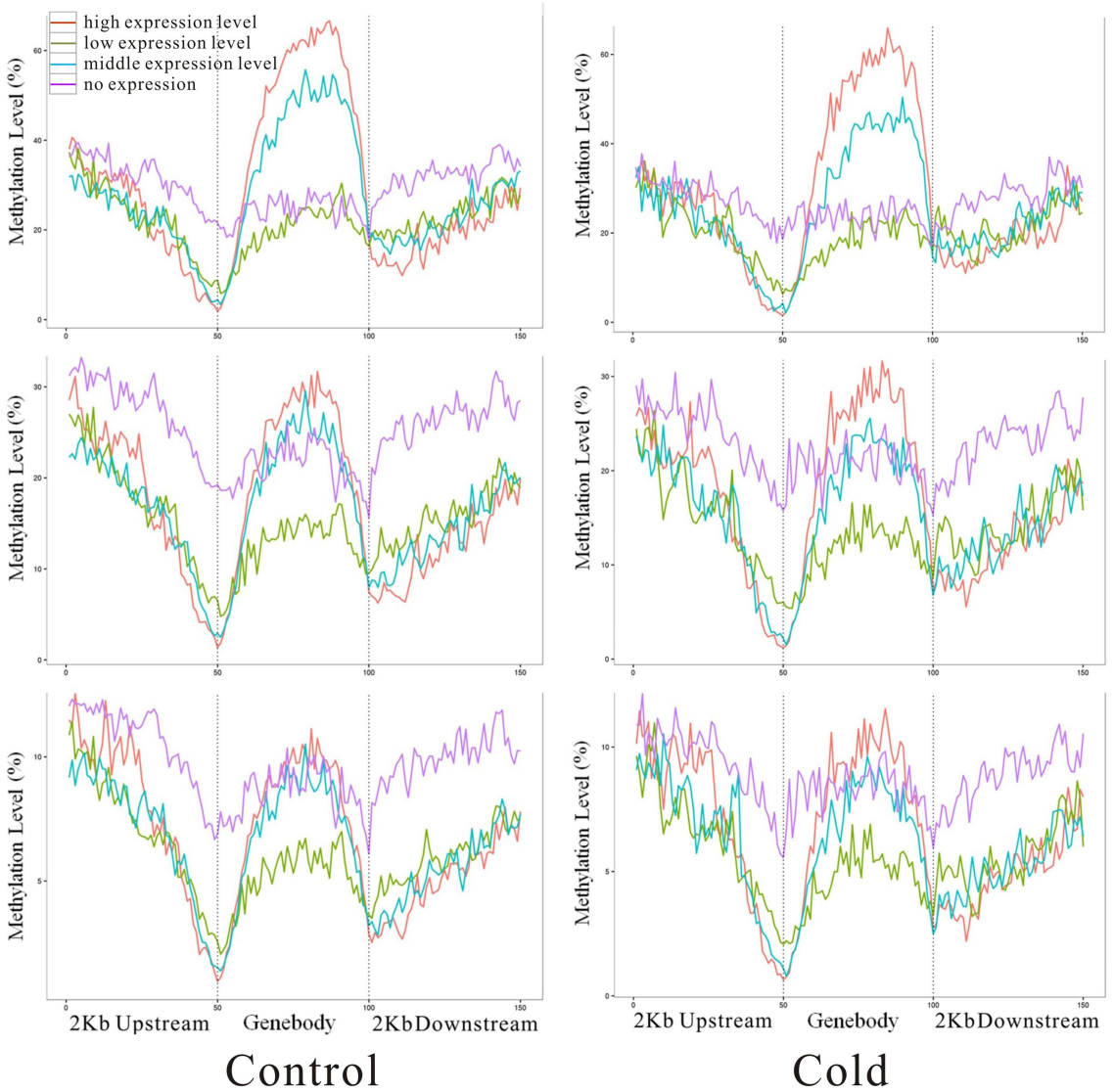


FIGURE 6 | Correlation of methylated cytosines with gene expression levels. The methylation level is indicated on the *y*-axis, and *x*-axis indicates the regions of 2 kb up and downstream of genebody. Red, green, blue, and pink colors represent high, moderate, low, and no expression of genes, respectively.

Preexisting DNA methylation can be lost as a consequence of passive or active demethylation processes. The active removal of cytosine methylation is catalyzed by the members of the DNA glycosylase family (Chinnusamy and Zhu, 2009a,b). Previous report showed that the demethylation-associated genes were strongly induced by cold treatment (Tang et al., 2018), suggesting that cold induced global DNA demethylation might be an active process in rubber tree. Similar phenomenon was observed in other plant species. When maize seedlings were exposed to cold stress, a genome-wide demethylation occurred in root tissues (Steward et al., 2002). Rapid alterations in cytosine methylation occurred throughout the periods of chilling and freezing associated with the cold tolerance of an alpine subnival plant, *Chorispora bungeana* (Song et al., 2015). It is the first time to report that cold treatment globally reprogrammed the epigenome of *H. brasiliensis*. Following the global DNA demethylation, extensive chromatin rearrangement and gene expression regulation might occur. Integrated analysis of the DNA methylome and the transcriptome revealed the correlation of gene expression with altered DNA methylation. DMR-related differentially expression genes involve in protein modification, response to stimulus, metabolic process, macromolecular modification, and so on. These genes might regulate physiological, biochemical, and metabolic processes to accumulate protective proteins and sugars and modify lipid metabolism to cope with cold stress. However, how such a demethylation locates its target regions remains to be explored.

Another interesting finding is the differential pattern of cold induced DNA methylation in TEs. Total mCG methylation was enhanced by cold treatment in the TEs. The differential

methylation patterns reflect the different roles that genes and TEs play in *H. brasiliensis* in coping with cold stress. TEs have long been known as useless "junk" DNA. These mysterious mobile elements of the genome might play some kind of regulatory role, determining which genes are turned on and when this activation takes place. TEs not only play a role in regulating gene expression, but also in generating different cell types and different biological structures (Pray, 2008). Enhanced methylation of TEs under cold stress influences the TEs transcription and decrease the "jumping" activity of TEs, which might affect the gene expression related to cold response. This phenomenon has not been described in previous reports. In contrast, hypomethylation and transposition of the Tam-3 transposon were induced by cold stress in *Antirrhinum majus* (Hashida et al., 2006). Repetitive elements were reported to be transcriptionally activated in a genotype-dependent manner in alfalfa during cold acclimation (Ivashuta et al., 2002).

Most of stress-induced epigenetic modifications are reset to the basal level once the stress is relieved, whereas some of the modifications may be stable. These epigenetic stress memories may help plants cope more effectively with subsequent stresses (Chinnusamy and Zhu, 2009a). Vernalisation is a well-documented "memory response" in which mCG methylation serves as a central epigenetic coordinator that ensures stable transgenerational inheritance (Mathieu et al., 2007). Integrated analysis of DEGs and changes in DNA methylation revealed several protective genes for tolerance to cold stress. *HbSTE3* is related to sugers transport and metabolism (Tang et al., 2010). *ERF* might involve in latex drainage related with ethylene signaling (Duan et al., 2013). *HbCRP1*, *HbERD1* and *HbDRE1* are well-known cold stress responsive genes of plant. NAC domain-containing protein gene (*HbNAC1/NAC2*) is involved in response to many abiotical stresses. These genes might help to maintain rubber production under cold stress.

A key question regarding the epigenetic regulation of the cold response in *H. brasiliensis* is whether cold-induced DNA demethylation is a transient or "memory" response. Under natural condition, the methylation status of *HbICE1* and *HbMET1* promoters changed seasonally (Tang et al., 2018), suggesting that these stress-induced epigenetic modifications were transient response. This work reveals a global cold induced DNA demethylation, it is still unknown whether it was transient or a stable "memory." Clarification of this question might help us to answer whether cold acclimation by methylation modification could be applied to tree breeding and production. Additionally, the stress-induced change of DNA methylation is usually coordinated with histone modification, e.g., cold response in vernalisation, which requires the epigenetic silencing of FLOWERING LOCUS C (FLC) by histone modification (Bastow et al., 2004; Sung and Amasino, 2004). Whether cold acclimation in the rubber tree also involves in histone modification requires further investigation.

CONCLUSION

Cold treatment induced global genomic DNA demethylation of *H. brasiliensis* and altered the sequence contexts of methylated cytosines, but increased the levels of mC methylation in TEs. Integrated analysis of the DNA methylome and transcriptome revealed 400 genes whose expression correlated with altered DNA methylation. DNA demethylation in the upstream region of gene seems to correlate with higher gene expression, whereas demethylation in the gene body has less association. Global change of the genomic DNA methylation status induced by cold treatment might coordinate reprogramming of the transcriptome in the rubber tree.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://www.ncbi.nlm.nih.gov/genbank/, NCBI SRA data: PRJNA741882.

AUTHOR CONTRIBUTIONS

XT performed most of the experiments. YZ and H-MY performed the bioinformatics analysis. XH and JZ supervised and wrote the manuscript. All authors have read and approved the final version of the manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (grant numbers 31860194 and 31860222).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.831839/full#supplementary-material

REFERENCES

Akalin, A., Kormaksson, M., Li, S., Garrett-Bakelman, F. E., Figueroa, M. E., Melnick, A., et al. (2012). methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol.* 13:R87. doi: 10.1186/gb-2012-13-10-r87

Bastow, R., Mylne, J. S., Lister, C., Lippman, Z., Martienssen, R. A., and Dean, C. (2004). Vernalization requires epigenetic silencing of FLC by histone methylation. *Nature* 427, 164–167. doi: 10.1038/nature02269

Cao, X., and Jacobsen, S. E. (2002). Role of the arabidopsis DRM methyltransferases in de novo DNA methylation and gene silencing. *Curr. Biol.* 12, 1138–1144. doi: 10.1016/S0960-9822(02)00925-9

Chen, X., Ge, X., Wang, J., Tan, C., King, G. J., and Liu, K. (2015). Genome-wide DNA methylation profiling by modified reduced representation bisulfite sequencing in Brassica rapa suggests that epigenetic modifications play a key role in polyploid genome evolution. *Front. Plant Sci.* 6:836. doi: 10.3389/fpls.2015.00836

Chinnusamy, V., and Zhu, J. K. (2009a). Epigenetic regulation of stress responses in plants. *Curr. Opin. Plant Biol.* 12, 133–139. doi: 10.1016/j.pbi.2008.12.006

Chinnusamy, V., and Zhu, J. K. (2009b). RNA-directed DNA methylation and demethylation in plants. *Sci China C Life Sci.* 52, 331–343. doi: 10.1007/s11427-009-0052-1

Choi, C. S., and Sano, H. (2007). Abiotic-stress induces demethylation and transcriptional activation of a gene encoding a glycerophosphodiesterase-like protein in tobacco plants. *Mol. Gen. Genomics* 277, 589–600. doi: 10.1007/s00438-007-0209-1

Cokus, S. J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C. D., et al. (2008). Shotgun bisulphite sequencing of the arabidopsis genome reveals DNA methylation patterning. *Nature* 452, 215–219. doi: 10.1038/nature06745

Duan, C., Argout, X., Gebelin, V., Summo, M., Dufayard, J. F., Leclercq, J., et al. (2013). Identification of the *Hevea brasiliensis* AP2/ERF superfamily by RNA sequencing. *BMC Genomics* 14:30. doi: 10.1186/1471-2164-14-30

Fowler, S., and Thomashow, M. F. (2002). Arabidopsis transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway. *Plant Cell* 14, 1675–1690. doi: 10.1105/tpc.003483

Gendall, A. R., Levy, Y. Y., Wilson, A., and Dean, C. (2001). The VERNALIZATION 2 gene mediates the epigenetic regulation of vernalization in arabidopsis. *Cell* 107, 525–535. doi: 10.1016/S0092-8674(01)00573-6

Gu, H., Smith, Z. D., Bock, C., Boyle, P., Gnirke, A., and Meissner, A. (2011). Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nat. Protoc.* 6, 468–481. doi: 10.1038/nprot.2010.190

Hashida, S. N., Uchiyama, T., Martin, C., Kishima, Y., Sano, Y., and Mikami, T. (2006). The temperature-dependent change in methylation of the antirrhinum transposon Tam3 is controlled by the activity of its transposase. *Plant Cell* 18, 104–118. doi: 10.1105/tpc.105.037655

Henderson, I. R., and Jacobsen, S. E. (2007). Epigenetic inheritance in plants. *Nature* 447, 418–424. doi: 10.1038/nature05917

Ivashuta, S., Naumkina, M., Gau, M., Uchiyama, K., Isobe, S., Mizukami, Y., et al. (2002). Genotype-dependent transcriptional activation of novel repetitive elements during cold acclimation of alfalfa (*Medicago sativa*). *Plant J.* 31, 615–627. doi: 10.1046/j.1365-313x.2002.01383.x

Jacob, J., Annmalainathan, K., Alam, B. M., Sathick, M. B., Thapaliyal, A. P., and Devakumar, A. S. (1999). Physiological constraints for cultivation of *Hevea brasiliensis* in certain unfavourable agroclimatic regions of India. *Indian J. Nat. Rub. Res.* 12, 1–16.

Jaglo-Ottosen, K. R., Gilmour, S. J., Zarka, D. G., Schabenberger, O., and Thomashow, M. F. (1998). Arabidopsis CBF1 overexpression induces COR genes and enhances freezing tolerance. *Science* 280, 104–106. doi: 10.1126/science.280.5360.104

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14:R36. doi: 10.1186/gb-2013-14-4-r36

Knight, M. R., and Knight, H. (2012). Low-temperature perception leading to gene expression and cold tolerance in higher plants. *New Phytol.* 195, 737–751. doi: 10.1111/j.1469-8137.2012.04239.x

Law, J. A., and Jacobsen, S. E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet.* 11, 204–220. doi: 10.1038/nrg2719

Li, H. L., Guo, D., Yang, Z. P., Tang, X., and Peng, S. Q. (2014). Genome-wide identification and characterization of WRKY gene family in *Hevea brasiliensis*. *Genomics* 104, 14–23. doi: 10.1016/j.ygeno.2014.04.004

Li, H. L., Guo, D., Zhu, J. H., Wang, Y., Chen, X. T., and Peng, S. Q. (2016). Comparative transcriptome analysis of latex reveals molecular mechanisms underlying increased rubber yield in *Hevea brasiliensis* self-rooting juvenile clones. *Front. Plant Sci.* 7:1204. doi: 10.3389/fpls.2016.01204

Lister, R., Pelizzola, M., Dowen, R. H., Hawkins, R. D., Hon, G., Tonti-Filippini, J., et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 462, 315–322. doi: 10.1038/nature08514

Liu, T., Li, Y., Duan, W., Huang, F., and Hou, X. (2017a). Cold acclimation alters DNA methylation patterns and confers tolerance to heat and increases growth rate in Brassica rapa. *J. Exp. Bot.* 68, 1213–1224. doi: 10.1093/jxb/erw496

Liu, Y., Song, Q., Li, D., Yang, X., and Li, D. (2017b). Multifunctional roles of plant dehydrins in response to environmental stresses. *Front. Plant Sci.* 8:1018. doi: 10.3389/fpls.2017.01018

Mai, J., Herbette, S., Vandame, M., Cavaloc, E., Julien, J. L., Ameglio, T., et al. (2010). Contrasting strategies to cope with chilling stress among clones of a tropical tree *Hevea brasiliensis*. *Tree Physiol.* 30, 1391–1402. doi: 10.1093/treephys/tpq075

Mai, J., Herbette, S., Vandame, M., Kositsup, B., Kasemsap, P., Cavaloc, E., et al. (2009). Effect of chilling on photosynthesis and antioxidant enzymes in *Hevea brasiliensis* Muell. *Arg. Trees* 23, 863–874. doi: 10.1007/s00468-009-0328-x

Mathieu, O., Reinders, J., Caikovski, M., Smathajitt, C., and Paszkowski, J. (2007). Transgenerational stability of the arabidopsis epigenome is coordinated by CG methylation. *Cell* 130, 851–862. doi: 10.1016/j.cell.2007.07.007

Mayr, C. (2017). Regulation by 3′-untranslated regions. *Annu. Rev. Genet.* 51, 171–194. doi: 10.1146/annurev-genet-120116-024704

Mayr, C. (2019). What are 3′ UTRs doing? *Cold Spring Harb. Perspect. Biol.* 11. doi: 10.1101/cshperspect.a034728

Penterman, J., Zilberman, D., Huh, J. H., Ballinger, T., Henikoff, S., and Fischer, R. L. (2007). DNA demethylation in the arabidopsis genome. *Proc. Natl. Acad. Sci. U. S. A.* 104, 6752–6757. doi: 10.1073/pnas.0701861104

Pray, L. (2008). Transposons, or jumping genes: not junk DNA? *Nat. Educ.* 1, 32.

Puli, M. R., and Raghavendra, A. S. (2012). Pyrabactin, an ABA agonist, induced stomatal closure and changes in signalling components of guard cells in abaxial epidermis of *Pisum sativum*. *J. Exp. Bot.* 63, 1349–1356. doi: 10.1093/jxb/err364

Rao, P. S., Saraswathyamma, C. K., and Sethuraj, M. R. (1998). Studies on the relationship between yield and meteorological parameters of Para rubber tree (*Hevea brasiliensis*). *Agric. For. Meteorol.* 90, 235–245. doi: 10.1016/S0168-1923(98)00051-3

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616

Sanghera, G. S., Wani, S. H., Hussain, W., and Singh, N. B. (2011). Engineering cold stress tolerance in crop plants. *Curr. Genom.* 12, 30–43. doi: 10.2174/138920211794520178

Song, Y., Liu, L., Feng, Y., Wei, Y., Yue, X., He, W., et al. (2015). Chilling- and freezing-induced alterations in cytosine methylation and its association with the cold tolerance of an alpine Subnival plant *Chorispora bungeana*. *PLoS One* 10:e0135485. doi: 10.1371/journal.pone.0135485

Steward, N., Ito, M., Yamaguchi, Y., Koizumi, N., and Sano, H. (2002). Periodic DNA methylation in maize nucleosomes and demethylation by environmental stress. *J. Biol. Chem.* 277, 37741–37746. doi: 10.1074/jbc.M204050200

Strand, A., Hurry, V., Henkes, S., Huner, N., Gustafsson, P., Gardestrom, P., et al. (1999). Acclimation of arabidopsis leaves developing at low temperatures. Increasing cytoplasmic volume accompanies increased activities of enzymes in the Calvin cycle and in the sucrose-biosynthesis pathway. *Plant Physiol.* 119, 1387–1398. doi: 10.1104/pp.119.4.1387

Sung, S., and Amasino, R. M. (2004). Vernalization in arabidopsis thaliana is mediated by the PHD finger protein VIN3. *Nature* 427, 159–164. doi: 10.1038/nature02195

Tanas, A. S., Borisova, M. E., Kuznetsova, E. B., Rudenko, V. V., Karandasheva, K. O., Nemtsova, M. V., et al. (2017). Rapid and affordable genome-wide bisulfite DNA sequencing by XmaI-reduced representation bisulfite sequencing. *Epigenomics* 9, 833–847. doi: 10.2217/epi-2017-0031

Tang, C., Huang, D., Yang, J., Liu, S., Sakr, S., Li, H., et al. (2010). The sucrose transporter HbSUT3 plays an active role in sucrose loading to laticifer and rubber productivity in exploited trees of *Hevea brasiliensis* (Para rubber tree). *Plant Cell Environ.* 33, 1708–1720. doi: 10.1111/j.1365-3040.2010.02175.x

Tang, X., Wang, Q., Yuan, H., and Huang, X. (2018). Chilling-induced DNA demethylation is associated with the cold tolerance of *Hevea brasiliensis*. *BMC Plant Biol.* 18:70. doi: 10.1186/s12870-018-1276-7

Tang, C., Yang, M., Fang, Y., Luo, Y., Gao, S., Xiao, X., et al. (2016). The rubber tree genome reveals new insights into rubber production and species adaptation. *Nat Plants* 2:16073. doi: 10.1038/nplants.2016.73

Theocharis, A., Clement, C., and Barka, E. A. (2012). Physiological and molecular changes in plants grown at low temperatures. *Planta* 235, 1091–1105. doi: 10.1007/s00425-012-1641-y

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., et al. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and cufflinks. *Nat. Protoc.* 7, 562–578. doi: 10.1038/nprot.2012.016

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., et al. (2010). Transcript assembly and quantification by RNA-Seq reveals

unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28, 511–515. doi: 10.1038/nbt.1621

Uthup, T. K., Ravindran, M., Bini, K., and Thakurdas, S. (2011). Divergent DNA methylation patterns associated with abiotic stress in *Hevea brasiliensis*. *Mol. Plant* 4, 996–1013. doi: 10.1093/mp/ssr039

Vanyushin, B. F., and Ashapkin, V. V. (2011). DNA methylation in higher plants: past, present and future. *Biochim. Biophys. Acta* 1809, 360–368. doi: 10.1016/j.bbagrm.2011.04.006

Wang, Y., Wu, J., Ma, X., Liu, B., Su, R., Jiang, Y., et al. (2015). Single base-resolution methylome of the dizygotic sheep. *PLoS One* 10:e0142034. doi: 10.1371/journal.pone.0142034

Wani, S. H., Singh, N. B., Haribhushan, A., and Mir, J. I. (2013). Compatible solute engineering in plants for abiotic stress tolerance - role of glycine betaine. *Curr. Genom.* 14, 157–165. doi: 10.2174/1389202911314030001

Xi, Y., and Li, W. (2009). BSMAP: whole genome bisulfite sequence MAPping program. *BMC Bioinform.* 10:232. doi: 10.1186/1471-2105-10-232

Xia, Z., Xu, H., Zhai, J., Li, D., Luo, H., He, C., et al. (2011). RNA-Seq analysis and de novo transcriptome assembly of *Hevea brasiliensis*. *Plant Mol. Biol.* 77, 299–308. doi: 10.1007/s11103-011-9811-z

Young, M. D., Wakefield, M. J., Smyth, G. K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 11:R14. doi: 10.1186/gb-2010-11-2-r14

Zhong, S., Fei, Z., Chen, Y. R., Zheng, Y., Huang, M., Vrebalov, J., et al. (2013). Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening. *Nat. Biotechnol.* 31, 154–159. doi: 10.1038/nbt.2462

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 22 March 2022
doi: 10.3389/fpls.2022.843107

Genome-Wide Analysis of Late Embryogenesis Abundant Protein Gene Family in *Vigna* Species and Expression of *VrLEA* Encoding Genes in *Vigna glabrescens* Reveal Its Role in Heat Tolerance

Chandra Mohan Singh[1]*, Mukul Kumar[1], Aditya Pratap[2]*, Anupam Tripathi[1], Smita Singh[1], Anuj Mishra[1], Hitesh Kumar[1], Ramkrishnan M. Nair[3] and Narendra Pratap Singh[1]

[1] Department of Genetics and Plant Breeding, Banda University of Agriculture and Technology, Banda, India, [2] ICAR-Indian Institute of Pulses Research, Kanpur, India, [3] World Vegetable Center South Asia, Hyderabad, India

Late embryogenesis abundant (LEA) proteins are identified in many crops for their response and role in adaptation to various abiotic stresses, such as drought, salinity, and temperature. The *LEA* genes have been studied systematically in several crops but not in *Vigna* crops. In this study, we reported the first comprehensive analysis of the *LEA* gene family in three legume species, namely, mung bean (*Vigna radiata*), adzuki bean (*Vigna angularis*), and cowpea (*Vigna unguiculata*), and the cross-species expression of *VrLEA* genes in a wild tetraploid species, *Vigna glabrescens*. A total of 201 *LEA* genes from three *Vigna* crops were identified harboring the LEA conserved motif. Among these 55, 64, and 82 *LEA* genes were identified in mung bean, adzuki bean, and cowpea genomes, respectively. These *LEA* genes were grouped into eight different classes. Our analysis revealed that the cowpea genome comprised all eight classes of *LEA* genes, whereas the LEA-6 class was absent in the mung bean genome. Similarly, LEA-5 and LEA-6 were absent in the adzuki bean genome. The analysis of *LEA* genes provides an insight into their structural and functional diversity in the *Vigna* genome. The genes, such as *VrLEA-2*, *VrLEA-40*, *VrLEA-47*, and *VrLEA-55,* were significantly upregulated in the heat-tolerant genotype under stress conditions indicating the basis of heat tolerance. The successful amplification and expression of *VrLEA* genes in *V. glabrescens* indicated the utility of the developed markers in mung bean improvement. The results of this study increase our understanding of *LEA* genes and provide robust candidate genes for future functional investigations and a basis for improving heat stress tolerance in *Vigna* crops.

Keywords: abiotic stress, candidate genes, expression analysis, heat stress, mung bean, wild *Vigna*, *LEA* genes

OPEN ACCESS

Edited by:
Rajeev K. Varshney, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), India

Reviewed by:
Harsh Nayyar, Panjab University, India
Monika Dalal, ICAR-NIPB, India
Charu Lata, National Institute of Science Communication and Information Resources (CSIR), India

***Correspondence:**
Chandra Mohan Singh
cmsingh.gpb@gmail.com
orcid.org/0000-0002-3857-7361
Aditya Pratap
Aditya.Pratap@icar.gov.in
orcid.org/0000-0001-7280-0953

Specialty section:
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

Received: *24 December 2021*
Accepted: *02 February 2022*
Published: *22 March 2022*

Citation:
Singh CM, Kumar M, Pratap A, Tripathi A, Singh S, Mishra A, Kumar H, Nair RM and Singh NP (2022) Genome-Wide Analysis of Late Embryogenesis Abundant Protein Gene Family in Vigna Species and Expression of VrLEA Encoding Genes in Vigna glabrescens Reveal Its Role in Heat Tolerance. Front. Plant Sci. 13:843107. doi: 10.3389/fpls.2022.843107

INTRODUCTION

Mung bean (*Vigna radiata*) is an important short-duration pulse crop occupying a large area, predominantly in the Indian subcontinent, besides many other places in Australia, Africa, and United States. It is primarily consumed as food and animal feed and also as a green manuring crop (Pratap et al., 2021). It has a high nutrition value (Pratap et al., 2020) with a prominence of quality protein besides having soil ameliorative properties (Allito et al., 2015; Sehrawat et al., 2015; Singh et al., 2017). Although several high-yielding and superior varieties have been developed over the last few decades in the mung bean due to their climate sensitivity, improving their productivity in extreme environments is still a major challenge (Singh et al., 2014, 2019a; Manivannan and Mahalingam, 2021). In legume species, such as *Vigna* crops, seed yield is highly affected by various abiotic stresses (HanumanthaRao et al., 2016; Douglas et al., 2020). Drought and high temperature are the most limiting factors (Douglas et al., 2020; Kumar et al., 2020; Singh et al., 2021) among the abiotic stresses, affecting the crops at different growth stages, leading to severe yield penalties. Varieties respond differentially to abiotic stresses as per stress duration, crop growth stage, and genetic potential of a variety, which leads to moderate-to-severe yield loss (Bangar et al., 2019). Hence, there is an urgent need to develop abiotic stress-tolerant varieties to mitigate the effects of climate change.

The heat stress negatively affects the seedling vigor, biomass accumulation, and reproductive development in mung bean (Tzudir et al., 2014; HanumanthaRao et al., 2016). It also affects all reproductive traits, such as flower initiation, pollen viability, stigma receptivity, ovule size and viability, fertilization, pod set, grain filling, and seed quality (Barnabas et al., 2008; Wassmann et al., 2009; Douglas et al., 2020). Flower shedding is very common in mung bean under heat stress (Kumari and Varma, 1983) while it severely affects flower bud initiation, and this sensitivity may prevail for 10–15 days (Bita and Gerats, 2013). Terminal heat stress is a major problem in spring- and summer-grown crops, whereas heat stress during early vegetative growth is frequent in *Kharif* (rainy) season.

A few potential donors have been identified and characterized among the cultivated germplasm and wild relatives in earlier studies against various abiotic stresses (Pratap et al., 2014; Sharma et al., 2016; Singh et al., 2021), which can be included in breeding programs for mung bean improvement. Pratap et al. (2014) characterized two wild accessions belonging to *Vigna umbellata* and *Vigna glabrescens* for photo- and thermoperiod insensitivity. In this study, we investigated the gene expression of *VrLEA* genes in heat-sensitive (HS) mung bean and heat-tolerant (HT) creole bean (*V. glabrescens*) genotypes for understanding the response of selected candidate genes under high-temperature stress. Wild species can be especially useful toward integrating crossable donors in breeding programs and help in improving stress tolerance in domesticated cultivars and broadening the genetic base of the crop.

Abiotic stresses decline crop productivity and seed quality by serious disruptions in plant growth and development (Xiong and Zhu, 2002; Ahuja et al., 2010). Therefore, to defend these adverse climatic conditions, plants have evolved complex regulatory mechanisms at molecular, physiological, and biochemical levels to minimize the effects of abiotic stresses (Zhu, 2016). Proteins play a major role in regulating the various signaling pathways that activate under abiotic stresses and indirectly improve stress tolerance. The late embryogenesis abundant (LEA) protein gene family is considered as an important group of functional proteins to reduce the scope of abiotic stresses in plants (Hirayama and Shinozaki, 2010; Debnath et al., 2011). These proteins can eliminate the cellular content of reactive oxygen species (ROS) to protect the macromolecular substances and alleviate damages caused by the various abiotic stresses (Yu et al., 2019).

The LEA genes were first identified in mature cotton seeds (Dure et al., 1981), which are found to be accumulated during dehydration and maturation of seeds, and protect them from damage. They have also been detected in various plant parts, such as seedlings, leaves, stems, roots, and other organs, under abiotic stress (Hincha and Thalhammer, 2012). Later, LEA proteins have also been identified in other plants, such as *Arabidopsis* (Finkelstein, 1993; Hundertmark and Hincha, 2008), *Brassica napus* (Liang et al., 2016), maize (Li and Cao, 2016), rice (Wang et al., 2007), soybean (Liu et al., 2018), wheat (Liu et al., 2019), grape wine (İbrahime et al., 2019), poplar (Cheng et al., 2021), and many more species. LEA proteins are highly hydrophilic and intrinsically unstructured, whereas they partially fold into mainly α-helical structures under dehydration state (Hand et al., 2011). This allows them to contribute the function as chaperones (Hanin et al., 2011) and also in the stabilization of membranes, calcium-binding, and metal homeostasis and the protection of functional proteins against aggregation (Krueger et al., 2002; Tolleter et al., 2010; Rahman et al., 2011; Rosales et al., 2014). LEA proteins are divided into eight different clusters, including LEA-1, LEA-2, LEA-3, LEA-4, LEA-5, LEA-6, dehydrin (DHN), and seed maturation protein (SMP) (Hundertmark and Hincha, 2008; Finn et al., 2010; Hunault and Jaspard, 2010).

While the LEA gene family also plays an important role in growth and development under multiple stress conditions, the study of this gene family in *Vigna* is still lacking. Many *Vigna* species are now sequenced, such as mung bean (Kang et al., 2014), adzuki bean (Kang et al., 2015), cowpea (Lonardi et al., 2019), beach pea (Singh et al., 2019b), moth bean (Takahashi et al., 2019), and black gram (Souframanien et al., 2020), with their sequence data available in the public database and can be accessed for analysis. Keeping this in view, this experiment was conducted to survey LEA genes and study their comparative structural diversity analysis in mung bean, adzuki bean, and cowpea. The expression of selected *VrLEA* genes through quantitative real-time PCR (qRT-PCR) was performed on HS mung bean and HT creole bean genotypes. It provides new insights for understanding the mechanism of heat tolerance in *Vigna* species.

MATERIALS AND METHODS

Plant Materials and Stress Treatment

The plant materials comprised a high-yielding and HS-advanced breeding line of mung bean, IPM 312-19, and one HT-wild creole

bean (*V. glabrescens*), accession IC251372. The plants of both the genotypes were grown in the net house during *Kharif* (rainy season) 2021 at Banda University of Agriculture and Technology, Banda, located at 24°53′ latitude and 25°55′ latitude. Both the genotypes were subjected to heat-shock treatment through the detached leaf method following the study by Kumar et al. (2009) with modification as well as *in planta*.

Experiment 1

At the early vegetative stage (20 days after germination), the expanded second trifoliate leaves from the top of the plants were subjected to heat stress treatment through the detached leaf method. This was performed in a growth chamber. The leaves from both the test genotypes (IPM 312-19 and IC 251372) were plucked and kept in Petri plates in distilled water, and heat shock was induced at two levels, namely, HSI 1: 45°C for 3 h with 60% relative humidity (RH); and HSI 2: 45°C for 6 h with 60% RH. For the control treatment (non-stressed), the detached trifoliate leaves were kept in Petri plates in distilled water and put into the growth chamber for 6 h at 25°C temperature with 60% RH to recover them from the injury induced by plucking. Immediately after the heat shock, the stressed and control (non-stressed) samples were taken out, blot-dried by slightly pressing between two layers of filter paper, frozen in liquid nitrogen, and stored at −80°C for further study.

Experiment 2

Another experiment was performed *in planta* under laboratory conditions in a plant growth chamber. The plants of both the genotypes were raised in a potting mixture in plastic pots with ideal conditions, such as 30°C temperature with 60% RH. The 7-day-old seedlings were subjected to heat stress. The leaf samples without induction of heat stress were kept as control (non-stress) while those taken 24 h after heat stress (45°C, 60% RH) were considered as stressed plants. The leaf samples from the stressed plants were immediately frozen in liquid nitrogen after plucking at different time points and stored at −80°C for further study.

RNA Extraction and cDNA Synthesis

A total of 200 mg of frozen leaf samples were homogenized in liquid nitrogen, and RNA was extracted using a Plant RNA Extraction Kit (RNeasy Mini Kit, QIAGEN) following the manufacturer's instructions. Subsequently, RNA was reverse-transcribed from 1 μg of RNA by the RevertAid First-Strand cDNA Synthesis Kit (Thermo Fisher Scientific). The cDNA was normalized by 50 ng/μl for qRT-PCR analysis.

Identification of Late Embryogenesis Abundant Protein Sequences

The genome-wide data of mung bean (*V. radiata*), adzuki bean (*Vigna angularis*), and cowpea (*Vigna unguiculata*) were obtained from Legume Information System (LIS) online website[1]. The typical LEA protein genes were retrieved through Pfam analysis with PF03760, PF03168, PF03242, PF02987, PF0477, PF10714, PF0492, and PF00257[2]. The redundancy of sequences was eliminated by ExPASy[3] to generate a unique set of LEA proteins. The LEA proteins were scanned against the mung bean, adzuki bean, and cowpea genomes by using the HMMER 3.0 program[4] followed by manual verification with the Pfam database[5]. The protein sequences that did not comprise LEA domains were removed and not included in the analysis.

Chromosomal Distribution and Phylogenetic Analysis of Late Embryogenesis Abundant Genes

The physical positions (in bps) of the *VrLEA, VaLEA,* and *VuLEA* genes were identified in this study, and the length of chromosomes was obtained from the LIS database. These genes from short arm to long arm were mapped onto their corresponding chromosomes in ascending order using MapChart 2.3 (Voorrips, 2002). The LEA genes identified in the scaffold were not assigned on the chromosome. The putative *VrLEA, VaLEA,* and *VuLEA* genes were classified into eight different groups. The phylogenetic trees were constructed with all the putative LEA genes identified in this study. Multiple alignments of the deduced amino acid sequences were performed using ClustalW (Thompson et al., 2002). The phylogenetic tree was constructed through the maximum likelihood (ML) method (Guindon and Gascuel, 2003) provided in the MEGA 7.0 tool (Tamura et al., 2013) based on the Jones-Taylor-Thornton (JTT) matrix-based model. The bootstrap analysis was set with 1,000 replications for the accuracy of the constructed tree.

Gene Structure Prediction and Protein Analysis of Late Embryogenesis Abundant Genes

The CDS and the genomic DNA sequences of the corresponding *LEA* genes were retrieved by blast analysis on the NCBI database. The gene structures of LEA were determined by comparing the genomic DNA sequences and their corresponding coding sequences using the Gene Structure Display Server version 2.0 (Hu et al., 2015)[6]. The molecular weight (kDa) and isoelectric point (pI) of LEA sequences were predicted with EndMemo[7] and pI calculator[8], respectively (Gasteiger et al., 2003). The motifs were analyzed through the MEME program[9] (Bailey et al., 2009). In motif analysis, the maximum number of motifs was preset to 10 while the optimum width of motifs was set to 6–50 amino acid residues, and other settings were kept as default.

[1]https://legumeinfo.org/

[2]https://legumeinfo.org/search/gene

[3]https://web.expasy.org/decrease_redundancy/

[4]https://www.ebi.ac.uk/Tools/hmmer/

[5]http://pfam.xfam.org/

[6]http://gsds.gao-lab.org/

[7]http://www.endmemo.com/bio/promw.php

[8]http://isoelectric.org/

[9]https://meme-suite.org/meme/tools/meme

TABLE 1 | The proportions of late embryogenesis abundant (*LEA*) genes in three *Vigna* and *Medicago truncatula* genomes*.

Description	Mung bean (*V. radiata*)	Adzuki bean (*V. angularis*)	Cowpea (*V. unguiculata*)	*Medicago truncatula*
Total LEA family proteins	55	64	82	96
LEA-1	01	02	05	06
LEA-2	38	45	55	68
LEA-3	06	10	08	05
LEA-4	02	02	03	02
LEA-5	02	00	02	02
LEA-6	00	00	01	02
SMP	05	02	04	06
DHN	01	03	04	05
Total protein-coding genes	26961	26634	28314	31927
LEA genes (%)	0.20	0.24	0.29	0.30
Genome size (Mb)	459.27	447.81	607.06	419.47
Average *LEA* genes per Mb	0.12	0.14	0.13	0.23

**Data accessed from the LIS database on May 10, 2021.*

Gene-Based Primer Designing and Quantitative Real-Time PCR Analysis

All the *VrLEA* candidate genes were blasted against the *V. glabrescens* genome, and the qRT-PCR primers were designed using the Primer-BLAST tool[10]. Six gene-specific qRT-PCR primers from each clade based on their functions were selected for expression analysis. For PCRs, 10 μl 2X SYBR Green qPCR Master Mix, 1 μl of 10 pmol each forward and reverse primers, and 6 μl nuclease-free water were used. The three-step program was run with 10 min: initial denaturation at 96°C, 40 cycles of 45 s; denaturation at 96°C, 45 s; annealing at 58°C, 45 s; and extension at 72°C, followed by melting at the default setting. The *Actin* gene was used as an internal control. The qRT-PCR analysis was carried out using a Rotor-Gene Q-6000 Real-Time PCR machine (QIAGEN). Three biological replicates were taken for expression profiling, and two technical replicates were used for analysis. The relative expression levels of the genes were calculated using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001).

RESULTS

Genome-Wide Identification of *VrLEA*, *VaLEA*, and *VuLEA* Genes

A total of 201 LEA genes were retrieved from three *Vigna* species, namely, mung bean, adzuki bean, and cowpea. A total of 55 non-redundant *LEA*-sequences from mung bean, 64 from adzuki bean, and 82 from cowpea genomes were identified and considered as putative LEA genes (**Table 1**). However, 96 *MtLEA*-genes were retrieved from the analysis of the *Medicago truncatula* genome. These sequences were classified into eight different groups, namely, LEA-1, LEA-2, LEA-3, LEA-4, LEA-5, LEA-6, SMP, and DHN. LEA genes accounted for 0.20% of all genes present in mung bean (*VrLEA*), 0.24% of genes present in adzuki bean (*VaLEA*), 0.29% of genes present in cowpea (*VuLEA*), and 0.30% of genes present in model legume *Medicago*. Among the eight classes of the LEA gene subfamily, *VrLEA-6* was absent in mung bean, whereas *VaLEA-5* and *VaLEA-6* were found absent in adzuki bean. The maximum proportion of LEA-2 was noticed in all the three *Vigna* species as 38 *VrLEA-2* genes were observed in mung bean, 45 *VaLEA-2* in adzuki bean, 55 *VuLEA-2* in cowpea, and 68 *VrLEA-2* genes in *Medicago*. The average *LEA* genes per Mb were found as 0.12 genes per Mb in mung bean, 0.15 in adzuki bean, 0.13 in cowpea, and 0.23 genes per Mb in *M. truncatula.*

Chromosomal Distribution of *VrLEA*, *VaLEA*, and *VuLEA* Genes

The study observed that 48 out of 55 *VrLEA*, 38 out of 62 *VaLEA*, and 80 out of 84 *VuLEA* genes were distributed across 11 chromosomes of mung bean, adzuki bean, and cowpea, respectively (**Figure 1**), whereas the rest of the *VrLEA*, *VaLEA*, *and VuLEA* genes were present in the scaffold (**Supplementary Tables 1–3**). Chromosome 10 (ch. 10) had a minimum number of two *VrLEA* genes, whereas ch. 8 had the maximum number of nine *VrLEA* genes in the mung bean genome. Four chromosomes, namely, 1, 2, 10, and 11 had only *VrLEA-2* genes, whereas the rest of the *VrLEA* genes were distributed over seven chromosomes. Similarly, in adzuki bean, ch. 5 comprised only one *VaLEA* gene, whereas ch. 4 had a maximum number of seven *VaLEA* genes. The ch. 1, 7, and 10 are comprised only *VaLEA* genes. In the cowpea genome, ch. 8 had a minimum of four *VuLEA* genes, whereas ch. 5 had a maximum of 11 *VuLEA* genes. Out of 11 chromosomes of cowpea, three chromosomes, namely, ch. 2, 8, and 11 consisted of only *VuLEA* genes while the rest of the 8 chromosomes comprised a mixture of different classes of *VuLEA* genes.

Phylogenetic Relationship of *VrLEA*, *VaLEA*, and *VuLEA* Genes

For the domain-based classification of LEA proteins, all the 201 LEA genes from three *Vigna* species identified in this study were considered for the construction of the phylogenetic tree (**Figure 2**). The LEA genes verified through Pfam for late embryogenesis-abundant domains were considered as putative *LEA* genes. All of the 201 LEA genes were clearly grouped into 15 major classes indicating their diversity. In mung bean, all the 55 LEA genes were grouped into eight clusters (**Figure 3**). Clade I comprised 27 *VrLEA*-2 genes, forming the largest cluster. Clade II consisted of 10 *VrLEA* genes, including 3 *VrLEA*-2 genes, 6 *VrLEA*-3 genes, and 1 *VrLEA*-4 gene. Clades III and IV were the single-cluster clades comprising 1 *VrLEA*-2 gene each. Clade V had three types of *LEA* genes, including 1 *VrLEA*-1 and *VrLEA*-4 gene each, and 4 *VrLEA*-2 genes. Clade VI had 4 *VrSMP* genes, and Clade VII had 2 *VrLEA*-2 genes. Similarly, Clade VIII had

[10]https://www.ncbi.nlm.nih.gov/tools/primer-blast/

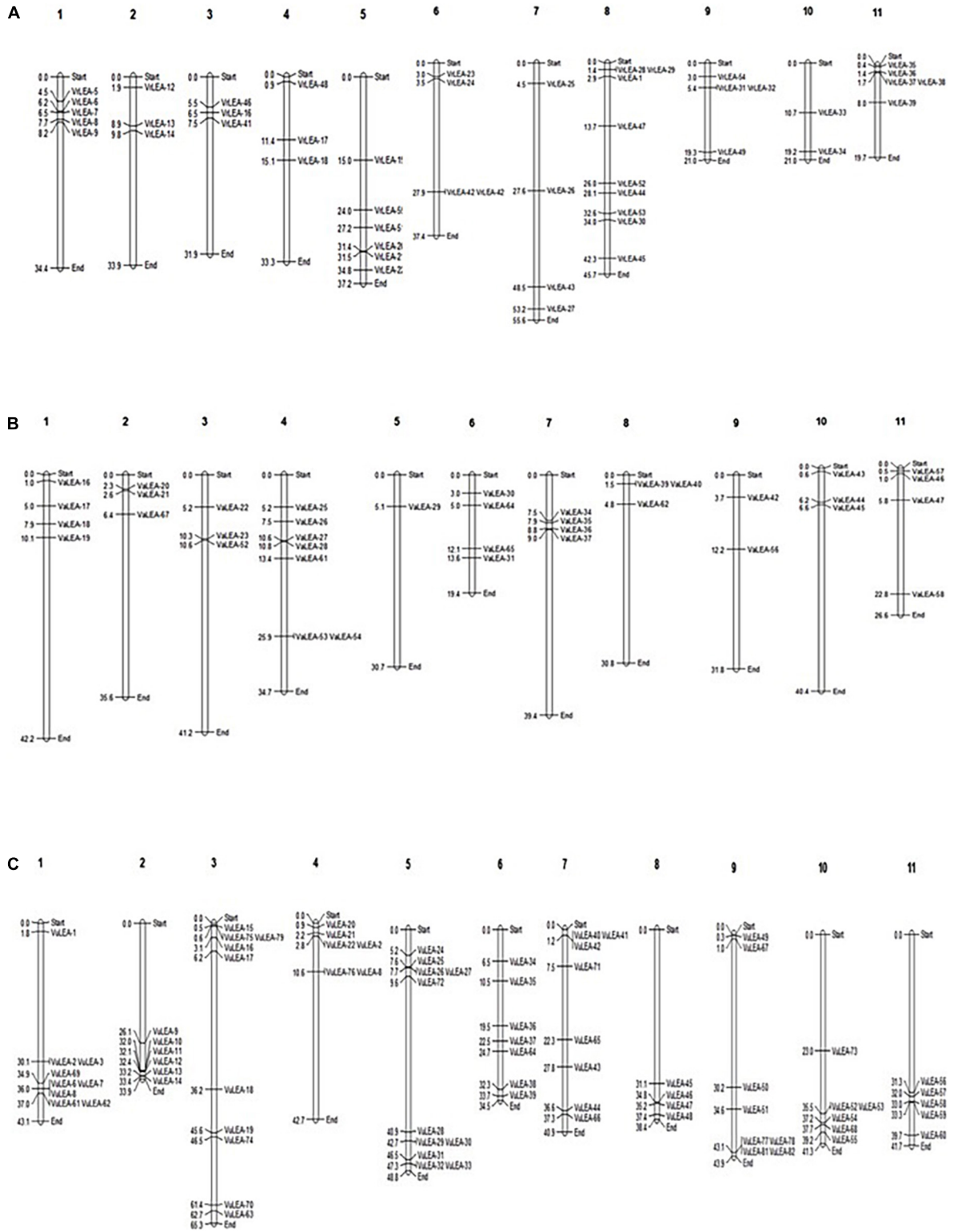


FIGURE 1 | (A) Chromosomal distribution of late embryogenesis abundant (*LEA*) genes on mung bean (*Vigna radiata*) genome. **(B)** Chromosomal distribution of *LEA* genes on adzuki bean (*Vigna angularis*) genome. **(C)** Chromosomal distribution of *LEA* genes on cowpea (*Vigna unguiculata*) genome.

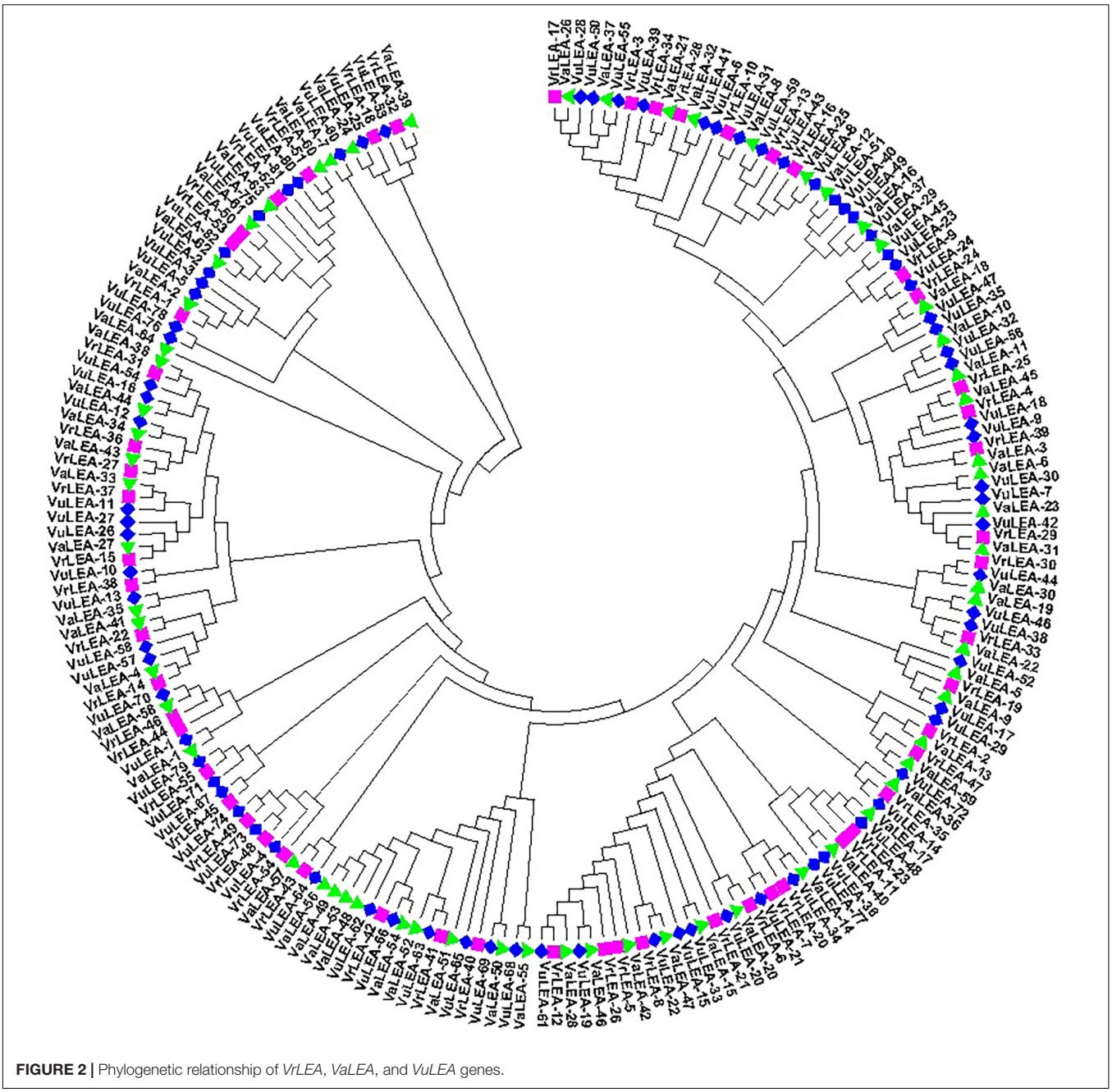


FIGURE 2 | Phylogenetic relationship of *VrLEA*, *VaLEA*, and *VuLEA* genes.

four *VrLEA* genes, including 2 *VrLEA*-5 and *VrLEA*-6 genes each. The LEA genes in adzuki bean clearly grouped into three major clades (**Figure 4**). Clade I comprised 44 *VaLEA*-2 genes and 1 *VaLEA*-3 gene. Clade II consisted of only one gene of *VrLEA*-1. Clade III had 2 *VaLEA-2*, 3 *VaDHN*, 9 *VaLEA*-3, 2 *VaLEA*-4, and SMP genes each. Similarly, all the 82 LEA genes from the cowpea genome were grouped into four clades (**Figure 5**). Clade I comprised 50 *VuLEA-2* genes, and 1 *VuLEA-3* and *VuSMP* gene each whereas Clade II consisted of five different types of *VuLEA* genes. Clade III had 4 *VuLEA-2* genes and 1 *VuLEA-3* gene. Six *VuLEA*-3 genes and 1 *VuLEA*-4 gene were grouped together in Clade IV. This phylogenetic analysis also provided information about diversification and evolutionary relationships among the LEA genes.

Characterization of *VrLEA*, *VaLEA*, and *VuLEA* Genes

The analyses about the length of LEA protein, genomic and CDS sequence, number of exons, molecular weight, and pI are presented in **Supplementary Tables 1**–**3**. The genomic length of *LEA* genes ranged from 282 bp (V*rLEA-40*) to 22,112 bp (*VrLEA-34*) in mung bean, 249 (*VaLEA-56*) to 30,173 bp (*VaLEA-21*) in adzuki bean, and 410 (*VuLEA-61*) to 4711 bp (V*uLEA-70*) in

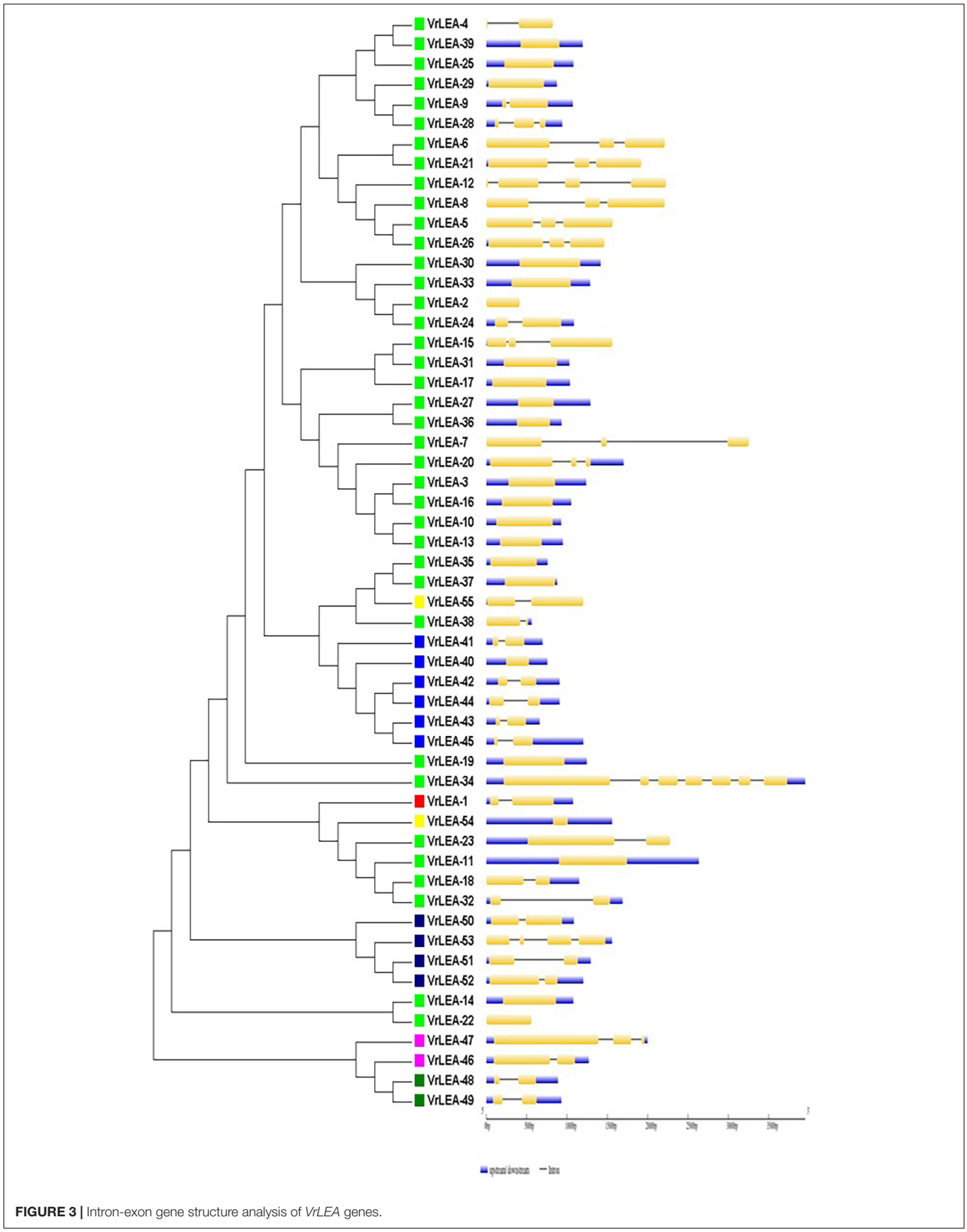


FIGURE 3 | Intron-exon gene structure analysis of *VrLEA* genes.

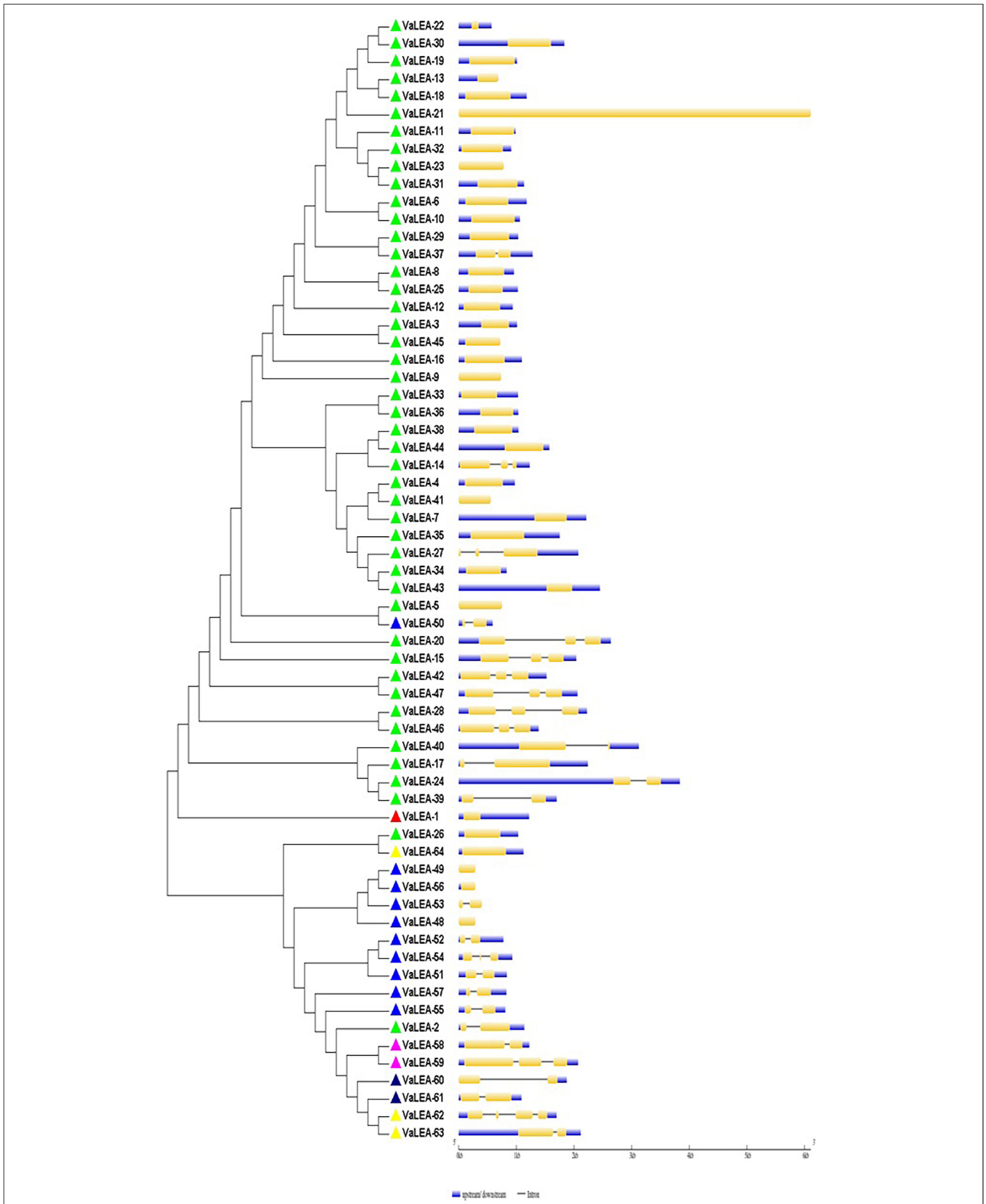


FIGURE 4 | Intron-exon gene structure analysis of *VaLEA* genes.

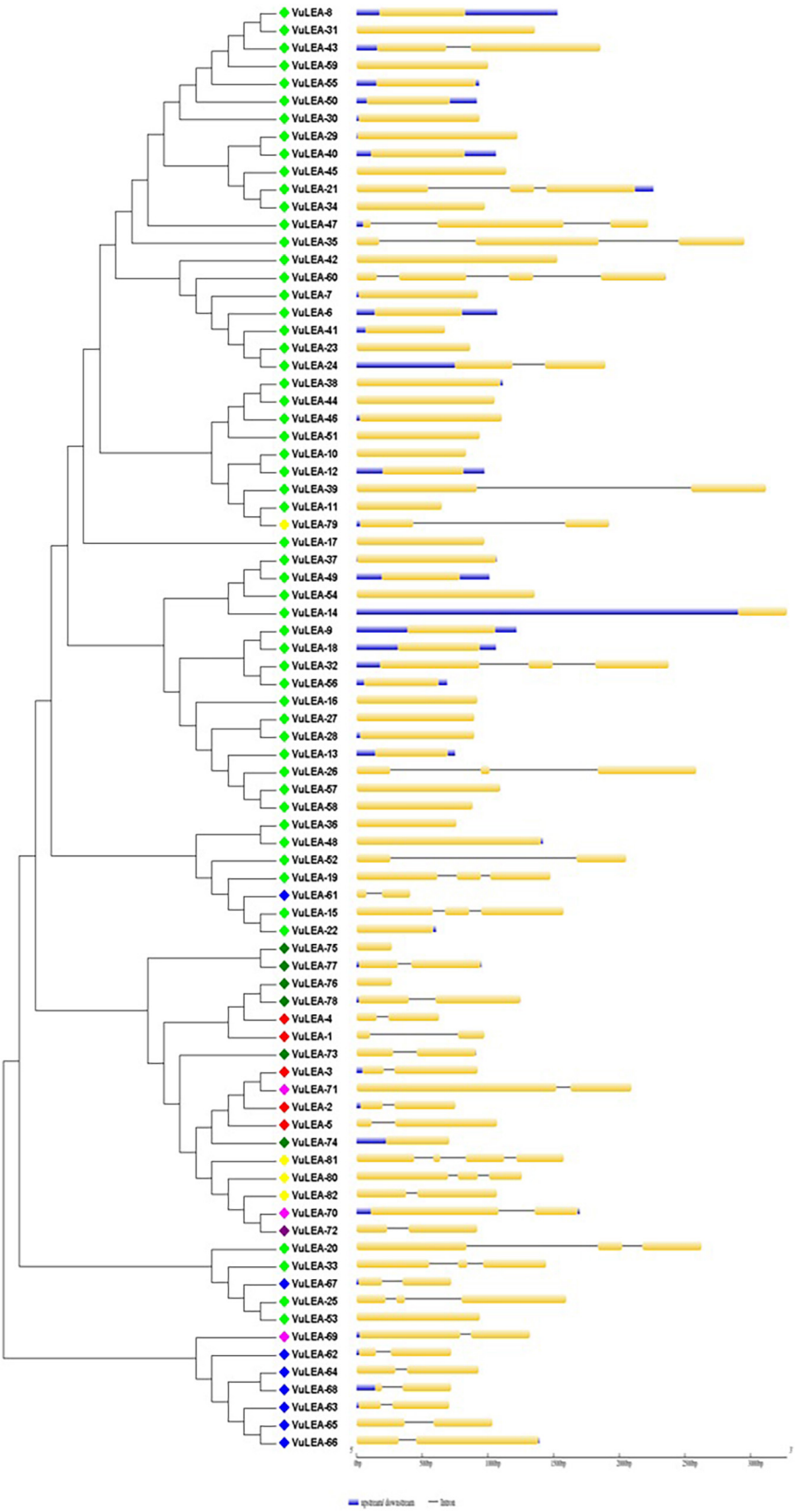


FIGURE 5 | Intron-exon gene structure analysis of *VuLEA* genes.

cowpea. Simultaneously, the length of coding sequences varied from 282 bp (*VrLEA-40*) to 7,113 bp (*VrLEA-34*) in mung bean, 249 (*VaLEA-56*) to 3,993 bp (*VaLEA-21*) in adzuki bean, and 270 (*VuLEA-76*) to 2,386 bp (*VuLEA-24*) in cowpea. The number of exons ranged from 1 to 7 in mung bean, whereas it varied from 1 to 4 in adzuki bean and cowpea, respectively. The LEA protein length varied from 94 (*VrLEA-34*) to 2,354 (*VrLEA-34*) in mung bean, 83 (*VrLEA-55*) to 1,050 (*VrLEA-64*) in adzuki bean, and 83 to 1,050 in cowpea. The predicted molecular weight of LEA proteins from mung bean, adzuki bean, and cowpea ranged from 10.22 kDa (*VrLEA-48*) to 263.50 kDa (*VrLEA-34*), 9.36 kDa (*VaLEA-55*) to 1063.50 kDa (*VaLEA-59*), and 21.11 kDa (*VuLEA-76*) to 199.46 kDa (*VuLEA-44*), respectively. Similarly, pI ranged from 4.47 (*VrLEA-50*) to 9.78 (*VrLEA-31*), 4.40 (*VaLEA-60*) to 11.02 (*VaLEA-54*), and 3.85 (*VuLEA-60*) to 4.37 (*VuLEA-76*) in the three *Vigna* species under study, indicating the structural diversity of LEA genes.

Gene Structure and Motif Analysis of *VrLEA*, *VaLEA*, and *VuLEA* Genes

The exon-intron organization of the 201 LEA genes from mung bean, adzuki bean, and cowpea was constructed using the coding and genomic sequences (**Figures 3–5**). The gene structures of LEA genes revealed that all the *LEA* genes exhibited variable exon-intron organization. It ranged from 1 to 7 exons in mung bean and 1 to 4 exons in both adzuki bean and cowpea. The maximum number of introns was found in mung bean as compared to adzuki bean and cowpea, indicating the diversification of the mung bean genome. A total of 10 distinct motifs were identified for LEA genes from mung bean, adzuki bean, and cowpea (**Figure 6**).

Expression Pattern of *VrLEA* Genes in Response to Heat Stress

All the 55 *VrLEA* candidate genes were used for primer designing, and subsequently, 55 primer pairs were developed. These primers were BLAST-searched against the *V. glabrescens* genome, and their annotation was performed (**Supplementary Table 4**). Seven representative candidates from each subclade were selected for the expression analysis (**Figure 7**). The expression of these seven genes (*VrLEA-1, VrLEA-2, VrLEA-40, VrLEA-47, VrLEA-48, VrLEA-54,* and *VrLEA-55*) in the HS (IPM 312-19) and HT (IC251372) genotypes was analyzed. The gene *VrLEA-1* exhibited a significant increase (about 15-fold) in HS genotypes on heat shock, whereas a decrease was noticed in the HT genotype. The *VrLEA-2* transcript significantly increased in HS genotype and decreased in HT genotype (IC 251372) at both the levels of heat stress (HSI-1 and HSI-2; 45°C for 3 and 6 h, respectively). The expression of *VrLEA-40* decreased in both the genotypes at HSI-1, whereas significantly increased in both the genotypes at HSI-2. The *VrLEA-47* exhibited a significant decrease in HS genotype, whereas it increased significantly to about 2.5-fold at HSI-2. The pattern of expression of *VrLEA-48* was similar to that of *VrLEA-47*, although the level of expression was higher than *VrLEA-47*. The *VrLEA-54* transcript was downregulated in the HS genotype after the heat sock, whereas it was upregulated significantly to about 2-fold in the HT genotype after the heat-shock induction. Similarly, about a 5-fold increase was noticed in HT genotypes at HSI-2 as compared to control.

In the second experiment, the 7-day-old seedlings of HS and HT genotypes were subjected to 24-h heat-shock induction and they exhibited differential expression of the genes (**Figure 8**). The gene *VrLEA-1* exhibited a similar pattern of expression as in the case of the detached leaf method although it was comparatively lower. The expression of three genes, namely, *VrLEA-2, VrLEA-40,* and *VrLEA-47*, was downregulated in the HS genotypes after 24 h of heat stress, whereas it was significantly upregulated in the HT genotype. In contrast, *VrLEA-48* and *VrLEA-54* exhibited upregulation in both HS and HT genotypes, although the level was higher in the HS genotype. Similarly, *VrLEA-55* exhibited upregulation in both HS and HT genotypes with a higher expression in the HT genotype as compared to the HS genotype.

DISCUSSION

Abiotic stresses encompass extreme environmental factors which alter a plant's biochemical and physiological processes constraining its productivity (Douglas et al., 2020). Abiotic stresses, either individually or in combination, hinder the crops from reaching their actual genetic yield potential. Among the abiotic stresses, heat stress is one of the most important and frequent constraints that limit the productivity of crop plants, especially during the current scenario of climate change, and is, therefore, a concern for crop cultivation worldwide. The crop yield of mung bean is highly affected by the fluctuating temperatures as the progression and the phenological development of the plant are mainly driven by temperature (Carberry, 2007; Chauhan et al., 2010). While the optimum temperatures for mung bean are 28–30°C (Kaur et al., 2015), the key temperatures are 7.5°C (baseline), 30°C (optimum), and 40°C (maximum). Most of the *Vigna* crops are very frequently subjected to extreme environmental conditions that are many times at the upper extremes as far as temperature is concerned. However, a very limited study has been conducted till present on the aspect of high-temperature tolerance in mung bean (Kumar et al., 2016, 2017). The linkage of high-throughput phenotyping platforms for the screening of target traits with omics interventions can give a clue for its mitigation (Pratap et al., 2019). Besides, the genome-wide analysis of gene families and transcriptional factors helps in designing crops for improved heat tolerance in mung bean.

The LEA gene family is large and widely diversified across the plant kingdom (Gao and Lan, 2016). It is an important gene family that responds to plant developmental activities, biotic and abiotic stresses (Hincha and Thalhammer, 2012), and hormonal regulations (Liu et al., 2019). This gene family has been characterized in many crops, such as *Arabidopsis* (Hundertmark and Hincha, 2008), wheat (Liu et al., 2019), tea (Jin et al., 2019), potato (Chen et al., 2019), and poplar (Cheng et al., 2021). Besides, this gene family was also identified in fungi and bacteria (Garay-Arroyo et al., 2000; Gal et al., 2004; Gao et al., 2016). In contrast, the genome-wide identification and

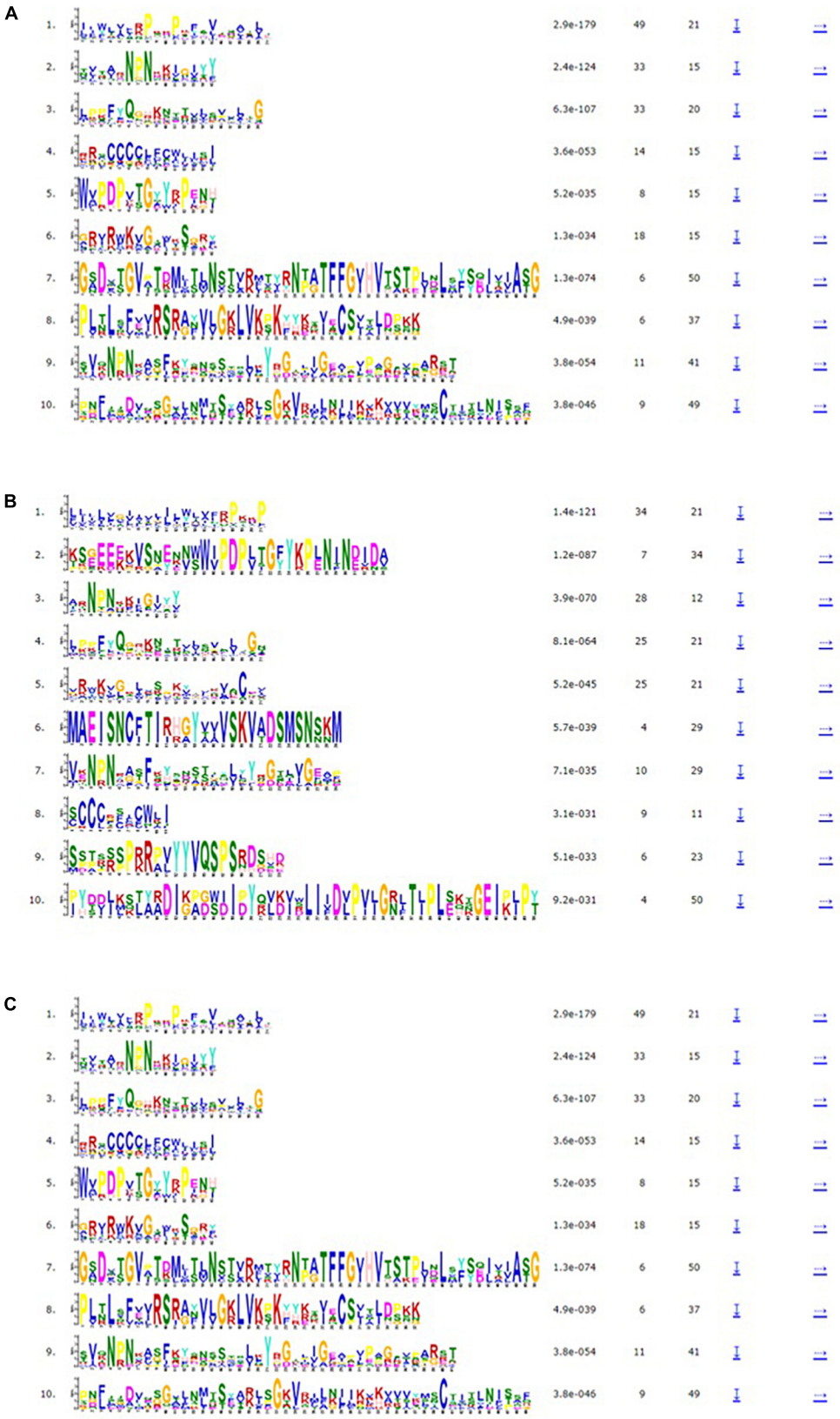


FIGURE 6 | Motif analysis of **(A)** *VrLEA*, **(B)** *VaLEA*, and **(C)** *VuLEA* genes.

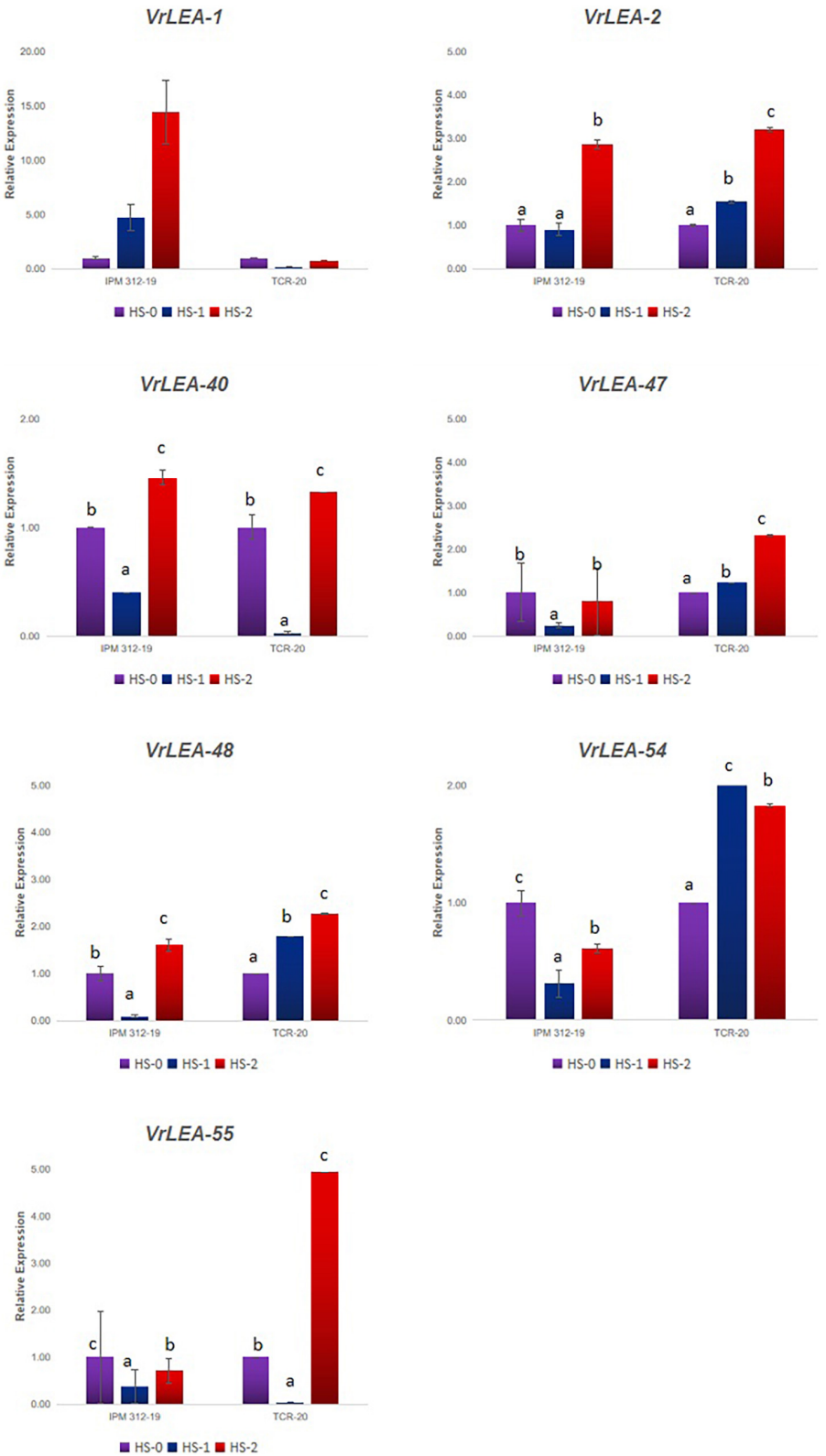


FIGURE 7 | qRT-PCR analysis of selected *VrLEA* candidate genes on HS-mungbean (IPM 312-19) and HR-*V. glabrescens* (IC251372) under heat shock stress through detached leaf method. The statistical test was performed separately in both genotypes. The same letters show non-significant differences, whereas different letters show significant differences.

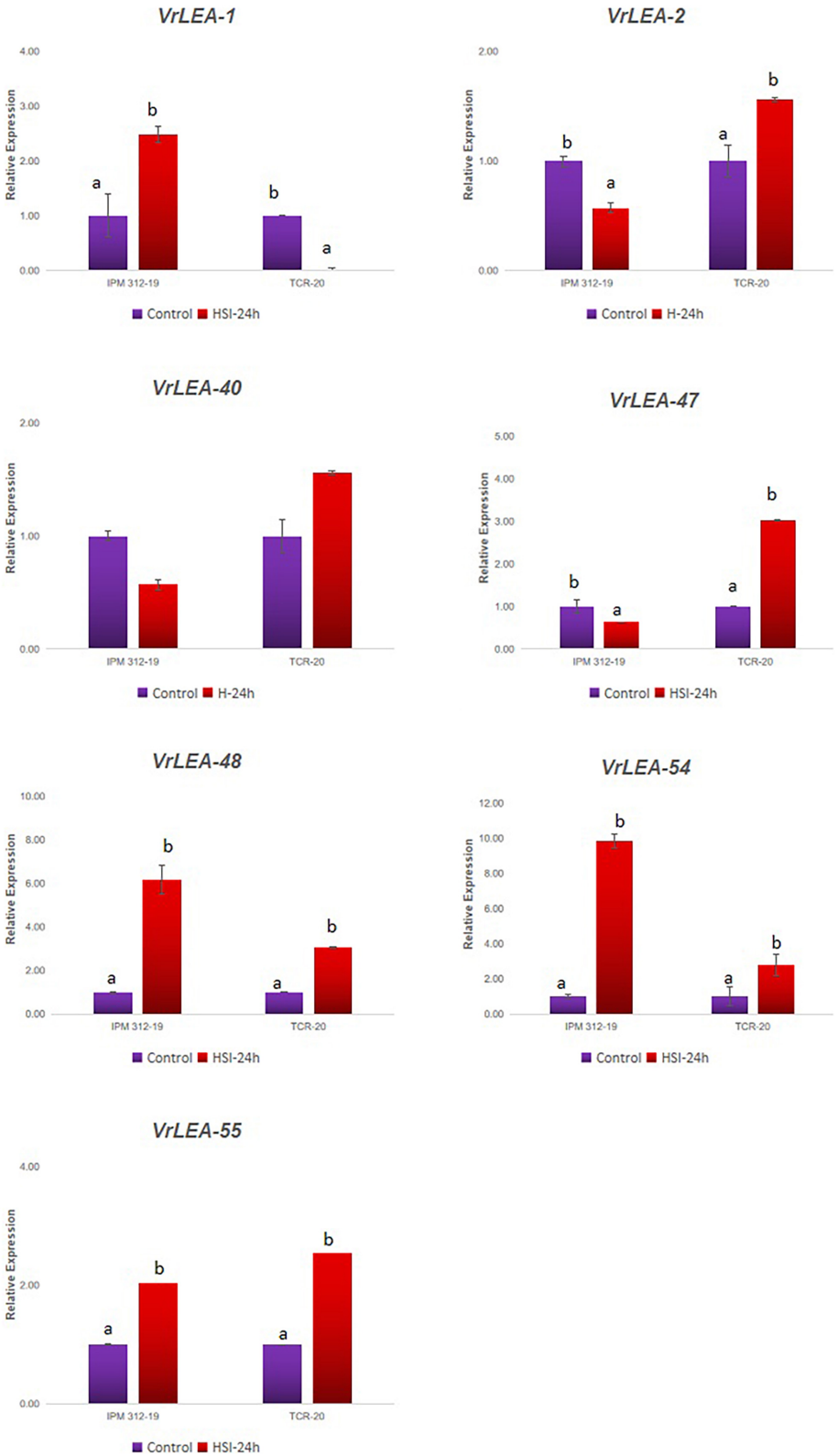


FIGURE 8 | qRT-PCR analysis of selected *VrLEA* candidate genes on HS-mungbean (IPM 312-19) and HR-*V. glabrescens* (IC251372) under heat shock stress at seedling stage. The statistical test was performed separately in both genotypes. The same letters show non-significant differences, whereas different letters show significant differences.

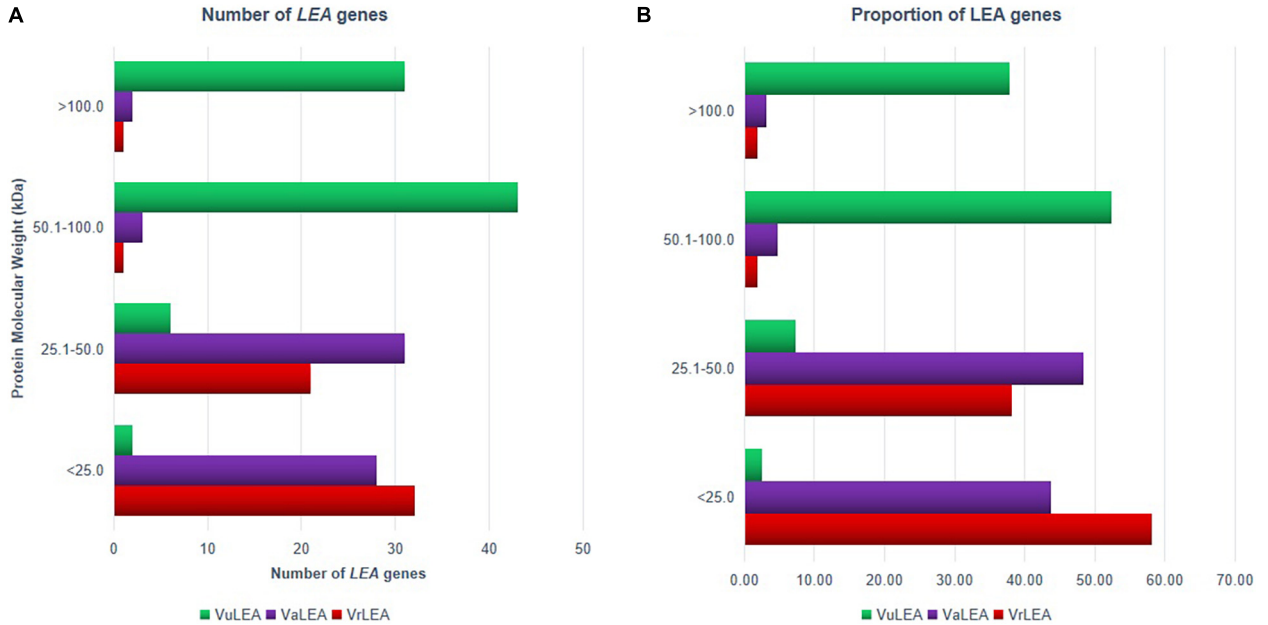


FIGURE 9 | (A) Number and **(B)** proportion of LEA protein in *VrLEA*, *VaLEA*, and *VuLEA* gene families based on their molecular weight.

characterization of the LEA gene family in mung bean, adzuki bean, and cowpea have not yet been performed systematically. In this study, a total of 201 LEA genes were identified, in which 55, 64, and 82 *LEA* genes were found in mung bean, adzuki bean, and cowpea genomes, respectively. The analysis revealed that these genes belong to six to eight classes of the LEA gene family. Based on the conserved domain and phylogenetic tree analyses, 55 *VrLEA* genes were categorized in seven distinct groups, 64 *VaLEA* genes in six groups, and 82 *VuLEA* genes in eight groups. The cowpea genome had all the eight types of LEA subfamilies whereas the adzuki bean genome lacked two subfamilies of *VaLEA* genes (i.e., *VaLEA*-5 and *VaLEA*-6), and the mung bean genome lacked the *VrLEA*-6 subfamily. It was indicated that LEA-1 to LEA-4, SMP, and DHN were highly conserved. The LEA-6 group has also been found to be absent in a few higher plants, such as *Dendrobium officinale* (Ling et al., 2016) and *Solanum lycopersicum* (Cao and Li, 2015). This finding indicated that the variation exists in the LEA protein gene family groups in some of the plant species. In the three *Vigna* species studied, the maximum proportion of the LEA gene family was accounted for LEA-2 as compared to other LEA groups. A similar type of contribution of LEA-2 has been observed earlier by Jin et al. (2019) in tea. In contrast, a similar proportion of the LEA-2 group was not observed earlier in many plant species, such as *Arabidopsis thaliana* (Hundertmark and Hincha, 2008), *Cleistogenes songorica* (Muvunyi et al., 2018), *Manihot esculenta* (Wu et al., 2018), *Pinus tabuliformis* (Gao et al., 2016), and *Populus trichocarpa* (Lan et al., 2013). This difference may be attributed to the improvement in the annotation of genomes of plants species. The differences were observed in all the groupings of *VrLEA, VaLEA, and VuLEA* genes in the phylogenetic analysis, indicating their diversification and differential genetic pattern. This finding also indicated the variation and complexity of the LEA gene family in plants. The gene structure prediction of LEA genes revealed that most of the LEA genes contained only a few or no introns, and all encoded the conserved LEA domain.

Based on the physiochemical properties of LEA genes, it was noticed that most of the *LEA* genes (58.18%) encode relatively small proteins (<25 kDa) in the mung bean genome (**Figure 9A**), whereas only 3.60% of them are large proteins (>50 kDa) (**Figure 9B**). Similarly, about 43.75% of genes encode small proteins in the adzuki bean genome and 7.81% of them encode larger proteins. On the contrary, only 2.44% of genes are <25 kDa in the cowpea genome, whereas most of the genes (about 52.44%) had a high molecular weight (50.1–100 kDa). This result was in conformity with the findings of Chen et al. (2019) in *Solanum tuberosum* (94.6%) and those of Liang et al. (2016) in *B. napus* (90.7%). Therefore, while this report is in conformity with the earlier reports in the case of mung bean and adzuki bean, it is contrasting in the case of the cowpea genome. The analysis also revealed that each LEA group contained conserved motifs that have been previously reported in other plant species (Liang et al., 2016; Magwanga et al., 2018). This was indicated that *LEA* genes may encode functional LEA proteins that have specific functions within the group, and the members of the same protein within the group may have originated from gene expansion. The members of the different groups may be due to the evolution from different ancestors.

Most of the LEA proteins are predicted to have no stable secondary structure in solution, but they may acquire α-helical structures on dehydration (Bremer et al., 2017), which allows them to conformational change according to the changes in their microenvironment, resulting in multiple functions in abiotic stresses (Hanin et al., 2011). Previous studies have shown that

LEA genes play an important role in low-temperature, drought, and high-salinity stresses (Jin et al., 2019). The expression of LEA proteins is often induced by abiotic stresses, such as drought, temperature, salt, and exogenous application of hormones, at different development stages and tissues of plants (Chen et al., 2019), and an increased expression of the *LEA* genes can improve stress tolerance. The overexpression of the wheat *DHN-5* was reported to enhance osmotic stress tolerance in *Arabidopsis* (Brini et al., 2011). Ding et al. (2021) reported the maximum expression of *LEA* genes under polyethylene glycol (PEG) treatment, low temperature, and light intensity deviations in the case of rye. Therefore, the above-mentioned studies demonstrate that the LEA genes could potentially be used to improve the abiotic stress tolerance of crops. The expression analysis of HS and HT genotypes in this study indicated that the genes *VrLEA-1, VrLEA-47, VrLEA-48, VrLEA-54*, and *VrLEA-55* play an important role toward heat stress tolerance in the HT-wild *Vigna*, accession IC 251372. Similarly, *LEA* genes, such as *VrLEA-2, VrLEA-40, VrLEA-47*, and *VrLEA-55,* were significantly overexpressed in the HT genotype, indicating the basis of heat tolerance in IC251372. The *VrLEA-55* (DHN protein) gene showed a 5-fold increase in the expression in HT genotype in Experiment-1 (detached leaf method), whereas about two-fold increase was noticed at the seedling stage under 24 h of heat induction in the intact seedling (*in planta* heat stress treatment). The differential expression of the *VrLEA-55* gene in both treatments could be attributed to mechanical injury in addition to heat stress in the detached leaf method as compared to the *in planta* method. Ding et al. (2021) also noticed the expression of this gene in different tissues, including seeds of rye, in drought stress although the expression was non-significant. Our study reports the expression of this candidate gene in the leaves of the HT-wild *Vigna* genotype and provides a new insight to heat tolerance. Therefore, the results of this study laid the foundation for further investigating the functional characterization of LEA proteins and their potential use in the genetic improvement of mung bean for heat tolerance.

CONCLUSION

A total of 201 LEA genes were identified in three *Vigna* species which were further grouped into seven distinct clades based on their phylogenetic relationships. The LEA-6 group was found absent in the mung bean genome, whereas all the seven groups were present in the cowpea genome. The gene structure exhibited diversification of candidates in the *Vigna* genome. Additionally, the expression profiling of seven *VrLEA* genes in leaves under four heat stress treatments revealed that these genes had an important role in response to heat stress. These results, therefore, provide valuable information for the future functional studies of *LEA* genes toward the genetic improvement of mung bean for heat stress tolerance.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

CMS, MK, and AP conceived the idea and planned the work. SS, AT, AM, and HK retrieved the data from the database. CMS and AP analyzed the data. CMS drafted the manuscript. AP, RN, and NS edited the manuscript. All authors read the manuscript and approved it for publication.

FUNDING

This study was financially supported by the Ministry of Agriculture, Agriculture Education and Research, Government of Uttar Pradesh, under the project "Center of Excellence in Dryland Agriculture (CEDA/2018-23)" and "ACIAR Funded International Mungbean Improvement Network (Grant No. CIM/2014/079)."

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.843107/full#supplementary-material

REFERENCES

Ahuja, I., de-Vos, R. C. H., Bones, A. M., and Hall, R. D. (2010). Plant molecular stress responses face climate change. *Trends Plant Sci.* 15, 664–674. doi: 10.1016/j.tplants.2010.08.002

Allito, B. B., Nana, E. M., and Alemneh, A. A. (2015). Rhizobia strain and legume genome interaction effects on nitrogen fixation and yield of grain legume: a review. *Mol. Soil Biol.* 2, 1–6. doi: 10.5376/msb.2015.06.0002

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., et al. (2009). MEME Suite: tools for motif discovery and searching. *Nucleic Acids Res.* 37, W202–W208. doi: 10.1093/nar/gkp335

Bangar, P., Chaudhury, A., Tiwari, B., Kumar, S., Kumari, R., and Bhat, K. V. (2019). Morphophysiological and biochemical response of mungbean [*Vigna radiata* (L.) Wilczek] varieties at different developmental stages under drought stress. *Turk. J Biol.* 43, 58–69. doi: 10.3906/biy-1801-64

Barnabas, B., Jäger, K., and Fehér, A. (2008). The effect of drought and heat stress on reproductive processes in cereals. *Plant Cell Environ.* 31, 11–38. doi: 10.1111/j.1365-3040.2007.01727.x

Bita, C. E., and Gerats, T. (2013). Plant tolerance to high temperature in a changing environment: scientific fundamentals and production of heat stress tolerant crops. *Front. Plant Sci.* 4:273. doi: 10.3389/fpls.2013.00273

Bremer, A., Wolff, M., Thalhammer, A., and Hincha, D. K. (2017). Folding of intrinsically disordered plant LEA proteins is driven by glycerol-induced crowding and the presence of membranes. *FEBS J.* 284, 919–936. doi: 10.1111/febs.14023

Brini, F., Yamamoto, A., Jlaiel, L., Takeda, S., Hobo, T., Dinh, H. Q., et al. (2011). Pleiotropic effects of the wheat Dehydrin DHN-5 on stress responses in *Arabidopsis*. *Plant Cell Physiol.* 52, 676–688. doi: 10.1093/pcp/pcr030

Cao, J., and Li, X. (2015). Identification and phylogenetic analysis of late embryogenesis abundant proteins family in tomato (*Solanum*

lycopersicum). *Planta* 241, 757–772. doi: 10.1007/s00425-014-2215-y

Carberry, P. (2007). "Crop development models," in *Encyclopedia of Science Encyclopedia of Water Science*, 2nd Edn. (Boca Raton, FL: CRC Press), 122–124.

Chauhan, Y. S., Douglas, C. A., Rachaputi, R. C. N., Agius, P., Martin, W. D., King, K., et al. (2010). "Physiology of mungbean and development of the mungbean crop model," in *Proceedings of the 1st Australian Summer Grains Conference*, Gold Coast, QLD, 21–24.

Chen, Y., Li, C., Zhang, B., Yi, J., Yang, Y., Kong, C., et al. (2019). The role of the late embryogenesis-abundant (LEA) protein family in development and the abiotic stress response: a comprehensive expression analysis of potato (*Solanum tuberosum*). *Genes* 10:148. doi: 10.3390/genes10020148

Cheng, Z., Zhang, X., Yao, W., Zhao, K., Liu, L., Fan, G., et al. (2021). Genome-wide search and structural and functional analyses for late embryogenesis-abundant (LEA) gene family in poplar. *BMC Plant Biol.* 21:110. doi: 10.1186/s12870-021-02872-3

Debnath, M., Pandey, M., and Bisen, P. S. (2011). An omics approach to understand the plant abiotic stress. *OMICS* 15, 739–762. doi: 10.1089/omi.2010.0146

Ding, M., Wang, L., Zhan, W., Sun, G., Jia, X., Chen, S., et al. (2021). Genome-wide identification and expression analysis of late embryogenesis abundant protein-encoding genes in rye (*Secale cereale* L.). *PLoS One* 16:e0249757. doi: 10.1371/journal.pone.0249757

Douglas, C., Pratap, A., Hanumantharao, B., Manu, B., Dubey, S., Singh, P., et al. (2020). "Breeding progress and future challenges: abiotic stresses," in *The Mungbean Genome, Compendium of Plant Genomes*, eds R. M. Nair, R. Schafleitner, and S. H. Lee (Cham: Springer). doi: 10.1007/978-3-030-20008-4_6

Dure, L. R., Greenway, S. C., and Galau, G. A. (1981). Developmental biochemistry of cottonseed embryogenesis and germination: changing messenger ribonucleic acid populations as shown by in vitro and in vivo protein synthesis. *Biochemistry-US* 20, 4162–4168. doi: 10.1021/bi00517a033

Finkelstein, R. R. (1993). Abscisic acid-insensitive mutations provide evidence for stage-specific signal pathways regulating expression of an *Arabidopsis* late embryogenesis-abundant (lea) gene. *Mol. Gen. Genet.* 238, 401–408. doi: 10.1007/BF00291999

Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., et al. (2010). The Pfam protein families database. *Nucleic Acids Res.* 38, 11–22. doi: 10.1093/nar/gkp985

Gal, T. Z., Glazer, I., and Koltai, H. (2004). An LEA group 3 family member is involved in survival of C. *elegans* during exposure to stress. *FEBS Lett.* 577, 21–26. doi: 10.1016/j.febslet.2004.09.049

Gao, J., and Lan, T. (2016). Functional characterization of the late embryogenesis abundant (LEA) protein gene family from *Pinus tabuliformis* (Pinaceae) in *Escherichia Coli*. *Nat. Publ. Gr. Nat. Publ. Group.* 6, 1–10. doi: 10.1038/srep19467

Gao, W., Huang, X. L., and Tao, Y. (2016). A critical review of NanoSIMS in analysis of microbial metabolic activities at single cell level. *Crit. Rev. Biotechnol.* 36, 884–890. doi: 10.3109/07388551.2015.1057550

Garay-Arroyo, A., Colmenero-Flores, J. M., Garciarrubio, A., and Covarrubias, A. A. (2000). Highly hydrophilic proteins in prokaryotes and eukaryotes are common during conditions of water deficit. *J. Biol. Chem.* 275, 5668–5674. doi: 10.1074/jbc.275.8.5668

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., and Bairoch, A. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 31, 3784–3788. doi: 10.1093/nar/gkg563

Guindon, S., and Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52, 696–704.

Hand, S. C., Menze, M. A., Toner, M., Boswell, L., and Moore, D. (2011). LEA proteins during water stress: not just for plants anymore. *Annu. Rev. Physiol.* 73, 115–134. doi: 10.1146/annurev-physiol-012110-142203

Hanin, M., Brini, F., Ebel, C., Toda, Y., Takeda, S., and Masmoudi, K. (2011). Plant dehydrins and stress tolerance: versatile proteins for complex mechanisms. *Plant Signal. Behav.* 6, 1503–1509. doi: 10.4161/psb.6.10.17088

HanumanthaRao, B., Nair, R. M., and Nayyar, H. (2016). Salinity and high temperature tolerance in mungbean from a physiological perspective. *Front. Plant Sci.* 7:957. doi: 10.3389/fpls.2016.00957

Hincha, D. K., and Thalhammer, A. (2012). LEA proteins: IDPs with versatile functions in cellular dehydration tolerance. *Biochem. Soc.* 40, 1000–1003. doi: 10.1042/BST20120109

Hirayama, T., and Shinozaki, K. (2010). Research on plant abiotic stress responses in the post-genome era: past, present and future. *Plant J.* 61, 1041–1052. doi: 10.1111/j.1365-313X.2010.04124.x

Hu, B., Jin, J., Guo, A. Y., Zhang, H., Luo, J., and Gao, G. (2015). GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics* 31, 1296–1297. doi: 10.1093/bioinformatics/btu817

Hunault, G., and Jaspard, E. (2010). LEAPdb: a database for the late embryogenesis abundant proteins. *BMC Genomics* 11:221. doi: 10.1186/1471-2164-11-221

Hundertmark, M., and Hincha, D. K. (2008). LEA. (Late embryogenesis abundant) proteins and their encoding genes in *Arabidopsis thaliana*. *BMC Genomics* 9:118. doi: 10.1186/1471-2164-9-118

İbrahime, M., Kibar, U., Kazan, K., Ozmen, C. Y., Mutaf, F., Asci, S. D., et al. (2019). Genome-wide identification of the LEA protein gene family in grapevine (*Vitis vinifera* L.). *Tree Genet. Genomes* 15:55. doi: 10.1007/s11295-019-1364-3

Jin, X., Cao, D., Wang, Z., Ma, L., Tian, K., Liu, Y., et al. (2019). Genome-wide identification and expression analyses of the LEA protein gene family in tea plant reveal their involvement in seed development and abiotic stress responses. *Sci. Rep.* 9:14123. doi: 10.1038/s41598-019-50645-8

Kang, J. Y., Kim, S. K., Kim, M. Y., Lestari, P., Kim, K. H., Ha, B.-K., et al. (2014). Genome sequence of mungbean and insights into evolution within *Vigna* species. *Nat. Commun.* 5:5443. doi: 10.1038/ncomms6443

Kang, Y. J., Satyawan, D., Shim, S., Lee, T., Lee, J., Hwang, W. J., et al. (2015). Draft genome sequence of adzuki bean, *Vigna angularis*. *Sci. Rep.* 5:8069. doi: 10.1038/srep08069

Kaur, R., Bains, T. S., Bindumadhava, H., and Nair, H. (2015). Responses of mungbean genotypes to heat stress: effects on reproductive biology, leaf function and yield traits. *Sci. Hortic.* 197, 527–541. doi: 10.1016/j.scienta.2015.10.015

Krueger, C., Berkowitz, O., Stephan, U. W., and Hell, R. (2002). A metal-binding member of the late embryogenesis abundant protein family transports iron in the phloem of *Ricinus communis* L. *J. Biol. Chem.* 277, 25062–25069. doi: 10.1074/jbc.M201896200

Kumar, C., Mishra, S. B., Nilanjaya, and Singh, C. M. (2017). Investigating morpho-physiological traits and agrometeorological indices in green gram for terminal heat tolerance. *Legume Res.* 40, 271–276.

Kumar, M., Busch, W., Birke, H., Kemmerling, B., Nurnberger, T., and Schoff, F. (2009). Heat shock factors HsfB1 and HsfB2are involved in the regulation of Pdf1.2 expression and pathogen resistance in *Arabodopsis*. *Mol. Plant* 2, 152–165. doi: 10.1093/mp/ssn095

Kumar, C., Mishra, S. B., Nilanjay, and Singh, C. M. (2016). Evaluation of selection indices for improving terminal heat tolerance in greengram (*Vignaradiata* L. Wilczek). *J. Agrometeorol.* 18, 216–221.

Kumar, R., Singh, C. M., Arya, M., Kumar, R., Mishra, S. B., Singh, U. K., et al. (2020). Investigating stress indices to discriminate the physiologically efficient heat tolerant genotypes of mungbean [*Vigna radiata* (L.) Wilczek]. *Legume Res.* 43, 43–49.

Kumari, P., and Varma, S. K. (1983). Genotypic differences in flower production/shedding and yield in mungbean (*Vigna radiata*). *Indian J. Plant Physiol.* 26, 402–405.

Lan, T., Gao, J., and Zeng, Q. Y. (2013). Genome-wide analysis of the LEA (late embryogenesis abundant) protein gene family in *Populus trichocarpa*. *Tree Genet. Genomes* 9, 253–264.

Li, X., and Cao, J. (2016). Late embryogenesis abundant (LEA) gene family in maize: identification, evolution, and expression profiles. *Plant Mol. Biol. Rep.* 34, 15–28. doi: 10.1007/s11105-015-0901-y

Liang, Y., Xiong, Z., Zheng, J., Xu, D., Zhu, Z., Xiang, J., et al. (2016). Genome-wide identification, structural analysis and new insights into late embryogenesis abundant (LEA) gene family formation pattern in *Brassica napus*. *Sci. Rep.* 6:24265. doi: 10.1038/srep24265

Ling, H., Zeng, X., and Guo, S. X. (2016). Functional insights into the late embryogenesis abundant (LEA) protein family from *Dendrobium ofcinale* (Orchidaceae) using an *Escherichia coli* system. *Sci. Rep.* 6:39693. doi: 10.1038/srep39693

Liu, H., Xing, M., Yang, W., Mu, X., Wang, X., Lu, F., et al. (2019). Genome-wide identification of and functional insights into the late embryogenesis abundant

(LEA) gene family in bread wheat (*Triticum aestivum*). *Sci. Rep.* 9:13375. doi: 10.1038/s41598-019-49759-w

Liu, P. L., Huang, Y., Shi, P. H., Yu, M., Xie, J. B., and Xie, L. (2018). Duplication and diversification of lectin receptor-like kinases (LecRLK) genes in soybean. *Sci. Rep.* 8:5861. doi: 10.1038/s41598-018-24266-6

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2-\Delta\Delta CT$ method. *Methods* 25, 402–408.

Lonardi, S., Munoz-Amatria, M., Liang, Q., Shu, S., Wanamaker, S. I., Lo, S., et al. (2019). The genome of cowpea (*Vigna unguiculata* (L.) Walp.). *Plant J.* 98, 767–782.

Magwanga, R. O., Lu, P., Kirungu, J. N., Lu, H., Wang, X., Cai, X., et al. (2018). Characterization of the late embryogenesis abundant (LEA) proteins family and their role in drought stress tolerance in upland cotton. *BMC Genet.* 19:6. doi: 10.1186/s12863-017-0596-1

Manivannan, N., and Mahalingam, A. (2021). "Cowpea," in *The Beans and the Peas: From Orphan to Mainstream Crops*, eds A. Pratap and S. Gupta (Dufford: Wood Head Publishing Elsevier), 241–254.

Muvunyi, B. P., Yan, Q., Wu, F., Min, X., Yan, Z. Z., Kanzana, G., et al. (2018). Mining late embryogenesis abundant (LEA) family genes in *Cleistogenes songorica*, a xerophyte perennial desert plant. *Int. J. Mol. Sci.* 19:3430. doi: 10.3390/ijms19113430

Pratap, A., Basu, P. S., Gupta, S., Malviya, N., Rajan, N., Tomar, R., et al. (2014). Identification and characterization of sources for photo- and thermo-insensitivity in *Vigna* species. *Plant Breed.* 133, 756–764. doi: 10.1111/pbr.12215

Pratap, A., Douglas, C., Prajapati, U., Kumari, G., War, A. R., Tomar, R., et al. (2020). "Breeding progress and future challenges: biotic stresses," in *The Mungbean Genome, Compendium of Plant Genomes*, eds R. Nair, R. Schafleitner, and S. H. Lee (Cham: Springer), 55–80. doi: 10.1007/978-3-030-20008-4_5

Pratap, A., Gupta, S., Rathore, M., Basavraj, T. S., Singh, C. M., Prajapati, U., et al. (2021). "Mungbean," in *The Beans and the Peas: From Orphan to Mainstream Crops*, eds A. Pratap and S. Gupta (Dufford: Wood Head Publishing Elsevier), 1–32.

Pratap, A., Gupta, S., Basu, P. S., Tomar, R., Dubey, S., Rathore, M., et al. (2019). "Towards development of climate smart mungbean: challenges and opportunities," in *Genomic Designing of Climate Smart Pulse Crops*, ed. C. Kole (Cham: Springer), 235–264.

Rahman, L. N., Smith, G. S. T., Bamm, V. V., Voyer-Grant, J. A. M., Moffatt, B. A., Dutcher, J. R., et al. (2011). Phosphorylation of *Thellungiella salsuginea* dehydrins TsDHN-1 and TsDHN-2 facilitates cation-induced conformational changes and actin assembly. *Biochemistry-US* 50, 9587–9604. doi: 10.1021/bi201205m

Rosales, R., Romero, I., Escribano, M. I., Merodio, C., and Sanchez-Ballesta, M. T. (2014). The crucial role of Phi- and K-segments in the in vitro functionality of *Vitis vinifera* dehydrin DHN1a. *Phytochemistry* 108, 17–25.

Sehrawat, N., Yadav, M., Bhat, K. V., Sairam, R. K., and Jaiwal, P. K. (2015). Effect of salinity stress on mungbean [*Vigna radiata* (L.) Wilczek] during consecutive summer and spring seasons. *J. Agric. Sci.* 60, 23–32.

Sharma, L., Priya, M., Bindumadhava, H., Nair, R. M., and Nayyar, H. (2016). Influence of high temperature stress on growth, phenology and yield performance of mungbean (*Vigna radiata* (L.) Wilczek) under managed growth conditions. *Sci. Hortic.* 213, 379–391.

Singh, A. K., Velmurugan, A., Gupta, D. S., Kumar, J., Kesari, R., Konda, A., et al. (2019b). Draft genome sequence of a less-known wild *Vigna*: beach pea (*V. marina* cv. ANBp-14-03). *Crop J.* 7, 660–666.

Singh, C. M., Mishra, S. B., Pandey, A., and Arya, M. (2014). Eberhart-Russell' and AMMI approaches of genotype by environment interaction (GEI) for yield and yield component traits in *Vigna radiata* (L.) Wilczek. *Int. J. Agric. Environ. Biotech.* 7, 272–292.

Singh, C. M., Singh, P., Pratap, A., Pandey, R., Purwar, S., Vibha, et al. (2019a). Breeding for enhancing Legumovirus resistance in mungbean: current understanding and future directions. *Agronomy* 9:622. doi: 10.3390/agronomy9100622

Singh, C. M., Singh, P., Tiwari, C., Purwar, S., Kumar, M., Pratap, A., et al. (2021). Improving drought tolerance in Mungbean (*Vigna radiata* L. Wilczek): morpho-physiological, biochemical and molecular Perspectives. *Agronomy* 11:1534. doi: 10.3390/agronomy11081534

Singh, D. P., Singh, B. B., and Pratap, A. (2017). Genetic improvement of Mungbean and Urdbean and their role in enhancing pulse production in India. *Indian J. Genet.* 76, 550–567.

Souframanien, J., Raizada, A., Dhanasekar, P., and Suprasanna, P. (2020). Draft genome sequence of the pulse crop blackgram [*Vigna mungo* (L.) Hepper] reveals potential R-genes. *bioRxiv* [preprint]. doi: 10.1101/2020.06.21.163923

Takahashi, Y., Sakai, H., Yoshitsu, Y., Muto, C., Anai, T., Pandiyan, M., et al. (2019). Domesticating *Vigna stipulacea*: a potential legume crop with broad resistance to biotic stresses. *Front. Plant Sci.* 10:1607. doi: 10.3389/fpls.2019.01607

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis, version 6.0. *Mol. Biol. Evol.* 30, 2725–2729.

Thompson, J. D., Gibson, T., and Higgins, D. G. (2002). Multiple sequence alignment using ClustalW and ClustalX. *Curr. Protoc. Bioinformatics* Chapter 2:Unit2.3. doi: 10.1002/0471250953.bi0203s00

Tolleter, D., Hincha, D. K., and Macherel, D. (2010). A mitochondrial late embryogenesis abundant protein stabilizes model membranes in the dry state. *Biochim. Biophys. Acta Biomembranes* 10:1926. doi: 10.1016/j.bbamem.2010.06.029

Tzudir, L., Bera, P. S., and Chakraborty, P. K. (2014). Impact of temperature on the reproductive development in mungbean (*Vigna radiata*) varieties under different dates of sowing. *Int. J. Bioresour. Stress Manage.* 5, 194–199. doi: 10.5958/0976-4038.2014.00555.7

Voorrips, R. E. (2002). Mapchart: sofware for the graphical presentation of linkage maps and QTLs. *J. Hered.* 93, 77–78. doi: 10.1093/jhered/93.1.77

Wang, X.-S., Zhu, H.-B., Jin, G.-L., Liu, H.-L., Wu, W.-R., and Zhu, J. (2007). Genome-scale identification and analysis of LEA genes in rice (*Oryza sativa* L.). *Plant Sci.* 172, 414–420. doi: 10.1016/j.plantsci.2006.10.004

Wassmann, R., Jagadish, S., Sumfleth, K., Pathak, H., Howell, G., and Ismail, A. (2009). Regional vulnerability of climate change impacts on Asian rice production and scope for adaptation. *Adv. Agron.* 102, 91–133. doi: 10.1016/S0065-2113(09)01003-7

Wu, C., Hu, W., Yan, Y., Tie, W., Ding, Z., Guo, J., et al. (2018). The late embryogenesis abundant protein family in cassava (*Manihot esculenta* Crantz): genome-wide characterization and expression during abiotic stress. *Molecules* 23:1196. doi: 10.3390/molecules23051196

Xiong, L., and Zhu, J. K. (2002). Molecular and genetic aspects of plant responses to osmotic stress. *Plant Cell Environ.* 25, 131–139. doi: 10.1046/j.1365-3040.2002.00782.x

Yu, X., Yue, W., Yang, Q., Zhang, Y., Han, X., Yang, F., et al. (2019). Identification of the LEA family members from *Caragana korshinskii* (Fabaceae) and functional characterization of CkLEA2–3 in response to abiotic stress in *Arabidopsis*. *Braz. J. Bot.* 42, 227–238.

Zhu, J. K. (2016). Abiotic stress signaling and responses in plants. *Cell* 67, 313–324.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 23 March 2022
doi: 10.3389/fpls.2022.845314

Genome-Wide Analysis of the Soybean TIFY Family and Identification of *GmTIFY10e* and *GmTIFY10g* Response to Salt Stress

Ya-Li Liu[1,2†], Lei Zheng[2†], Long-Guo Jin[2†], Yuan-Xia Liu[1], Ya-Nan Kong[1], Yi-Xuan Wang[1], Tai-Fei Yu[2], Jun Chen[2], Yong-Bin Zhou[2], Ming Chen[2], Feng-Zhi Wang[3], You-Zhi Ma[2], Zhao-Shi Xu[2] and Jin-Hao Lan[1]**

[1]College of Agronomy, Qingdao Agricultural University, Qingdao, China, [2]Institute of Crop Science, Chinese Academy of Agricultural Sciences (CAAS)/National Key Facility for Crop Gene Resources and Genetic Improvement, Key Laboratory of Biology and Genetic Improvement of Triticeae Crops, Ministry of Agriculture, Beijing, China, [3]Hebei Key Laboratory of Crop Salt-Alkali Stress Tolerance Evaluation and Genetic Improvement/Cangzhou Academy of Agriculture and Forestry Sciences, Cangzhou, China

OPEN ACCESS

Edited by:
Rohini Garg,
Shiv Nadar University, India

Reviewed by:
Bing-kai Hou,
Shandong University, China
Satyabrata Nanda,
Centurion University of Technology and Management, India

***Correspondence:**
Zhao-Shi Xu
xuzhaoshi@caas.cn
Jin-Hao Lan
jinhao2005@qau.edu.cn

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

Received: 29 December 2021
Accepted: 23 February 2022
Published: 23 March 2022

Citation:
Liu Y-L, Zheng L, Jin L-G, Liu Y-X, Kong Y-N, Wang Y-X, Yu T-F, Chen J, Zhou Y-B, Chen M, Wang F-Z, Ma Y-Z, Xu Z-S and Lan J-H (2022) Genome-Wide Analysis of the Soybean TIFY Family and Identification of GmTIFY10e and GmTIFY10g Response to Salt Stress. Front. Plant Sci. 13:845314. doi: 10.3389/fpls.2022.845314

TIFY proteins play crucial roles in plant abiotic and biotic stress responses. Our transcriptome data revealed several *TIFY* family genes with significantly upregulated expression under drought, salt, and ABA treatments. However, the functions of the *GmTIFY* family genes are still unknown in abiotic stresses. We identified 38 *GmTIFY* genes and found that *TIFY10* homologous genes have the most duplication events, higher selection pressure, and more obvious response to abiotic stresses compared with other homologous genes. Expression pattern analysis showed that *GmTIFY10e* and *GmTIFY10g* genes were significantly induced by salt stress. Under salt stress, *GmTIFY10e* and *GmTIFY10g* transgenic *Arabidopsis* plants showed higher root lengths and fresh weights and had significantly better growth than the wild type (WT). In addition, overexpression of *GmTIFY10e* and *GmTIFY10g* genes in soybean improved salt tolerance by increasing the PRO, POD, and CAT contents and decreasing the MDA content; on the contrary, RNA interference plants showed sensitivity to salt stress. Overexpression of *GmTIFY10e* and *GmTIFY10g* in *Arabidopsis* and soybean could improve the salt tolerance of plants, while the RNAi of *GmTIFY10e* and *GmTIFY10g* significantly increased sensitivity to salt stress in soybean. Further analysis demonstrated that *GmTIFY10e* and *GmTIFY10g* genes changed the expression levels of genes related to the ABA signal pathway, including *GmSnRK2*, *GmPP2C*, *GmMYC2*, *GmCAT1*, and *GmPOD*. This study provides a basis for comprehensive analysis of the role of soybean *TIFY* genes in stress response in the future.

Keywords: soybean, TIFY, salt tolerance, ABA, transcription factor

INTRODUCTION

Environmental stresses affect both growth and yield in soybean (Bohnert et al., 1995). To adapt to environmental stresses, several regulatory pathways gradually formed during the evolution of plants (Zhu et al., 2011). In previous studies, TIFY proteins were found to respond to abiotic and biotic stresses through regulatory pathways (Thines et al., 2007; Ebel et al., 2018). Studying TIFY proteins were useful for protecting soybean (*Glycine max*) growth and yield under various environmental stresses.

TIFY proteins were defined with conservative amino acid (aa) sequence (TIF[F/Y] XG) (Vanholme et al., 2007). The TIFY family genes were divided into four subfamilies, including TIFY, Jasmonate ZIM domain (JAZ), PEAPOD (PPD), and ZIM-like (ZML) according to their specific domains (Bai et al., 2011). TIFY subfamily members contain only one TIFY domain; JAZ subfamily members have a C-terminal Jas ($SLX_2FX_2KRX_2RX_5PY$) domain (also named CCT_2 domain) in addition to the TIFY domain (Staswick, 2008); ZML subfamily members contain the CCT domain (CONSTANS, CO-like, and TOC1) and the GATA zinc finger domain ($CX_2CX_{20}CX_2C$), except for the TIFY domain (Nishii et al., 2000); PPD subfamily members contain the N-terminal PPD domain, TIFY domain and the C-terminal Jas domain without PY motif ($SLX_2FX_2KRX_2RX_5$) (Chung et al., 2009).

In the early research, TIFY proteins could respond to biotic stresses, such as insects and pathogens by jasmonic acid (JA) signaling pathway (Thines et al., 2007; Barah and Bones, 2015; Thireault et al., 2015; Mao et al., 2017; Dhakarey et al., 2018). Recent studies have demonstrated that TIFY proteins play an important role in regulating plants resistance to abiotic stresses (Demianski et al., 2012; Zhu et al., 2013; Fu et al., 2017; Sun et al., 2017; Peethambaran et al., 2018; Meng et al., 2019; Luo et al., 2020; Zhao et al., 2020a). The *Arabidopsis AtTIFY10a* and *AtTIFY10b* genes and their wild soybean homologous genes *GsTIFY10a, GsTIFY10b,* and *GsJAZ2* positively regulated the response to salt and alkali stresses (Zhu et al., 2013; Zhao et al., 2020a). Overexpression of *GhJAZ2* in cotton plants can significantly enhance sensitivity to salt stress (Sun et al., 2017). During the seedling and reproductive stages of rice, overexpression of *OsJAZ1* in rice can improve sensitivity to drought stress, while *JAZ1* t-DNA inserted in mutant plants had higher drought resistance than wild type (WT) plants (Fu et al., 2017). The rice *OsJAZ8* gene was confirmed to improve the salt tolerance of transgenic tobacco through the JA signaling pathway (Peethambaran et al., 2018). The hard wheat *TdTIFY11a* gene can improve salt tolerance when overexpressed in *Arabidopsis* (Ebel et al., 2018). *Arabidopsis AtJAZ7* gene was identified to mediate drought tolerance through comparative proteomics and metabolomics analysis (Meng et al., 2019). Cotton *GbJAZ1* gene was confirmed to interact with ABA-insensitive1 (*ABI1*) and involved in regulation the tolerance of salt and drought through the ABA signaling pathway (Luo et al., 2020).

Soybean is one of the most important commercial crops worldwide and an important source of vegetable protein and oil for humans. Salt stress is an important factor which could affect the growth and yield of soybean (Zhu et al., 2013). Studying salt stress-related genes and their functions are of great significance to soybean molecular breeding. After analyzing the transcriptome data in previous studies, we found that the expression levels of many *TIFY* family genes were significantly upregulated under drought, salt, and ABA treatments (Shi et al., 2018). In our study, we performed a genome-wide identification of the *TIFY* family genes in soybean and identified 38 *GmTIFY* genes. We analyzed the structure characteristics, expression patterns, duplication events, and physical and chemical properties of *GmTIFY* family genes. During transcriptome data analysis, we found six significantly upregulated genes under salt treatment, which were all *GmTIFY10* and *GmTIFY11* homologous genes in the JAZ subfamily. The gene function analysis of *GmTIFY10e* and *GmTIFY10g* showed that they have a positive regulatory effect on salt stress tolerance in *Arabidopsis* and soybean. Further analysis demonstrated that overexpression of *GmTIFY10e* and *GmTIFY10g* could influence the expression levels of ABA-related genes, which suggested that *GmTIFY10e* and *GmTIFY10g* may regulate the salt tolerance in plants by participating in ABA signaling pathway.

MATERIALS AND METHODS

Screening and Identification of *TIFY* Genes

The nucleic acid and protein databases of *Arabidopsis*, rice, soybean, apple, and grape were downloaded from the Ensemble Plants database.[1] The hidden Markov model (HMM) of TIFY domain (PF06200) was obtained from Pfam.[2] We then used the hmm-search program HMMER3.1 (Prince and Pickett, 2002; Xia et al., 2017) to identify the TIFY HMM for the TIFY proteins in the resulting protein databases. The 18 *Arabidopsis* AtTIFY protein sequences were obtained from TAIR[3] and used to search the TIFY proteins from rice, soybean, apple, and grape protein databases by the BLASTp program of basic local alignment search tool (BLAST; Wang et al., 2017). We compared the results of the two methods to confirm *TIFY* candidate genes in these species. These candidate genes were identified in their domains with SMART[4] and CDD[5] to ensure that the TIFY domain was in sequence (Letunic et al., 2002; Marchler-Bauer et al., 2002). Finally, the ExPASy[6] ProtParam tool was used to query the physical and chemical properties of the *GmTIFYs* (Appel et al., 1994; Wang et al., 2020a).

Phylogenetic Tree Analysis of TIFY Proteins

The TIFY protein sequences of *Arabidopsis*, rice, soybean, apple, and grape were compared using ClustalW in the MEGA-X software. The maximum likelihood (ML) method was used to construct a phylogenetic tree for analyzing the phylogenetic relationship between *TIFYs* (Kumar et al., 2016; Leng et al., 2021). The bootstrap method was used with 1,000 replicates. The methods and parameters were the same as the Jones–Taylor–Thornton (JTT) model, gamma-distributed rates (G), and the gamma parameter 1.

Chromosomal Location, Gene Duplication, and Selective Pressure Analysis

The position information of the *GmTIFY* family genes was extracted from the GFF3 file of the soybean genome. The location and distribution of *GmTIFY* family genes were visualized on the chromosomes using Map Gene 2 Chromosomal (Jiangtao et al., 2015; Wang et al., 2020a).[7]

For gene duplication analysis, the TBtools software was used to identify the duplication events of the soybean genome and *GmTIFY* genes. The collinearity pairs of *GmTIFY* genes were

[1]http://plants.ensembl.org/index.html
[2]http://pfam.Xfam.org/
[3]http://www.arabidopsis.org/
[4]http://smart.embl-heidelberg.de/
[5]https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi
[6]https://www.expasy.org/
[7]http://mg2c.iask.in/mg2c_v2.0/

extracted and used to visualize a synteny map with the CIRCOS software (Lestari et al., 2013).

The TIFY coding sequences were aligned using ClustalW software. The alignment results were converted to PAML format using EasyCodeML and a tree file in Newick format was built using MEGA-X. The selection pressure was estimated using the branch model of EasyCodeML. The ratio of non-synonymous to synonymous substitution rates (ω) was determined by the free-ratio model and the two-ratio model among the branches of the TIFY tree file (Gao et al., 2019).

Gene Structure, Motif, and Promoter Sequence Analysis

The motif information of the soybean GmTIFY proteins was analyzed using the MEME online tool (Bailey et al., 2009; Wang et al., 2019).[8] The resulting files and soybean gene structure annotation files were imported into TBtools for visualization.

The 2,000 bp promoter sequences were submitted to the PlantCARE website[9] to analyze the *cis*-acting elements of its family members (Guo et al., 2007; Su et al., 2020; Wang et al., 2020a). The resulting file was imported into GSDS[10] for visualization.

Expression Patterns of *GmTIFY* Genes

The RNA-seq data of GmTIFY family members in different tissues and organs were downloaded from the Phytozome database.[11] The transcriptome data of several abiotic stresses were obtained from previous studies (NCBI SRA accession: PRJNA694374; Shi et al., 2018). TBtools software was used to visualize the expression levels of GmTIFYs.

Plant Materials, Stress Treatments, and Real-Time Fluorescence Quantitative PCR

The soybean variety Zhonghuang39 was used for this study. The soybeans were planted in a greenhouse in a mixture of humus and vermiculite (humus: vermiculite = 1:1). Seven-day-old soybean seedlings were treated with 10% PEG6000 and 250 mM NaCl, respectively. The samples were collected at 0, 0.5, 1, 2, 4, 8, 12, and 24 h after treatments (Li et al., 2017; Zhang et al., 2019).

An RNA plant extraction kit (Zhuangmeng, Beijing, China) was used to extract total RNA from soybean leaves and *Transcript*R All-in-One First-Strand cDNA Synthesis SuperMix (TransGen Biotech, Beijing, China) was used for reverse transcription. The primers designed by Primer Premier 5.0 software were listed in **Supplementary Table 1**. The eukaryotic elongation factor 1-β (*GmELF1b*) was used as the internal control (Jian et al., 2008). An Applied Biosystems 7500 Real-Time PCR System was used to perform RT-qPCR. The $2^{-\Delta\Delta CT}$ method was used to analyze the quantitative results analysis (Udvardi, 2008PC). Each experiment was performed with three biological replicates.

Subcellular Localization Assay

Pectinase and cellulase were used to lyse fresh *Arabidopsis* leaves and obtain *Arabidopsis* protoplasts. The gene coding regions were cloned into the 16318hGFP expression vector. The fusion expression vector *GmTIFY10e*-hGFP and *GmTIFY10g*-hGFP were transformed into *Arabidopsis* protoplasts mediated by PEG4000, respectively (He et al., 2016). After 18 h of incubation at 22°C in the dark, the GFP fluorescence signal was detected using a laser confocal microscope (Zeiss LSM 700, Germany; Riechmann et al., 2000).

Obtaining Transgenic *Arabidopsis* and Salt Stress Treatment

Arabidopsis (Col-0) seeds were sterilized with 75% alcohol for 15 min. Sterilized *Arabidopsis* seeds were sprinkled on ½ MS medium and maintained at 4°C for 4 days, after which they were moved to a growth incubator at 22°C under a 16 h light and 8 h dark cycle. When the seedlings grew to four leaves, they were transferred to a mixture of humus and vermiculite for subsequent experiments (Riechmann et al., 2000; Du et al., 2018).

The coding regions of the *GmTIFY* genes were subcloned into the pCAMBIA1302 vector. The constructed pCAMBIA1302-*GmTIFY10e* and pCAMBIA1302-*GmTIFY10g* were transformed into *Arabidopsis* using the floral dip method, respectively (Clough and Bent, 1998). Positive lines were selected on ½ MS medium plates containing hygromycin (35 mg/L) and were further verified using PCR. The same method was used until transgenic three generation (T_3). The expression levels of transgenic lines were determined by RT-qPCR and three homozygous T_3 lines with the highest expression levels were used for the subsequent phenotypic analysis (Li et al., 2017).

For the experiment of root growth, 5-day-old seedings were transferred to MS medium and MS medium with 125 mM NaCl for another 7 days, after which the lengths of primary root and fresh weights were measured (Wang et al., 2019). For salt treatment, 5-day-old seedlings were transferred to the soil, and then, 21-day-old seedlings were treated with 250 mM NaCl for 14 days (Wang et al., 2019). All experiments contained three independent replicates.

Obtaining Soybean Hairy Roots by *Agrobacterium rhizogenes*-Mediated (*A. rhizogenes*-Mediated)

To obtain the overexpression vector of *GmTIFY* genes, the coding regions of *GmTIFY* genes were ligated with the pCAMBIA3301 vector to obtain recombinant plasmids (Kereszt et al., 2007; Zhao et al., 2017). To obtain the RNA interference expression vector, a 546 bp interference fragment consisting of a 200 bp target fragment and its antisense sequence connected by 146 bp zeol dehydrogenase gene sequence was synthesized and inserted into pCAMBIA3301 (Wang et al., 2019).

The constructed overexpression vector, interference expression vector, and empty pCAMBIA3301 vector were transferred to the *Bacillus rhizobacillus* (*B. rhizobacillus*) strain K599, and the recombinant vector was transferred to the hypocotyl of soybean *via* the *A. rhizogenes*-mediated method (Wang et al., 2015). The injected plants were cultured under high humidity conditions in a greenhouse until hairy roots grew at the infected site. When the hairy root reached about 5 cm long, the hypocotyl was removed below 0.5–1 cm of the infection site. At the same time, the seedlings were transplanted in mixed soil and cultured in the greenhouse for 7 days (Wang et al., 2020a). The positive soybean plants were subjected to salt stress test.

[8] http://meme-suite.org/
[9] http://bioinformatics.psb.ugent.be/webtools/plantcare/html/
[10] http://gsds.cbi.pku.edu.cn/
[11] https://phytozome-next.jgi.doe.gov/pz/portal.html

Measurement of Physiological Indicators and Nitroblue Tetrazolium Staining

The leaves and roots of plants were used to determine physiological indicators under salt stress. A Physiological Index Test Kit (Cominbio, Suzhou, China) was used to test the contents of malondialdehyde (MDA), proline (PRO), catalase (CAT), and peroxidase (POD) in leaves and roots (Shi et al., 2018). The whole leaves and roots were soaked in nitroblue tetrazolium (NBT) for overnight staining. After staining, the samples were soaked in a decolorizing solution (30% glycerol and 70% ethanol) to decolorize the sample until it turned white (Du et al., 2018). All experiments were performed in three biological replicates.

Enrichment Analysis of Co-expression Genes

The co-expression genes were obtained from the Phytozome database. The enrichment analysis was performed by Database for Annotation, Visualization and Integrated Discovery (DAVID) online tools.[12] R software was used to visualize the results of the enrichment analysis.

[12]https://david.ncifcrf.gov/

Statistical Analysis

One-way ANOVA test analysis was performed in Microsoft Excel 2007. Data were shown as means ± standard deviation (SD), with a *p*-value cutoff of 0.05 and 0.01. The method was used to analyze the RT-qPCR results and physiological indicators.

RESULTS

Screening and Identification of *GmTIFY* Genes

The BLASTp program and hmmsearch program were used to search for *TIFY* genes in databases of *Arabidopsis*, rice, soybean, apple, and grape. We then compared the results of two programs and identified the TIFY domain using SMART and CDD to confirm TIFY members in five species. Finally, 38 *GmTIFY* genes were identified in the soybean genome. The number of *TIFY* genes in *Arabidopsis* (18), rice (20), grape (19), and apple (30) are consistent with previous reports (Ye et al., 2009; Bai et al., 2011; Li et al., 2014). The *GmTIFY* genes were named according to their relationship with *Arabidopsis* and their location on the chromosomes (**Table 1**).

TABLE 1 | Details of the 38 soybean *TIFY* genes.

Gene name	Gene ID	ORF (aa)	MW (kD)	Chromosome	p*I*
GmTIFY1a	Glyma_02G215600	310	33.34	2	6.08
GmTIFY1b	Glyma_04G096200	324	35.51	4	5.68
GmTIFY1c	Glyma_06G097900	304	33.44	6	5.70
GmTIFY1d	Glyma_14G182800	307	33.28	14	6.23
GmTIFY2a	Glyma_04G096300	350	37.87	4	4.63
GmTIFY2b	Glyma_06G098000	351	38.08	6	4.67
GmTIFY2c	Glyma_07G182700	355	39.59	7	4.91
GmTIFY2d	Glyma_08G067200	358	39.90	8	5.24
GmTIFY2e	Glyma_08G221800	334	36.69	8	6.56
GmTIFY3a	Glyma_09G077500	206	22.11	9	6.90
GmTIFY3b	Glyma_13G116100	207	22.71	13	9.89
GmTIFY3c	Glyma_15G184900	201	21.48	15	6.43
GmTIFY3d	Glyma_17G043700	197	21.47	17	9.79
GmTIFY4a	Glyma_10G244400	346	38.10	10	8.89
GmTIFY4b	Glyma_20G150000	350	38.31	20	8.89
GmTIFY5a	Glyma_05G141200	134	15.54	5	9.30
GmTIFY5b	Glyma_08G096500	150	17.34	8	9.81
GmTIFY5c	Glyma_13G219100	138	15.91	13	9.10
GmTIFY5d	Glyma_15G093100	133	15.20	15	8.68
GmTIFY6	Glyma_09G174200	386	41.54	9	9.44
GmTIFY7a	Glyma_05G235500	371	39.01	5	8.76
GmTIFY7b	Glyma_08G043000	369	39.39	8	8.98
GmTIFY7c	Glyma_09G123600	319	33.64	9	9.37
GmTIFY8a	Glyma_04G071400	415	43.69	4	7.28
GmTIFY8b	Glyma_06G072700	413	43.50	6	9.39
GmTIFY8c	Glyma_17G205200	379	40.53	17	6.83
GmTIFY8d	Glyma_08G264700	370	40.99	8	5.59
GmTIFY8e	Glyma_16G081800	392	43.29	16	5.75
GmTIFY10a	Glyma_01G204400	195	21.67	1	8.99
GmTIFY10b	Glyma_04G013800	201	21.93	4	8.75
GmTIFY10c	Glyma_06G013700	160	17.75	6	9.56
GmTIFY10d	Glyma_09G071600	258	27.83	9	8.83
GmTIFY10e	Glyma_11G038600	203	23.03	11	9.01
GmTIFY10f	Glyma_13G112000	242	26.29	13	7.74
GmTIFY10g	Glyma_15G179600	258	27.65	15	9.03
GmTIFY10h	Glyma_17G047700	242	26.36	17	9.00
GmTIFY11a	Glyma_07G041400	232	25.09	7	9.18
GmTIFY11b	Glyma_16G010000	230	24.90	16	9.00

The protein lengths, molecular weights (MV), and isoelectric points (p*I*) are provided in **Table 1**. In 38 *GmTIFY* genes, the coding sequences range from 133 (*GmTIFY5d*) to 415 aa (*GmTIFY8a*); the MW ranges from 15.20 (*GmTIFY5d*) to 43.69 kD (*GmTIFY8a*) and the p*I* ranges from 4.63 to 9.89.

Phylogenetic Analysis of *GmTIFY* Genes

To investigate the phylogenetic relationships of GmTIFYs, we constructed a phylogenetic tree using *TIFY* genes from *Arabidopsis*, rice, soybean, grape, and apple (**Figure 1**). The *TIFY* genes were divided into four subfamilies of TIFY, JAZ, ZIM, and PPD according to specific domains (Zhang et al., 2012). The JAZ subfamily contains the largest number of *TIFY* genes including 12 *AtTIFY* genes, 11 *VvTIFY* genes, 22 *GmTIFY* genes, 22 *MdTIFY* genes, and 16 *OsTIFY* genes. The ZIM subfamily contains three *AtTIFY* genes, five *VvTIFY* genes, nine *GmTIFY* genes, two *MdTIFY* genes, and two *OsTIFY* genes. The PPD and TIFY subfamilies have no genes in monocotyledon plants which is consistent with the previous research results (Zhang et al., 2012; Li et al., 2014). There are two *AtTIFY* genes, two *VvTIFY* genes, two *GmTIFY* genes, and two *MdTIFY* genes in PPD subfamily and one *AtTIFY* genes, one *VvTIFY* genes, five *GmTIFY* genes, and four *MdTIFY* genes in TIFY subfamily. Further analysis revealed that the number of *TIFY* genes in soybeans and apples was almost twice than that of *Arabidopsis* and grapes. This may result from more events of chromosomes doubling, fusion and rearrangement occurred in soybean and apple evolution. Interestingly, we found the expression levels of *GmTIFYs* were significantly upregulated in our previous transcriptome data of drought, salt, and ABA treatments which were all *TIFY10* and *TIFY11* homologous genes. The transgenic *Arabidopsis* lines of *AtTIFY10a* and *AtTIFY10b* showed higher salt tolerance than WT plants (Zhu et al., 2014). We speculated that the *GmTIFY10* and *GmTIFY11* homologous genes may be involved in responding to salt stress.

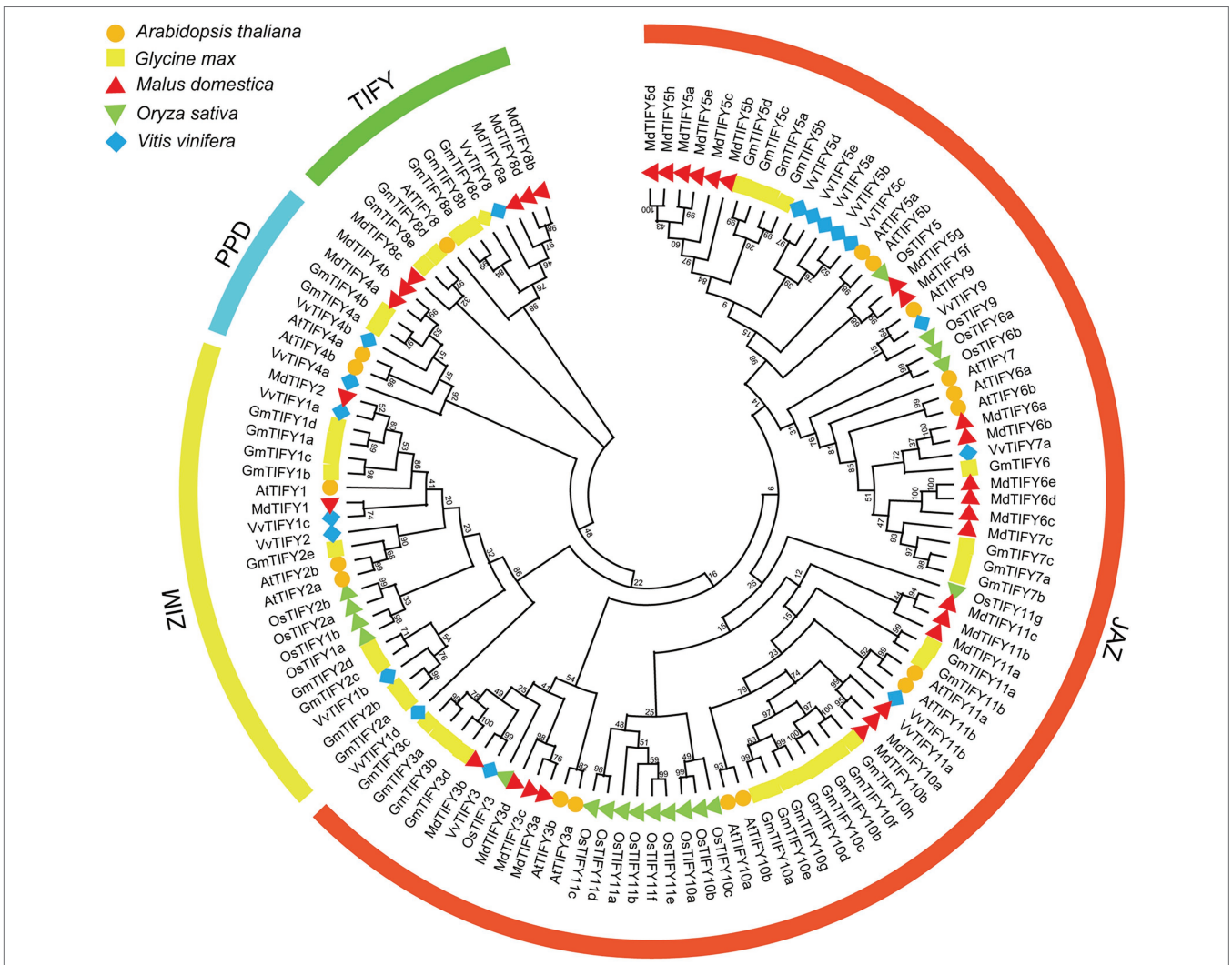


FIGURE 1 | The phylogenetic tree of TIFY proteins from *Arabidopsis*, rice, grape, apple, and soybean. Multiple sequence alignment was performed by MEGA-X and the phylogenetic tree was constructed using the ML method. The TIFY, JAZ, PPD, and ZIM subfamilies are represented by different colors.

Chromosomal Location and Gene Duplication

In the soybean genome, 38 *GmTIFY* genes are irregularly distributed on 16 chromosomes (**Figure 2**). The duplication events analysis demonstrated that nine *GmTIFY* genes were singletons and 29 *GmTIFY* genes experienced gene duplication events, including eight genes with segmental duplication and 21 genes with dispersed duplication (**Figure 3**). The most duplication events occurred in the JAZ subfamily (29 times). Only one duplication event occurred in the PPD subfamily. Further analysis demonstrated that large segmental chromosome duplication events occurred between chromosomes 4/6 and chromosomes 10/20. The most duplication events were identified in the *TIFY10* homologous genes in the JAZ subfamily (15 times).

To determine the significance of *GmTIFY* family genes during evolution, the EasyCodeML software was used to test the selective pressure of *GmTIFY* genes including purification selection, positive selection, and negative selection (**Table 2**). Since the *TIFY1* and *TIFY2* homologous genes are always grouped together in the evolutionary tree, we calculated their total ω. The positive selection of *GmTIFY1*, *GmTIFY2*, *GmTIFY3*, *GmTIFY7*, and *GmTIFY10* homologous genes exceeded 1, which indicated that these *TIFY* homologous genes experienced positive selection during their evolutionary history.

Gene Structure, Motif Composition, and *cis*-Element Analysis of *GmTIFY* Genes

The structural characteristics of *GmTIFY* family genescan be obtained by analyzing the phylogenetic tree, motifs, and positions of exons and introns (**Figure 4**). These results showed that genes belonging to the same phylogenetic group have similar motifs and exon/intron structures. According to previous studies of *TIFY* genes, *GmTIFY* family genes were divided into four subfamilies (TIFY, JAZ, PPD, and ZML) with different structural features. Five conserved motifs were found based on the analysis of TIFY protein sequences (**Figure 4B**). Motif 1 was identified as TIFY domain and distributed in all genes. Motif 2 (Jas domain) was contained by the JAZ and PPD subfamily genes. Motif 3 and 4 were identified as the sequence of CCT domain and the ZnF_GATA domain in ZIM subfamily, respectively. The PPD subfamily contains motif 1, 2, and 5, while motif 2 of the PPD subfamily lacks a PY motif which different from the original Jas domain. There was no PPD motif sequence information in the Pfam database. The conservative motif 5 domain sequence was constructed with PPD subfamily genes which were previously annotated in *Arabidopsis*, apples, and grapes. The motif 5 was confirmed with sequences from the report of earliest defined PPD genes (Bai et al., 2011). Conserved sequences were submitted to the MEME online tool to generate the domain logo (**Figure 4D**). Analysis of the promoters through PlantCARE website revealed

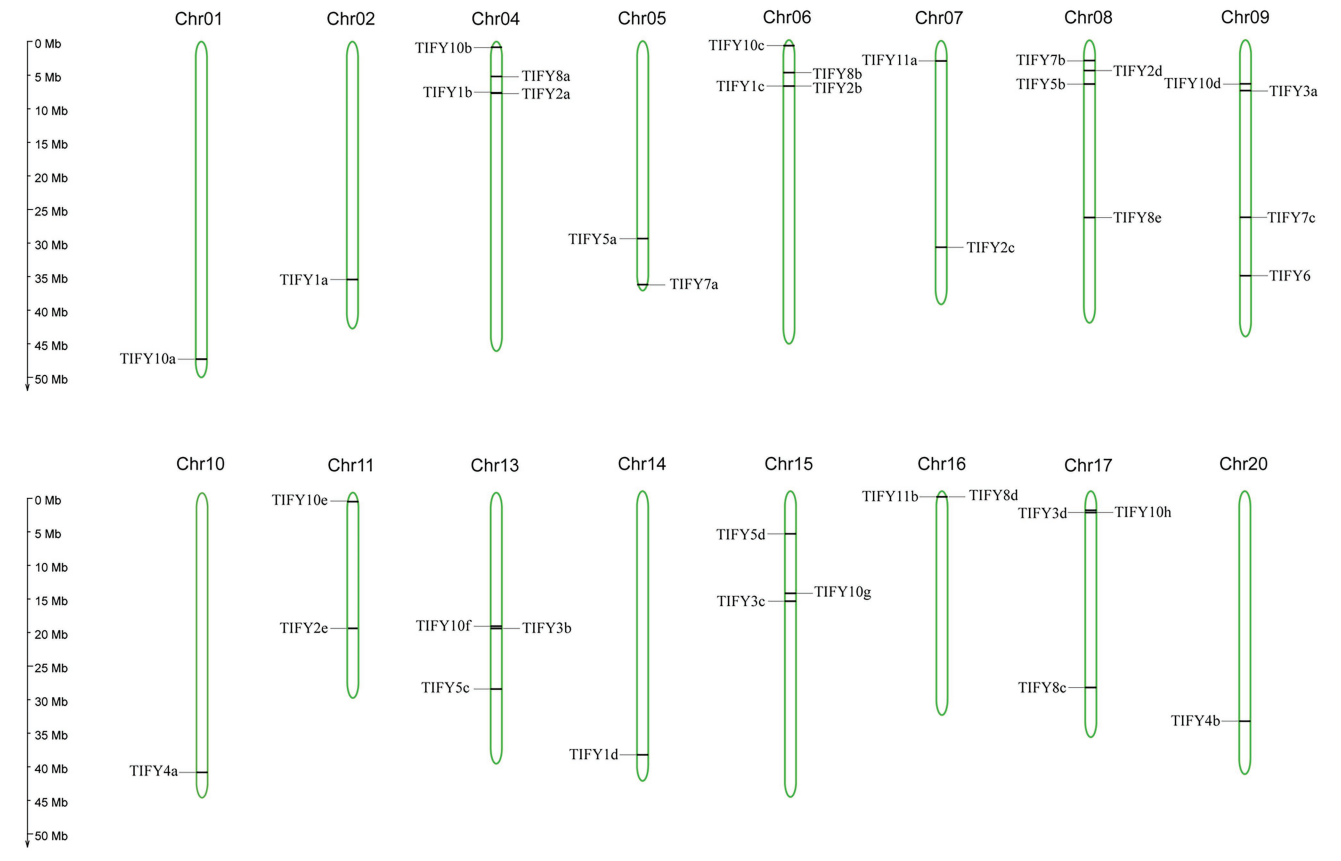


FIGURE 2 | The location of 38 *GmTIFY* genes on the soybean chromosome. The scale bar on the left indicates the size of the chromosomes.

that *GmTIFY* family genes contain many abiotic stresses responsive *cis*-elements, such as ABA-responsive element (ABRE), MYB banding site (MBS), and methyl jasmonate-responsive element (MeJA element) (**Supplementary Figure 1**).

Expression Patterns of *GmTIFY* Genes

To obtain the expression profiles of soybean *GmTIFY* genes in different tissues, RNA-seq data were downloaded from the Phytozome database and visualized by using the TBtools software (**Figure 5**). Results demonstrated that *GmTIFY* genes had significant expression differences in multiple tissues. The expression levels of most genes are relatively low in all tissues or are not even expressed. The homologous genes of *TIFY3*, *TIFY5*, *TIFY10*, and *TIFY11* in the JAZ subfamily showed higher expression in the roots, stems, and leaves.

We analyzed the transcriptome data of *GmTIFY* genes under drought, salt and ABA treatment (**Figure 6**). Most *TIFY10* and *TIFY11* homologous genes significantly responded to salt treatment. From these homologous genes, six significantly upregulated candidate genes (*GmTIFY10a*, *GmTIFY10e*, *GmTIFY10f*, *GmTIFY10g*, *GmTIFY11a*, and *GmTIFY11b*) were selected and confirmed their expression level under drought and salt stress by RT-qPCR (**Figure 7**). The six genes were not significantly changed after drought treatment. After salt stress, the expression levels of six genes were significantly upregulated (>10-fold), especially *GmTIFY10e* and *GmTIFY10g*, which reached the peak at 1 h.

GmTIFY10e and *GmTIFY10g* Are Localized in the Nucleus

The results of the phylogenetic, duplication events, selective pressure, and expression profiles of *GmTIFY* genes indicated that *GmTIFY10e* and *GmTIFY10g* could play an important role in soybean. We selected *GmTIFY10e* and *GmTIFY10g* for further study, both of which significantly responded to salt stress. To

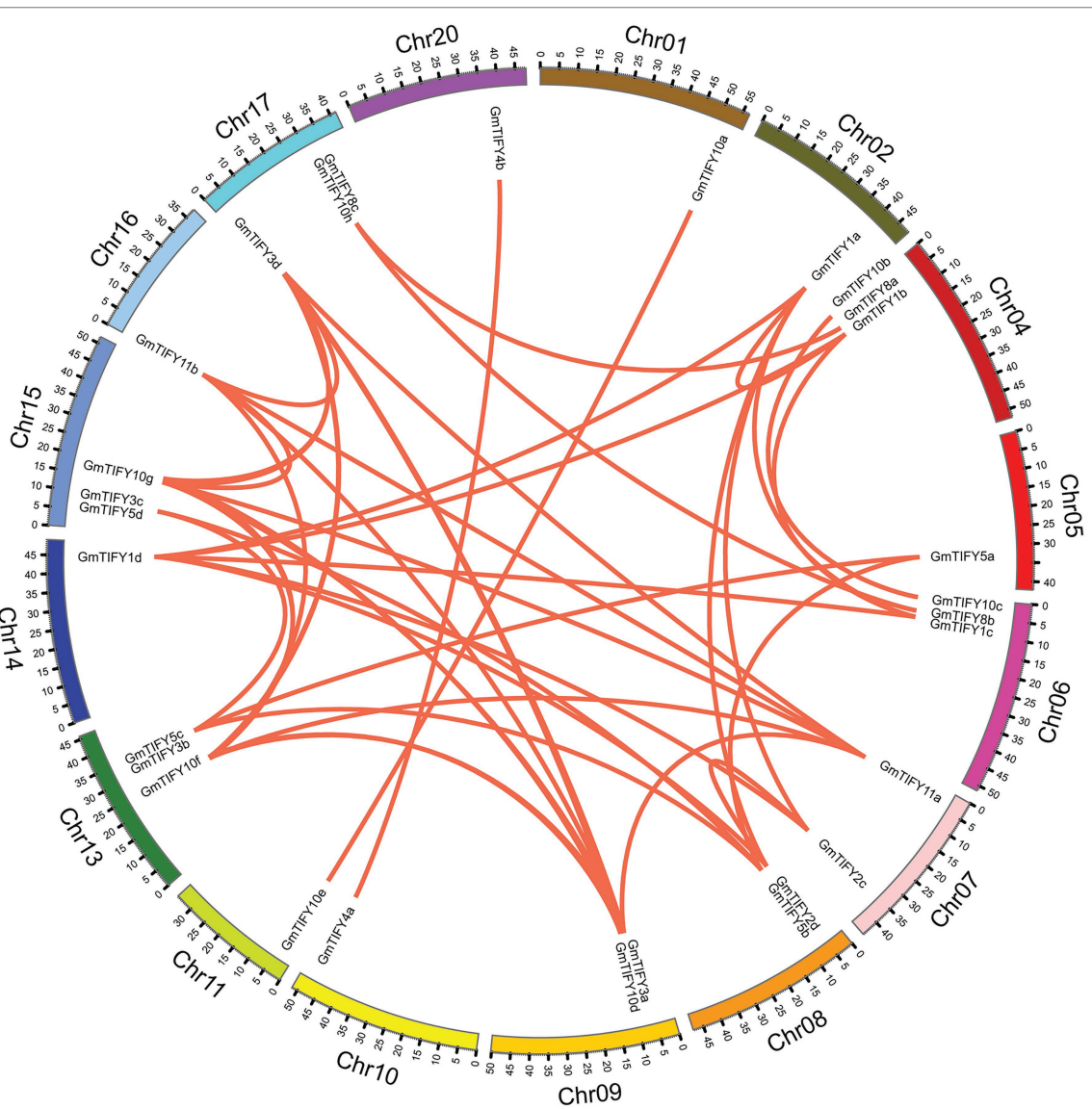


FIGURE 3 | Distribution of segmentally duplicated *GmTIFY* genes on soybean chromosomes. Red lines indicate duplicated *GmTIFY* gene pairs.

TABLE 2 | Analysis of natural selection patterns using PAML.

Subfamily	Group	Model	LnL	Estimates of Parameters	
				Background (ω)	Foreground (ω)
ZIM	1 + 2	Two-ratio Model 2	−1,672.124876	0.60068	1.28945
		Model 0	−1,672.422571	0.61229	
JAZ	3	Two-ratio Model 2	−1,672.472447	0.61143	1.59695
		Model 0	−1,672.422571	0.61229	
	4	Two-ratio Model 2	−1,672.402008	0.61570	0.47849
		Model 0	−1,672.422571	0.61229	
	5	Two-ratio Model 2	−1,672.420049	0.61160	0.68454
		Model 0	−1,672.422571	0.61229	
	6	Two-ratio Model 2	−1,672.411868	0.60902	0.66408
		Model 0	−1,672.422571	0.61229	
	7	Two-ratio Model 2	−1,672.422571	0.61229	1.47477
		Model 0	−1,672.422571	0.61229	
TIFY	8	Two-ratio Model 2	−1,672.417476	0.60916	0.74566
		Model 0	−1,672.422571	0.61229	
JAZ	10	Two-ratio Model 2	−1,672.422575	0.61229	1.65538
		Model 0	−1,672.422571	0.61229	
	11	Two-ratio Model 2	−1,672.422571	0.61230	0.61216
		Model 0	−1,672.422571	0.61229	

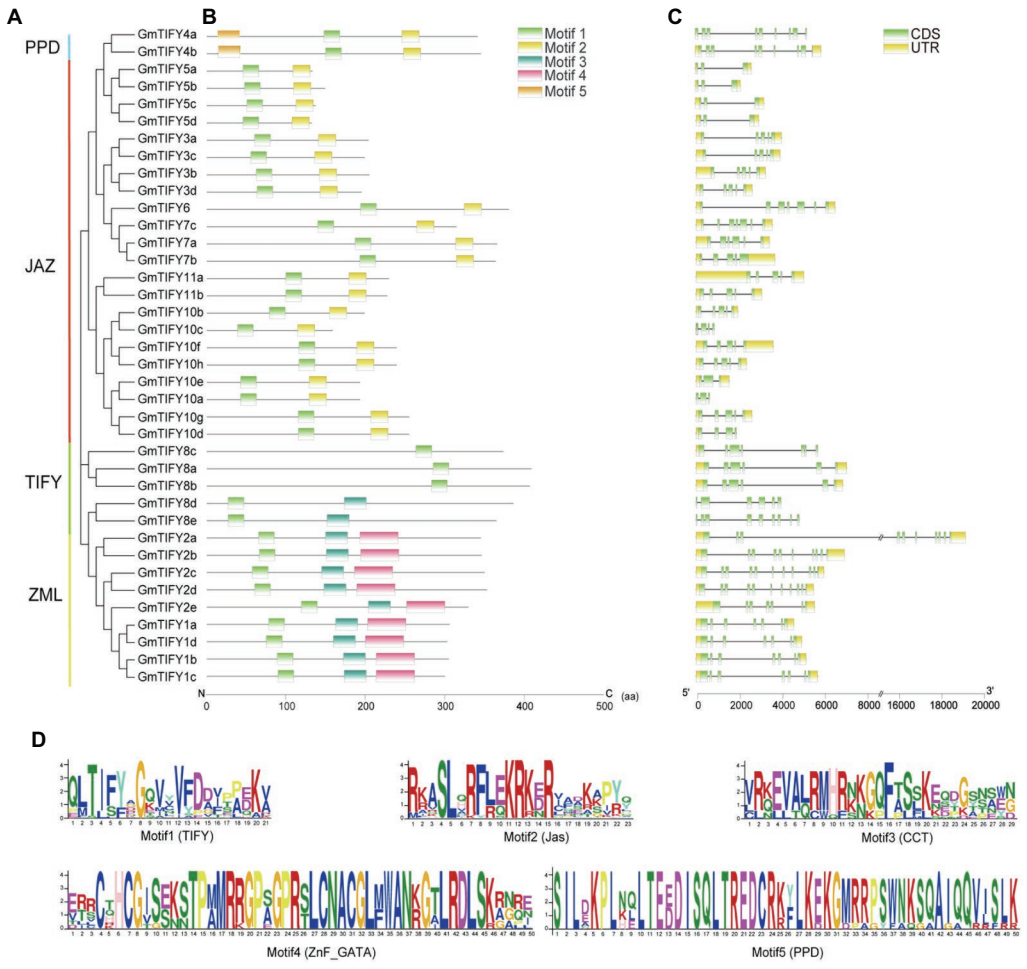


FIGURE 4 | Phylogenetic relationships, conserved motifs, and gene structures of GmTIFY. **(A)** Phylogenetic tree of TIFY proteins from soybean constructed using the ML method. **(B)** Distribution of conserved motifs in GmTIFY proteins. Five putative motifs are indicated by colored boxes. **(C)** Exon/intron organization of *GmTIFY* genes. Green boxes represent exons and black lines represent introns. The upstream/downstream regions of *GmTIFY* genes are indicated by yellow boxes. The scale at the bottom is used to infer exons length. **(D)** Motif logos of five conservative motifs detected in GmTIFY proteins *via* MEME analysis.

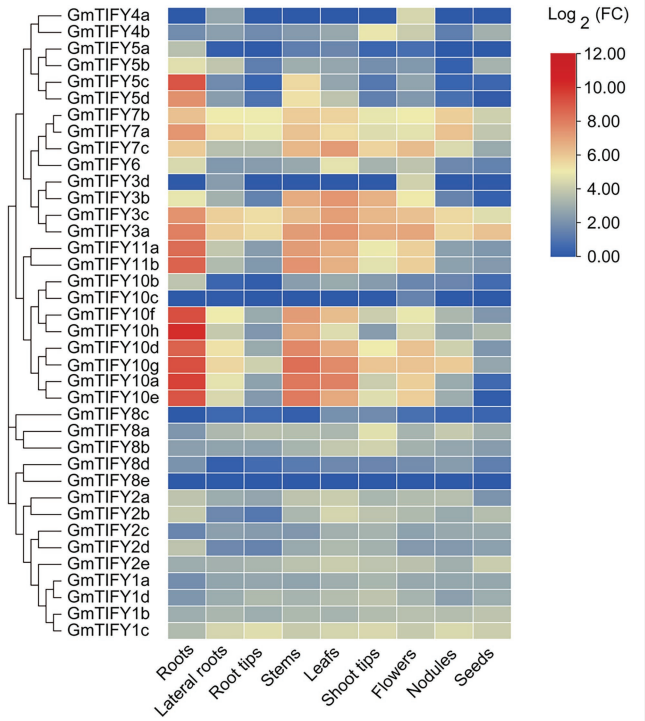


FIGURE 5 | The expression profiles of 38 *GmTIFY* genes in different soybean tissues. The expression data of 38 *GmTIFY* genes in different tissues were obtained from the Phytozome database.

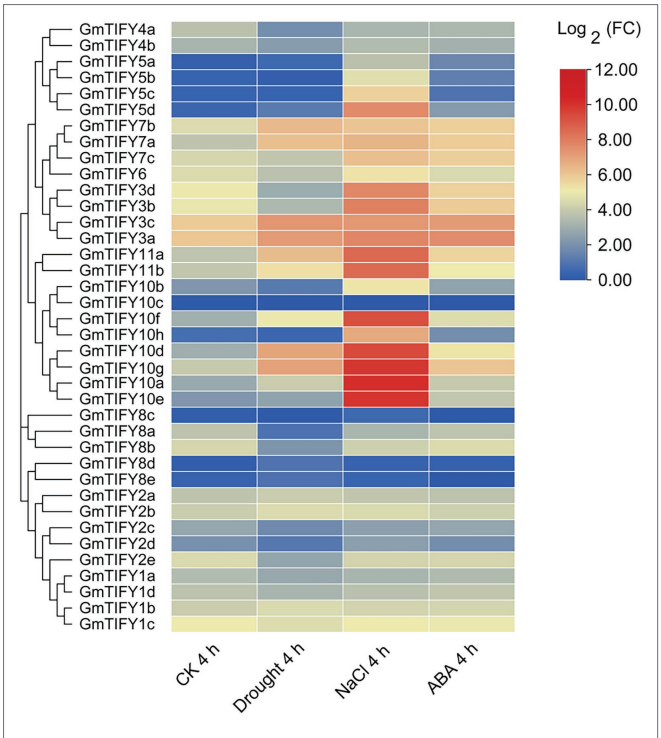


FIGURE 6 | The expression profiles of 38 *GmTIFY* genes under different conditions including drought, salt and ABA treatments.

determine the subcellular localization of *GmTIFY10e* and *GmTIFY10g*, we transformed the recombinant 16318hGFP vector linked to the *GmTIFY10e* and *GmTIFY10g* into *Arabidopsis* protoplasts using the PEG4000-mediated method, respectively (**Figure 8**). Both *GmTIFY10e*-hGFP and *GmTIFY10g*-hGFP fusion proteins are located in the nucleus.

GmTIFY10e and *GmTIFY10g* Can Improve Salt Tolerance in Transgenic *Arabidopsis*

To investigate the function of *GmTIFY10e* and *GmTIFY10g* in plant salt tolerance, we obtained transgenic *Arabidopsis* lines with high expression levels of *GmTIFY10e* and *GmTIFY10g*, respectively. For root length assay, the 5-day-old *Arabidopsis* plants were transferred to MS medium containing 125 mM NaCl and we calculated the primary root lengths and fresh weights. Under normal conditions, there was no significant difference between the WT plants and transgenic lines. However, the growth of the WT plants was significantly repressed when treated with 125 mM NaCl and the transgenic lines showed better growth than WT plants, with longer root lengths and heavier fresh weights (**Figures 9A,C,D**). To verify the salt tolerance of plants in soil, 21-day-old plants were subjected to salt stress (**Figure 9B**). There is no significant difference between the transgenic and WT plants under normal growth conditions. After 14 days of 250 mM NaCl treatment, the transgenic lines showed higher salt tolerance than WT, when WT is seriously wilted or even dead.

Salt stress can reduce the scavenging function of reactive oxygen species (ROS) in cells (Raza et al., 2021). Accumulation of ROS will lead to membrane lipid peroxidation, forming MDA, and activating oxygen enzymatic scavenging system including CAT, POD (Mhamdi and Van Breusegem, 2018; Wang et al., 2021a; Youssef et al., 2021). At the same time, plant cells will accumulate a large number of PRO under stress to maintain normal cell swelling pressure, prevent excessive water loss of protoplasm and enhance the adaptability of plants to adversity (Wang et al., 2021b). We measured the MDA, PRO, CAT, and POD contents of transgenic and WT lines under normal and salt treatment conditions (**Figures 9E–H**). Compared with WT plants, the contents of PRO, CAT, and POD in the transgenic lines were significantly increased, while the MDA contents of the transgenic lines were significantly reduced. These results all indicated that *GmTIFY10e* and *GmTIFY10g* play a role in improving the tolerance of salt stress.

Overexpression of *GmTIFY10e* and *GmTIFY10g* Improve Salt Tolerance in Soybean Hairy Roots

To verify the function of *GmTIFY10e* and *GmTIFY10g* in response to salt stress in soybean, we obtained overexpressing plants (OE), empty vector plants (EV) and RNAi plants with *A. rhizogenes*-mediated transformation of soybean hairy roots and treated with 250 mM NaCl. RT-qPCR analysis demonstrated that the expression levels of OE plants were significantly higher than that of the EV, while the expression levels of RNAi plants

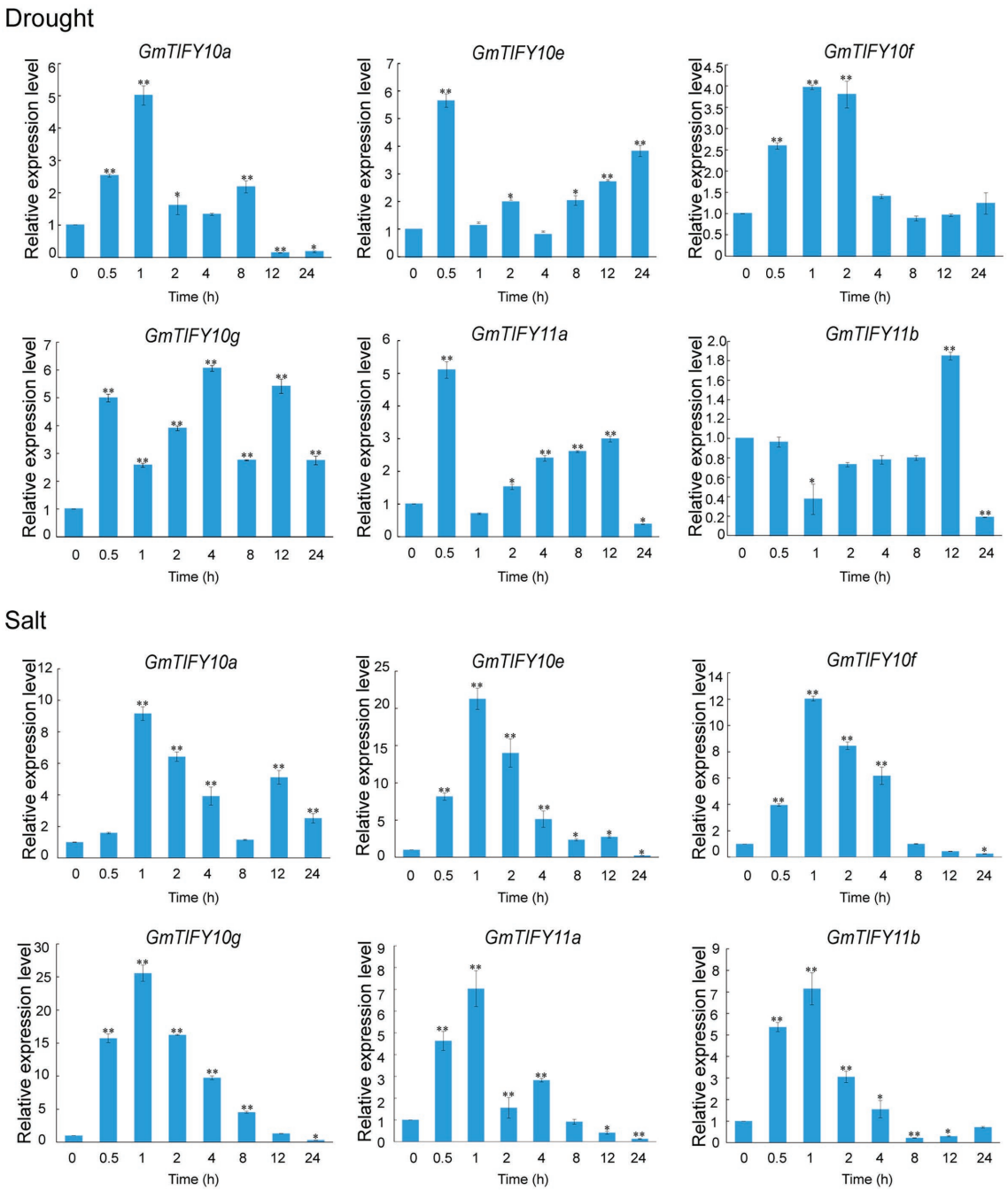


FIGURE 7 | Expression patterns of *GmTIFY* genes under drought and salt stress conditions. Expression levels of six *GmTIFY* genes were measured using RT-qPCR at different times under drought and salt treatments. RT-qPCR data were normalized using *GmELF1b* as the reference gene and were displayed relative to 0 h. The *x*-axes show the duration of treatment, and the y-axes depict relative expression levels (error bars indicate SD). The data are shown as the means of three biological replicates ±SD. ANOVA test demonstrates that there are significant differences ($^{*}p < 0.05$, $^{**}p < 0.01$).

were lower than that of the EV (**Supplementary Figure 3**). Phenotypic identification showed that there was no significant difference among the OE plants, RNAi plants, or EV plants under normal growth conditions. After 5 days treatment of salt, the leaves of the RNAi plants turned yellow and were more wilted than the EV plants. In the OE lines, only a few bottom leaves showed yellowing and wilting, which indicated significantly increased salt tolerance (**Figures 10A,B**). It indicated that overexpression of *GmTIFY10e* and *GmTIFY10g* improved the salt stress tolerance in soybean.

Reactive oxygen species is the most important signal substance for plants to respond to abiotic stresses. A large amount of ROS

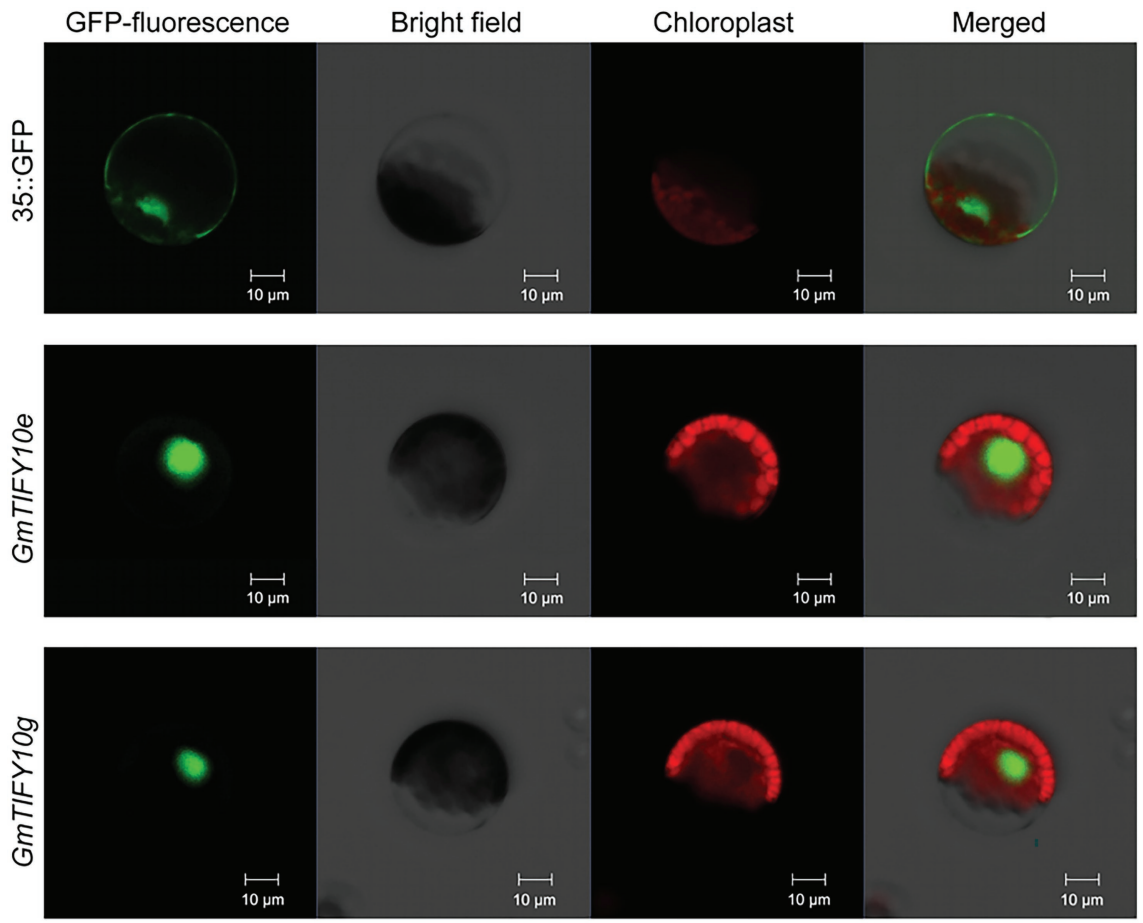


FIGURE 8 | Subcellular localization of *GmTIFY10e*-hGFP and *GmTIFY10g*-hGFP fusion protein. 35S::16318hGFP was used as a control. The scale bar of 35S::16318hGFP, *GmTIFY10e*-hGFP and *GmTIFY10g*-hGFP indicates 10 μm.

will accumulate in plants under abiotic stresses. NBT staining can reveal the levels of ROS accumulation in plants (**Figures 10C,D**). We measured the levels of ROS accumulation of plants roots. After salt treatment, the staining levels of OE plants were significantly lighter than that of EV plants and the RNAi plants showed more significant NBT stains. The same results were obtained from staining the leaves of OE plants, RNAi plants, and EV plants (**Supplementary Figures 4A,F**).

We further measured the contents of MDA, PRO, CAT, and POD in the roots of OE plants, RNAi plants and EV plants (**Figures 10E–H**). Under normal growth conditions, the MDA, PRO, CAT, and POD contents in OE plants and RNAi plants were not significantly different than in EV plants. After salt treatment, the PRO, POD, and CAT contents in OE plants were significantly higher than in EV plants and RNAi plants, while the MDA content in OE plants was lower than in EV plants and RNAi plants. We obtained similar results by measuring the MDA, PRO, CAT, and POD contents of the leaves (**Supplementary Figures 4C–F**). These results further confirmed that overexpression of *GmTIFY10e* and *GmTIFY10g* can enhance salt stress tolerance in soybean, which is consistent with the results of transgenic *Arabidopsis*.

Overexpression of *GmTIFY10e* and *GmTIFY10g* Can Influence the Expression Levels of ABA-Related Genes

Previous transcriptome data demonstrated that the expression levels of *GmTIFY10e* and *GmTIFY10g* were upregulated under ABA treatment. Enrichment analysis of *GmTIFY10e* and *GmTIFY10g* co-expression genes demonstrated that *GmTIFY10e* and *GmTIFY10g* can participate in the JA signal pathway and the ABA signal pathway (**Supplementary Figure 5**). In previous studies, the JAZ proteins were identified to participate in the ABA-dependent signal pathway through the target protein *MYC*/*MYB* transcription factor or ABA signal receptor *PYL4* (Fu et al., 2017; Luo et al., 2020). To study the possible mechanisms regulated by *GmTIFY10e* and *GmTIFY10g* during salt stress responses, we analyzed the expression levels of *GmCAT1*, *GmPOD,* and *GmERF115* which were genes in the ABA-mediated osmotic stress signals, and the *GmSnRK2*, *GmPP2C,* and *GmMYC2* which were key genes in the ABA signal transduction pathway (**Figure 11**). Under normal growth conditions, the expression levels of *GmCAT1*, *GmPOD,* and *GmERF115* genes were higher in OE plants and were lower in RNAi plants compared to EV plants, while *GmSnRK2*,

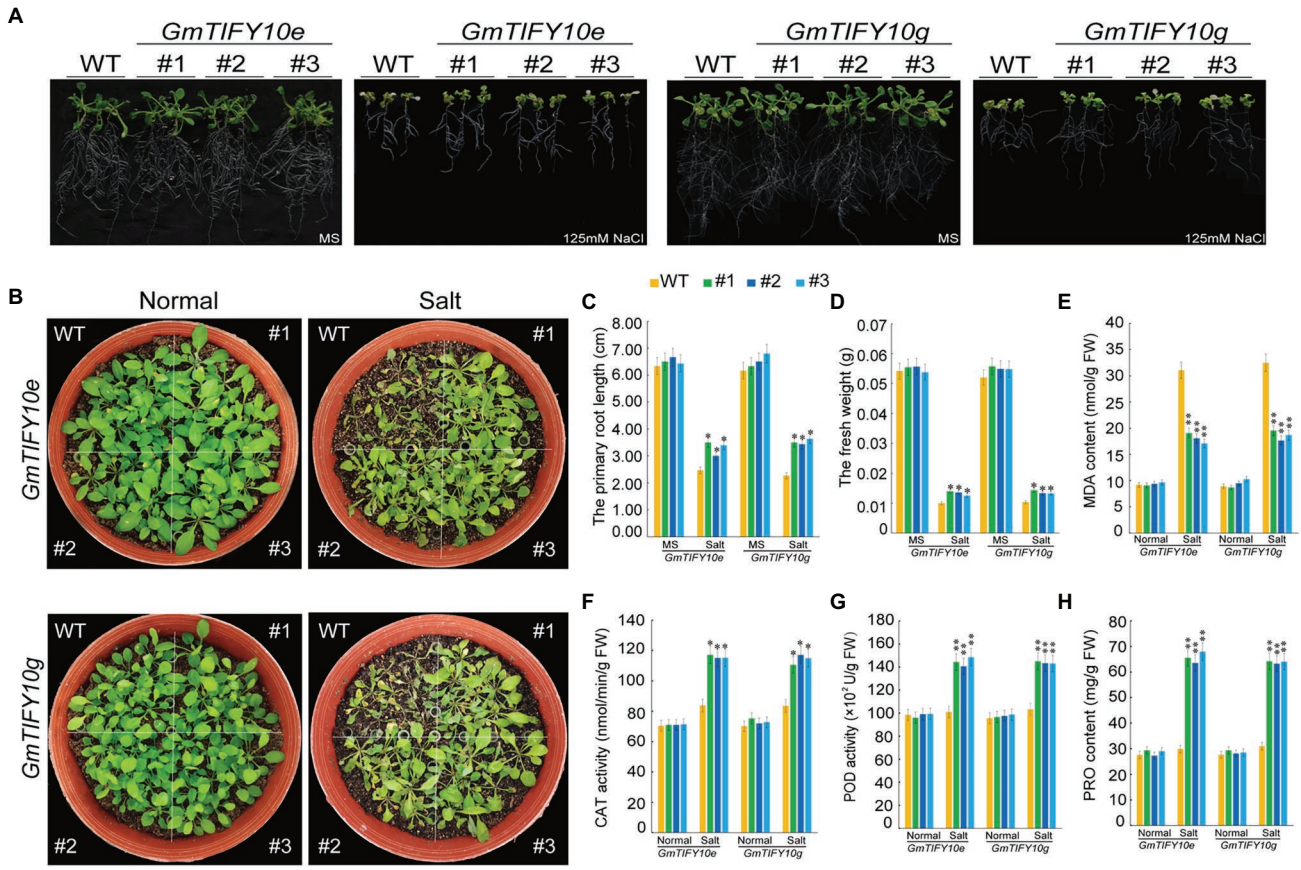


FIGURE 9 | Overexpression of *GmTIFY10e* and *GmTIFY10g* in *Arabidopsis* plants enhanced tolerance to salt stress. **(A)** Phenotypic analysis of transgenic *Arabidopsis* lines and WT lines under normal and 125 mM NaCl treatment. The scale bar indicates 2 cm. **(B)** Phenotypes of WT and transgenic plants under salt stress. **(C)** The primary root length analysis of the transgenic *Arabidopsis* lines and WT plants under normal and 125 mM NaCl treatment. **(D)** The fresh weights of transgenic *Arabidopsis* lines and WT lines under normal and 125 mM NaCl conditions. **(E–H)** The leaf contents of MDA **(E)**, CAT **(F)**, POD **(G),** and PRO **(H)** in WT and transgenic plants under normal and salt stress. The data are shown as the means ± SD obtained from three biological replicates. ANOVA test demonstrates significant differences compared with WT ($^{*}p < 0.05$, $^{**}p < 0.01$).

GmPP2C, and *GmMYC2* were higher in RNAi plants. For salt stress, the expression levels of *GmCAT1*, *GmPOD,* and *GmERF115* were significantly upregulated in OE plants and downregulated in RNAi plants, while *GmSnRK2*, *GmPP2C,* and *GmMYC2* significantly upregulated in RNAi plants and downregulated in OE plants. These results indicated that *GmTIFY10e* and *GmTIFY10g* may be involved in responding to salt stress through the ABA regulation pathway.

DISCUSSION

As a kind of specific plant proteins, TIFY proteins play an important role in the plant growth and response to environmental changes. However, the information about the expression and function of TIFY family in soybean are very limited. This study systematically identified and analyzed 38 *GmTIFY* genes and their responses to abiotic stresses in soybean. Our results confirmed that soybean *GmTIFY10e and GmTIFY10g* genes can positively regulate the salt stress tolerance in plants. *AtTIFY10a* and *AtTIFY10b* significantly improved tolerance to salt and alkali stresses in *Arabidopsis* plants (Zhu et al., 2014). In cotton, overexpression of *TIFY10* homologous genes *GaJAZ1* could significantly improve the tolerance to salt stress. In addition, the *GaJAZ1* was confirmed to interact with *GaMYC2* to repress expression of downstream genes related to ABA signaling pathways, affecting plant tolerance to salinity stress. Therefore, *TIFY10* homologous genes may be involved in abiotic stresses tolerance, including salt tolerance (Zhao et al., 2020b).

ABA and ABA signaling pathway play important roles in regulating various stress responses (Stone et al., 2006; Nilson and Assmann, 2007; Sirichandra et al., 2009). ABA-related *cis*-elements can combine with transcription factors to regulate the expression of corresponding genes and regulate the sensitivity of plants to ABA signaling pathway (Kim et al., 2011). *GmTIFY10e* and *GmTIFY10g* have several ABA-related elements and its co-expression genes were involved in the ABA-activated signaling pathway and the JA signaling pathway (**Supplementary Figures 1, 5**). Transcriptome data demonstrated that the expression levels of

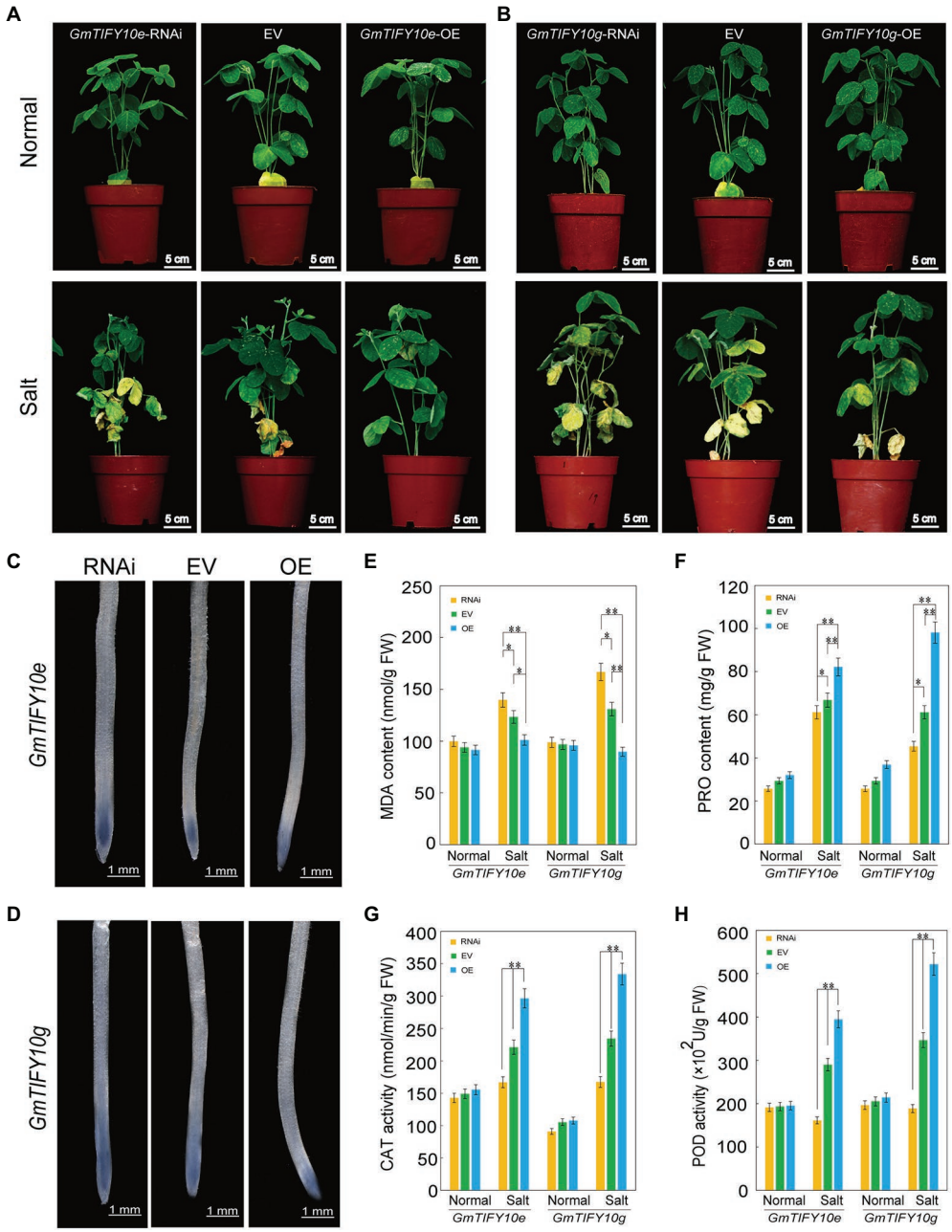


FIGURE 10 | Analysis of the function of soybean *GmTIFY10e* and *GmTIFY10g* under normal and salt stress. **(A,B)** Phenotype analysis of EV and transgenic plants under normal and salt stress. The scale bar indicates 5 cm. **(C,D)** NBT staining of EV and transgenic plant roots under salt stress. The scale bar indicates 1 mm. **(E–H)** The MDA **(E)**, PRO **(F)**, CAT **(G)**, and POD **(H)** contents of EV and transgenic plant leaves under normal and salt stress. The data are shown as the means ± SD obtained from three biological replicates. ANOVA test demonstrates that there are significant differences compared with EV ($^{*}p < 0.05$, $^{**}p < 0.01$).

GmTIFY10e and *GmTIFY10g* were upregulated under ABA treatment (**Figure 6**). These results indicated that these two genes may be related to ABA signaling pathway.

The tolerance of ABA to environmental stresses mainly depends on antioxidant protection system (Wu et al., 2021). Abiotic stresses can induce the production of H_2O_2 in plants. H_2O_2 can directly act on the negative regulatory factor *PP2Cs* of ABA signaling pathway and promote the expression of *CAT1* and *POD* genes. The *AtERF115* can mediate the ROS pathway and maintain the root stem and root growth through phytosulfokine (PSK) peptide incorporation (Kong et al., 2018). *GmTIFY10e* and *GmTIFY10g* could significantly affect the expression levels of *GmCAT1*, *GmPOD*, and *GmERF115* genes under salt stress (**Figures 11A–C, G–I**). To further analyze

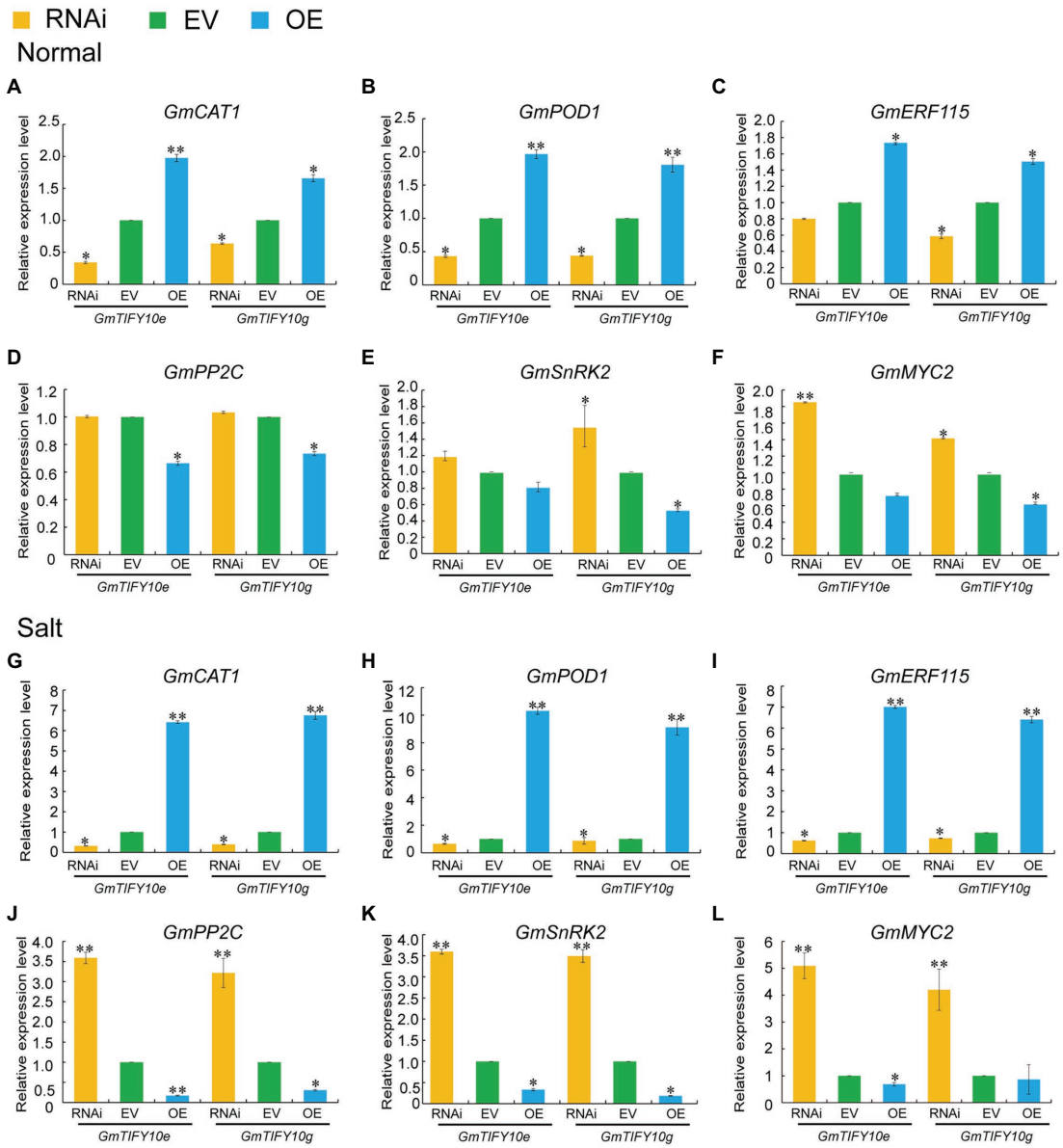


FIGURE 11 | (A–F) The GmCAT1 **(A)** , GmPOD **(B)** , GmERF115 **(C)** , GmPP2C **(D)**, GmSnRK2 **(E)** and GmMYC2 **(F)** expression levels of EV and transgenic plants under normal conditions. **(G–L)** The GmCAT1 **(G)**, GmPOD **(H)**, GmERF115 **(I)**, GmPP2C **(J)**, GmSnRK2 **(K)** and GmMYC2 **(L)** expression levels of EV and transgenic plants under salt stress. Vertical bars indicate ±SD of three replicates. ANOVA test demonstrates that there are significant differences compared with EV ($^{*}p<0.05$, $^{**}p<0.01$).

the regulation mechanism of *GmTIFY10e* and *GmTIFY10g*, we measured the key genes of ABA signal transduction pathway, which mainly includes *PP2Cs*, *SnRK2s,* and *MYCs*. The JAZ protein participated in the ABA-dependent signal pathway through its target protein *MYC*/*MYB* transcription factor and can interact with the ABA signal receptor *PYL4*. In presence of ABA, it will combine with *PYLs* and inhibit the phosphatase of *PP2Cs* and inhibit the *SnRK2s* (Lackman et al., 2011; Fu et al., 2017; Luo et al., 2020; Wang et al., 2020b). Our results showed that overexpression of *GmTIFY10e* and *GmTIFY10g* in soybean could significantly decrease the expression levels of *GmSnRK2*, *GmPP2C,* and *GmMYC2* genes compared with EV plants under salt stress (**Figures 11J–L**). Therefore, *GmTIFY10e* and *GmTIFY10g* maybe affect the salt stress tolerance through ABA pathway in plants.

CONCLUSION

We identified 38 *GmTIFY* genes in soybean genome, among which *GmTIFY10e* and *GmTIFY10g* were significantly upregulated by salt stress. Overexpression of *GmTIFY10e* and *GmTIFY10g* could improve the salt tolerance of transgenic plants by inhibiting the expression of key genes of ABA pathway. This research provides a basis for further study on how TIFY family members affect salt tolerance.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, and further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

Z-SX coordinated the project, conceived and designed the experiments, and edited the manuscript. Y-LL performed the experiments and wrote the first draft. LZ, Z-SX, and J-HL revised the manuscript. L-GJ, Y-XL, Y-NK, Y-XW, T-FY, JC, Y-BZ, F-ZW, and MC contributed to data analysis and managed reagents. Y-ZM and J-HL contributed with valuable discussions. All authors contributed to the article and approved the submitted version.

FUNDING

This research was financially supported by the National Natural Science Foundation of China (31871624), the Agricultural Science and Technology Innovation Program (CAAS-ZDRW202109 and CAAS-ZDRW202002), Chinese Academy of Agricultural Sciences and the central government of Shandong Province guides the development of local science and technology (YDZX20203700002548), and Key R&D Projects of Shandong Province (2021LZGC014).

ACKNOWLEDGMENTS

We are grateful to Li-Juan Qiu and Shi Sun of the Institute of Crop Science, Chinese Academy of Agricultural Sciences (CAAS) for kindly providing soybean seeds. Also, we are grateful to Wen-Sheng Hou and Hui Zhang of the Institute of Crop Science, CAAS for kindly providing vectors and the protocol for high-efficiency *A. rhizogenes*-mediated transformation, respectively.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.845314/full#supplementary-material

REFERENCES

Appel, R., Bairoch, A., and Hochstrasser, D. (1994). A new generation of information retrieval tools for biologists: the example of the ExPASy WWW server. *Trends Biochem. Sci.* 19, 258–260. doi: 10.1016/0968-0004(94)90153-8

Bai, Y., Meng, Y., Huang, D., Qi, Y., and Chen, M. (2011). Origin and evolutionary analysis of the plant-specific TIFY transcription factor family. *Genomics* 98, 128–136. doi: 10.1016/j.ygeno.2011.05.002

Bailey, T., Boden, M., Buske, F., Frith, M., Grant, C., Clementi, L., et al. (2009). MEME suite: tools for motif discovery and searching. *Nucleic Acids Res.* 37, W202–W208. doi: 10.1093/nar/gkp335

Barah, P., and Bones, A. (2015). Multidimensional approaches for studying plant defence against insects: from ecology to omics and synthetic biology. *J. Exp. Bot.* 66, 479–493. doi: 10.1093/jxb/eru489

Bohnert, H., Nelson, D., and Jensen, R. (1995). Adaptations to environmental stresses. *Plant Cell* 7, 1099–1111. doi: 10.2307/3870060

Chung, H., Niu, Y., Browse, J., and Howe, G. (2009). Top hits in contemporary JAZ: an update on jasmonate signaling. *Phytochemistry* 70, 1547–1559. doi: 10.1016/j.phytochem.2009.08.022

Clough, S., and Bent, A. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16, 735–743. doi: 10.1046/j.1365-313x.1998.00343.x

Demianski, A., Chung, K., and Kunkel, B. (2012). Analysis of *Arabidopsis JAZ* gene expression during *Pseudomonas syringae* pathogenesis. *Mol. Plant Pathol.* 13, 46–57. doi: 10.1111/j.1364-3703.2011.00727.x

Dhakarey, R., Raorane, M., Treumann, A., Peethambaran, P., Schendel, R., Sahi, V., et al. (2018). cpm2Corrigendum: Physiological and proteomic analysis of the rice mutant suggests a negative regulatory role of jasmonic acid in drought tolerance. *Front. Plant Sci.* 9:465. doi: 10.3389/fpls.2018.00465

Du, Y., Zhao, M., Wang, C., Gao, Y., Wang, Y., Liu, Y., et al. (2018). Identification and characterization of *GmMYB118* responses to drought and salt stress. *BMC Plant Biol.* 18:320. doi: 10.1186/s12870-018-1551-7

Ebel, C., BenFeki, A., Hanin, M., Solano, R., and Chini, A. (2018). Characterization of wheat (*Triticum aestivum*) TIFY family and role of *Triticum Durum TdTIFY11a* in salt stress tolerance. *PLoS One* 13:e0200566. doi: 10.1371/journal.pone.0200566

Fu, J., Wu, H., Ma, S., Xiang, D., Liu, R., and Xiong, L. (2017). *OsJAZ1* attenuates drought resistance by regulating JA and ABA signaling in rice. *Front. Plant Sci.* 8:2108. doi: 10.3389/fpls.2017.02108

Gao, F., Chen, C., Arab, D., Du, Z., He, Y., and Ho, S. (2019). EasyCodeML: a visual tool for analysis of selection using CodeML. *Ecol. Evol.* 9, 3891–3898. doi: 10.1002/ece3.5015

Guo, A., Zhu, Q., Chen, X., and Luo, J. (2007). GSDS: a gene structure display server. *Hereditas* 29, 1023–1026. doi: 10.1360/yc-007-1023

He, G., Xu, J., Wang, Y., Liu, J., Li, P., Chen, M., et al. (2016). Drought-responsive WRKY transcription factor genes *TaWRKY1* and *TaWRKY33* from wheat confer drought and/or heat resistance in *Arabidopsis*. *BMC Plant Biol.* 16:116. doi: 10.1186/s12870-016-0806-4

Jian, B., Liu, B., Bi, Y., Hou, W., Wu, C., and Han, T. (2008). Validation of internal control for gene expression study in soybean by quantitative real-time PCR. *BMC Mol. Biol.* 9:59. doi: 10.1186/1471-2199-9-59

Jiangtao, C., Yingzhen, K., Qian, W., Yuhe, S., Daping, G., Jing, L., et al. (2015). MapGene2Chrom, a tool to draw gene physical map based on Perl and SVG languages. *Yi Chuan* 37, 91–97. doi: 10.16288/j.yczz.2015.01.013

Kereszt, A., Li, D., Indrasumunar, A., Nguyen, C., Nontachaiyapoom, S., Kinkema, M., et al. (2007). *Agrobacterium rhizogenes*-mediated transformation of soybean to study root biology. *Nat. Protoc.* 2, 948–952. doi: 10.1038/nprot.2007.141

Kim, J. S., Mizoi, J., Yoshida, T., Fujita, Y., Nakajima, J., Ohori, T., et al. (2011). An ABRE promoter sequence is involved in osmotic stress-responsive expression of the DREB2A gene, which encodes a transcription factor regulating drought-inducible genes in *Arabidopsis*. *Plant Cell Physiol.* 52, 2136–2146. doi: 10.1093/pcp/pcr143

Kong, X., Tian, H., Yu, Q., Zhang, F., Wang, R., Gao, S., et al. (2018). *PHB3* maintains root stem cell niche identity through ROS-responsive *AP2*/*ERF* transcription factors in *Arabidopsis*. *Cell Rep.* 22, 1350–1363. doi: 10.1016/j.celrep.2017.12.105

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Lackman, P., González-Guzmán, M., Tilleman, S., Carqueijeiro, I., Pérez, A., Moses, T., et al. (2011). Jasmonate signaling involves the abscisic acid receptor *PYL4* to regulate metabolic reprogramming in *Arabidopsis* and tobacco. *Proc. Natl. Acad. Sci. U. S. A.* 108, 5891–5896. doi: 10.1073/pnas.1103010108

Leng, Z., Liu, Y., Chen, Z., Guo, J., Chen, J., Zhou, Y., et al. (2021). *GmDUF4228* genome-wide analysis of the DUF4228 family in soybean and functional identification of in response to drought and salt stresses. *Front. Plant Sci.* 12:628299. doi: 10.3389/fpls.2021.628299

Lestari, P., Van, K., Lee, J., Kang, Y. J., and Lee, S. H. (2013). Gene divergence of homeologous regions associated with a major seed protein content QTL in soybean. *Front. Plant Sci.* 4:176. doi: 10.3389/fpls.2013.00176

Letunic, I., Goodstadt, L., Dickens, N., Doerks, T., Schultz, J., Mott, R., et al. (2002). Recent improvements to the SMART domain-based sequence annotation resource. *Nucleic Acids Res.* 30, 242–244. doi: 10.1093/nar/30.1.242

Li, W., Wang, T., Zhang, Y., and Li, Y. (2017). Overexpression of soybean miR172c confers tolerance to water deficit and salt stress, but increases ABA sensitivity in transgenic *Arabidopsis thaliana*. *J. Exp. Bot.* 68, 4727–4729. doi: 10.1093/jxb/erw404

Li, X., Yin, X., Wang, H., Li, J., Guo, C., Gao, H., et al. (2014). Genome-wide identification and analysis of the apple (*Malus* × *domestica* Borkh.) TIFY gene family. *Tree Genet. Genomes* 11:808. doi: 10.1007/s11295-014-0808-z

Luo, X., Li, C., He, X., Zhang, X., and Zhu, L. (2020). ABA signaling is negatively regulated by *GbWRKY1* through *JAZ1* and *ABI1* to affect salt and drought tolerance. *Plant Cell Rep.* 39, 181–194. doi: 10.1007/s00299-019-02480-4

Mao, Y., Liu, Y., Chen, D., Chen, F., Fang, X., Hong, G., et al. (2017). Jasmonate response decay and defense metabolite accumulation contributes to age-regulated dynamics of plant insect resistance. *Nat. Commun.* 8:13925. doi: 10.1038/ncomms13925

Marchler-Bauer, A., Panchenko, A., Shoemaker, B., Thiessen, P., Geer, L., and Bryant, S. (2002). CDD: a database of conserved domain alignments with links to domain three-dimensional structure. *Nucleic Acids Res.* 30, 281–283. doi: 10.1093/nar/30.1.281

Meng, L., Zhang, T., Geng, S., Scott, P., Li, H., and Chen, S. (2019). Comparative proteomics and metabolomics of *JAZ7*-mediated drought tolerance in *Arabidopsis*. *J. Proteome* 196, 81–91. doi: 10.1016/j.jprot.2019.02.001

Mhamdi, A., and Van Breusegem, F. (2018). Reactive oxygen species in plant development. *Development* 145:dev164376. doi: 10.1242/dev.164376

Nilson, S., and Assmann, S. (2007). The control of transpiration. Insights from *Arabidopsis*. *Plant Physiol.* 143, 19–27. doi: 10.1104/pp.106.093161

Nishii, A., Takemura, M., Fujita, H., Shikata, M., Yokota, A., and Kohchi, T. (2000). Characterization of a novel gene encoding a putative single zinc-finger protein, ZIM, expressed during the reproductive phase in *Arabidopsis thaliana*. *Biosci. Biotechnol. Biochem.* 64, 1402–1409. doi: 10.1271/bbb.64.1402

Peethambaran, P., Glenz, R., Höninger, S., Shahinul Islam, S., Hummel, S., Harter, K., et al. (2018). Salt-inducible expression of *OsJAZ8* improves resilience against salt-stress. *BMC Plant Biol.* 18:311. doi: 10.1186/s12870-018-1521-0

Prince, V., and Pickett, F. (2002). Splitting pairs: the diverging fates of duplicated genes. *Nat. Rev. Genet.* 3, 827–837. doi: 10.1038/nrg928

Raza, A., Su, W., Gao, A., Mehmood, S. S., Hussain, M. A., Nie, W., et al. (2021). Catalase (CAT) gene family in rapeseed (*Brassica napus* L.): genome-wide analysis, identification, and expression pattern in response to multiple hormones and abiotic stress conditions. *Int. J. Mol. Sci.* 22:4281. doi: 10.3390/ijms22084281

Riechmann, J., Heard, J., Martin, G., Reuber, L., Jiang, C., Keddie, J., et al. (2000). *Arabidopsis* transcription factors: genome-wide comparative analysis among eukaryotes. *Science* 290, 2105–2110. doi: 10.1126/science.290.5499.2105

Shi, W., Du, Y., Ma, J., Min, D., Jin, L., Chen, J., et al. (2018). The WRKY transcription factor *GmWRKY12* confers drought and salt tolerance in soybean. *Int. J. Mol. Sci.* 19:4087. doi: 10.3390/ijms19124087

Sirichandra, C., Wasilewska, A., Vlad, F., Valon, C., and Leung, J. (2009). The guard cell as a single-cell model towards understanding drought tolerance and abscisic acid action. *J. Exp. Bot.* 60, 1439–1463. doi: 10.1093/jxb/ern340

Staswick, P. (2008). JAZing up jasmonate signaling. *Trends Plant Sci.* 13, 66–71. doi: 10.1016/j.tplants.2007.11.011

Stone, S., Williams, L., Farmer, L., Vierstra, R., and Callis, J. (2006). Keep on going, a ring E3 ligase essential for *Arabidopsis* growth and development, is involved in abscisic acid signaling. *Plant Cell* 18, 3415–3428. doi: 10.1105/tpc.106.046532

Su, H. G., Zhang, X. H., Wang, T. T., Wei, W. L., Wang, Y. X., Chen, J., et al. (2020). Genome-wide identification, evolution, and expression of GDSL-type esterase/lipase gene family in soybean. *Front. Plant Sci.* 11:726. doi: 10.3389/fpls.2020.00726

Sun, H., Chen, L., Li, J., Hu, M., Ullah, A., He, X., et al. (2017). The Jasmonate ZIM-domain gene family mediates JA signaling and stress response in cotton. *Plant Cell Physiol.* 58, 2139–2154. doi: 10.1093/pcp/pcx148

Thines, B., Katsir, L., Melotto, M., Niu, Y., Mandaokar, A., Liu, G., et al. (2007). JAZ repressor proteins are targets of the SCF(COI1) complex during jasmonate signalling. *Nature* 448, 661–665. doi: 10.1038/nature05960

Thireault, C., Shyu, C., Yoshida, Y., St Aubin, B., Campos, M., and Howe, G. (2015). Repression of jasmonate signaling by a non-TIFY JAZ protein in *Arabidopsis*. *Plant J.* 82, 669–679. doi: 10.1111/tpj.12841

Vanholme, B., Grunewald, W., Bateman, A., Kohchi, T., and Gheysen, G. (2007). The tify family previously known as ZIM. *Trends Plant Sci.* 12, 239–244. doi: 10.1016/j.tplants.2007.04.004

Wang, D., Liu, Y. X., Yu, Q., Zhao, S. P., Zhao, J. Y., Ru, J. N., et al. (2019). Functional analysis of the soybean *GmCDPK3* gene responding to drought and salt stresses. *Int. J. Mol. Sci.* 20:5909. doi: 10.3390/ijms20235909

Wang, F., Chen, H., Li, Q., Wei, W., Li, W., Zhang, W., et al. (2015). *GmWRKY27* interacts with *GmMYB174* to reduce expression of *GmNAC29* for stress tolerance in soybean plants. *Plant J.* 83, 224–236. doi: 10.1111/tpj.12879

Wang, F., Niu, H., Xin, D., Long, Y., Wang, G., Liu, Z., et al. (2021a). *OsIAA18*, an aux/IAA transcription factor gene, is involved in salt and drought tolerance in rice. *Front. Plant Sci.* 12:738660. doi: 10.3389/fpls.2021.738660

Wang, J., Ma, Z., Tang, B., Yu, H., Tang, Z., Bu, T., et al. (2021b). Tartary buckwheat (*Fagopyrum tataricum*) NAC transcription factors *FtNAC16* negatively regulates of pod cracking and salinity tolerant. *Int. J. Mol. Sci.* 22. doi: 10.3390/ijms22063197

Wang, T. T., Yu, T. F., Fu, J. D., Su, H. G., Chen, J., Zhou, Y. B., et al. (2020a). Genome-wide analysis of the GRAS gene family and functional identification of *GmGRAS37* in drought and salt tolerance. *Front. Plant Sci.* 11:604690. doi: 10.3389/fpls.2020.604690

Wang, Y., Qiao, L., Bai, J., Wang, P., Duan, W., Yuan, S., et al. (2017). Genome-wide characterization of Jasmonate-ZIM domain transcription repressors in wheat (*Triticum aestivum* L.). *BMC Genomics* 18:152. doi: 10.1186/s12864-017-3582-0

Wang, Z., Ren, Z., Cheng, C., Wang, T., Ji, H., Zhao, Y., et al. (2020b). Counteraction of ABA-mediated inhibition of seed germination and seedling establishment by ABA signaling terminator in *Arabidopsis*. *Mol. Plant* 13, 1284–1297. doi: 10.1016/j.molp.2020.06.011

Wu, H., Li, H., Zhang, W., Tang, H., and Yang, L. (2021). Transcriptional regulation and functional analysis of Nicotiana tabacum under salt and ABA stress. *Biochem. Biophys. Res. Commun.* 570, 110–116. doi: 10.1016/j.bbrc.2021.07.011

Xia, W., Yu, H., Cao, P., Luo, J., and Wang, N. (2017). Identification of TIFY family genes and analysis of their expression profiles in response to phytohormone treatments and melampsora larici-Populina infection in poplar. *Front. Plant Sci.* 8:493. doi: 10.3389/fpls.2017.00493

Ye, H., Du, H., Tang, N., Li, X., and Xiong, L. (2009). Identification and expression profiling analysis of TIFY family genes involved in stress and phytohormone responses in rice. *Plant Mol. Biol.* 71, 291–305. doi: 10.1007/s11103-009-9524-8

Youssef, M., Raafat, A., El-Yazied, A., Selim, S., Azab, E., Khojah, E., et al. (2021). Exogenous application of alpha-Lipoic acid mitigates salt-induced oxidative damage in sorghum plants through regulation growth, leaf pigments, ionic homeostasis, antioxidant enzymes, and expression of salt stress responsive genes. *Plan. Theory* 10. doi: 10.3390/plants10112519

Zhang, Y., Gao, M., Singer, S. D., Fei, Z., Wang, H., and Wang, X. (2012). Genome-wide identification and analysis of the TIFY gene family in grape. *PLoS One* 7:e44465. doi: 10.1371/journal.pone.0044465

Zhang, X., Zheng, W., Cao, X., Cui, X., Zhao, S., Yu, T., et al. (2019). *GmSAP16* genomic analysis of stress associated proteins in soybean and the role of in abiotic stress responses in soybean. *Front. Plant Sci.* 10:1453. doi: 10.3389/fpls.2019.01453

Zhao, C., Pan, X., Yu, Y., Zhu, Y., Kong, F., Sun, X., et al. (2020a). Overexpression of a TIFY family gene, *GsJAZ2*, exhibits enhanced tolerance to alkaline stress in soybean. *Mol. Breed.* 40. doi: 10.1007/s11032-020-01113-z

Zhao, G., Song, Y., Wang, Q., Yao, D., Li, D., Qin, W., et al. (2020b). *Gossypium hirsutum* salt tolerance is enhanced by overexpression of *JAZ1*. *Front. Bioeng. Biotechnol.* 8:157. doi: 10.3389/fbioe.2020.00157

Zhao, S., Xu, Z., Zheng, W., Zhao, W., Wang, Y., Yu, T., et al. (2017). *GmRAV03* genome-wide analysis of the RAV family in soybean and functional identification of involvement in salt and drought stresses and exogenous ABA treatment. *Front. Plant Sci.* 8:905. doi: 10.3389/fpls.2017.00905

Zhu, D., Bai, X., Luo, X., Chen, Q., Cai, H., Ji, W., et al. (2013). Identification of wild soybean (*Glycine soja*) TIFY family genes and their expression profiling analysis under bicarbonate stress. *Plant Cell Rep.* 32, 263–272. doi: 10.1007/s00299-012-1360-7

Zhu, D., Li, R., Liu, X., Sun, M., Wu, J., Zhang, N., et al. (2014). The positive regulatory roles of the TIFY10 proteins in plant responses to alkaline stress. *PLoS One* 9:e111984. doi: 10.1371/journal.pone.0111984

Zhu, J. K., Hasegawa, P. M., Bressan, R. A., and Bohnert, H. J. (2011). Molecular aspects of osmotic stress in plants. *Crit. Rev. Plant Sci.* 16, 253–277. doi: 10.1080/07352689709701950

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 31 March 2022
doi: 10.3389/fpls.2022.798118

Comparative Transcriptome Analysis Reveals New Insight of Alfalfa (*Medicago sativa* L.) Cultivars in Response to Abrupt Freezing Stress

Xia Wang, Wenjuan Kang, Fang Wu, Jiamin Miao and Shangli Shi**

College of Grassland Science, Gansu Agricultural University, Lanzhou, China

Freezing stress is a major limiting environmental factor that affects the productivity and distribution of alfalfa (*Medicago sativa* L.). There is growing evidence that enhancing freezing tolerance through resistance-related genes is one of the most efficient methods for solving this problem, whereas little is known about the complex regulatory mechanism of freezing stress. Herein, we performed transcriptome profiling of the leaves from two genotypes of alfalfa, freezing tolerance "Gannong NO.3" and freezing-sensitive "WL326GZ" exposure to −10°C to investigate which resistance-related genes could improve the freezing tolerance. Our results showed that a total of 121,366 genes were identified, and there were 7,245 differentially expressed genes (DEGs) between the control and treated leaves. In particular, the DEGs in "Gannong NO.3" were mainly enriched in the metabolic pathways and biosynthesis of secondary metabolites, and most of the DEGs in "WL326GZ" were enriched in the metabolic pathways, the biosynthesis of secondary metabolites, and plant-pathogen interactions. Moreover, the weighted gene co-expression network analysis (WGCNA) showed that ATP-binding cassette (ABC) C subfamily genes were strongly impacted by freezing stress, indicating that *ABCC8* and *ABCC3* are critical to develop the freezing tolerance. Moreover, our data revealed that numerous Ca^{2+} signal transduction and CBF/DREB1 pathway-related genes were severely impacted by the freezing resistance, which is believed to alleviate the damage caused by freezing stress. Altogether, these findings contribute the comprehensive information to understand the molecular mechanism of alfalfa adaptation to freezing stress and further provide functional candidate genes that can adapt to abiotic stress.

Keywords: ***Medicago sativa*, freezing stress, transcriptome, ABC gene family, Ca^{2+} signal transduction, CBFs/DREB1s**

OPEN ACCESS

Edited by:
Rohini Garg,
Shiv Nadar University, India

Reviewed by:
Jing Bo Jin,
Institute of Botany (CAS), China
Parviz Heidari,
Shahrood University of Technology, Iran

****Correspondence:***
Jiamin Miao
miaojm@gsau.edu.cn
Shangli Shi
shishl@gsau.edu.cn

Specialty section:
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

Received: *19 October 2021*
Accepted: *14 March 2022*
Published: *31 March 2022*

Citation:
Wang X, Kang W, Wu F, Miao J and Shi S (2022) Comparative Transcriptome Analysis Reveals New Insight of Alfalfa (Medicago sativa L.) Cultivars in Response to Abrupt Freezing Stress. Front. Plant Sci. 13:798118. doi: 10.3389/fpls.2022.798118

INTRODUCTION

Alfalfa (*Medicago sativa* L.) is one of the most important cultivated perennial forage legume species in the world (Yang et al., 2010), and it is grown extensively in different temperature zones based on its high biomass yield, rich nutritional value, good palatability, high capacity for nitrogen fixation, and strong ecological adaptability (Chao et al., 2019). However, under

extreme environmental conditions, freezing stress is a major factor that substantially attenuates alfalfa growth, development, productivity, and distribution (Song et al., 2016; Zhou et al., 2018; Castonguay et al., 2020), especially in northern cold regions (Anower et al., 2016). Moreover, unusual abrupt temperature changes lead to serious economic losses in the winter and later spring frost events (Zhou et al., 2018). Therefore, clarifying the mechanisms responsible for freezing tolerance during periods of abrupt freezing stress is important for the breeding of novel alfalfa varieties.

Freezing stress (below 0°C) initially causes ice formation in the cell wall, which directly affects cellular metabolic activities. However, as the ice crystals grow, water uptake by the call is reduced leading to cellular acute dehydration and severe damage to the cell membrane (Ahmed et al., 2015). Oxidative stress severely affects the activities of enzymes in plants, such as reactive oxygen species (ROS) scavenging system enzymes. Furthermore, freezing stress induces instability of protein complexes and RNA secondary structures in cell physiology. Finally, plants perish because freezing stress damages the photoinhibition and destroys the metabolic balance in plants (Gong et al., 2020). To adapt to freezing stress, plants have evolved a series of strategies to increase their freezing tolerance (Tsutsui et al., 2009). However, the freezing tolerance level is a diversiform change in physiological, biochemical, molecular, and morphological characteristics in different plant species (Eremina et al., 2016). ROS are important biochemical changes that severely affect plant growth. Plant cells scavenge ROS to induce related systems, including superoxide dismutase (SOD), catalase (CAT), peroxidase (POD), and ascorbate peroxidase (APX; Karimzadeh et al., 2021). Enhanced antioxidant activities increase freezing tolerance in many plants, such as maize (*Zea mays* L.; Hodges et al., 1997), rice (*Oryza sativa* L.; Huang and Guo, 2005), chickpeas (*Cicer arietinum* L.; Kaur et al., 2009), and alfalfa (Wang et al., 2009).

In addition, freezing stress can trigger rapid increase in calcium (Ca^{2+}) in the cytosol (Gong et al., 2020). Ca^{2+}, as a second messenger, strongly affects primary low-temperature stress signal transduction (Monroy and Dhindsa, 1995) and decoded by downstream effector proteins, including calmodulins (CaMs), CaM-like proteins (CMLs), calcium-dependent protein kinases (CDPKs), and calcineurin B-like proteins (CBLs), to generate related responses to stress (Zhu et al., 2015). CaM is present in all intracellular events in plants; it is a ubiquitous Ca^{2+} sensor protein. Moreover, CaMs interact with more than 300 target proteins and modulate their activities to defend against environmental stress (Yang and Poovaiah, 2003). Recent studies have demonstrated that many CaMs/CML genes are induced by freezing stress to improve the freezing tolerance, such as *NpCaM-1* (Zeng et al., 2015), *AtCaM3* (Townley and Knight, 2002), *AtCML24* (Delk et al., 2005), *AtCML10* (Cho et al., 2016), *CsCML16*, *CsCML42* (Ma et al., 2019), *MtCML24*, and *ShCML44* (Sun et al., 2021). In contrast, CDPKs directly modulate Ca^{2+} signals at the cytosolic level (Mall et al., 2011) to involved in resistance to abiotic stresses, including cold stress, and many CDPK-related genes can be induced under freezing stress to adapt to changes in temperature (Komatsu et al., 2007; Ray et al., 2007; Dubrovina et al., 2015). However, the mechanisms of CDPK-positive or negative regulated resistance to freezing stress are not clear. In *Populus euphratica, PeCPK10* positively regulates freezing tolerance (Chen et al., 2013). However, *ZmCPK1* in maize negatively regulates cold tolerance through the Ca^{2+} signaling pathways (Weckwerth et al., 2015). Similar to the CBL genes, multiple CBL genes conferring stress have been identified in different plant species (Kolukisaoglu et al., 2004; Li et al., 2012; Sun et al., 2015; Zhang et al., 2016; Wang et al., 2020). Several genes related to CBLs, such as *BrCBL1* (Lee et al., 2013), *PeCBL6*, and *PeCBL10* are upregulated in response to cold stress (Li et al., 2012) and can increase *CBF3*, *COR15A*, and *COR47A* expression levels to confer freezing stress (Zhou et al., 2016). However, Cheong et al. (2003) reported that *AtCBL1* is downregulated under cold stress inhibiting the expression of *AtCBF3* to cope with environmental changes. In addition to Ca^{2+} signal transduction, which is important in freezing stress, the mitogen-activated protein kinase (MPK) cascades and phytohormones also play a key role in the freezing signal process (Li et al., 2017; Liu et al., 2020). To mitigate freezing stress, several transcription factors are activated, regulating the expression of downstream cold-regulated (COR) genes (Peng et al., 2021).

CBF/DREB1 (C-repeat binding factors/Dehydration responsive element binding protein 1) is an important signal transduction pathway involved in freezing tolerance mechanism. CBFs, including CBF1, CBF2, and CBF3, are the AP2/ERF family transcription factors, which regulate COR genes to confer freezing stress. Accordingly, overexpressed CBFs increase the freezing tolerance in diverse plant species (Song et al., 2016). In Arabidopsis, *CBF1* and *CBF3* positively regulate freezing tolerance (Tsutsui et al., 2009), and overexpressing *MtCBF3* could enhance freezing tolerance in *Medicago truncatula* (Tayeh et al., 2013). Moreover, Shu et al. (2017) observed nine CBF unigenes of alfalfa homology at the *Mt-FTQTL6* site, which positively regulated freezing tolerance. Additionally, Kanchupati et al. (2017) suggested that CBFs may strongly impact the freezing tolerance of alfalfa. However, because the alfalfa genome has not been fully sequenced, the molecular mechanism of CBF cluster responses to freezing stress is not clear.

Recently, with the development of high-throughput RNA-sequencing (RNA-seq) technology applications, the mechanisms of freezing stress response in numerous plants have been revealed, and thousands of freezing tolerance-related genes are involved in signal transduction. These studies suggest that transcription factors regulate gene expression to defend against freezing stress (Nah et al., 2016). However, little is known about the response to abrupt freezing stress in alfalfa (Song et al., 2016). In this study, we utilized RNA-seq to evaluate changes in freezing stress response genes in leaves of "Gannong NO.3" (freezing tolerance) and "WL326GZ" (freezing sensitive) with abrupt freezing stress based on the referenced alfalfa genome (Chen et al., 2020), and identified several genes that may strongly impact the freezing stress response in alfalfa. These results elucidate the mechanism of freezing stress response and can be used to compare the freezing tolerance and freezing

sensitivity in alfalfa. This study will provide valuable resources for practitioners performing freezing tolerance molecular breeding in alfalfa.

MATERIALS AND METHODS

Plant Materials and Culture Conditions

Two freezing tolerance alfalfa cultivars were studied. The first was the freezing-tolerance cultivar "Gannong NO.3" (fall dormancy score of 3.0), which is produced in Gansu, Ningxia, Xinjiang, and eastern margin of the Qinghai-Tibet plateau. Its seeds were provided by Gansu Agricultural University. The second was the freezing-sensitive "WL326GZ" (fall dormancy score 3.8), which was purchased from Zhengdao Ecological Technology Co. (Beijing, China). All seeds were sown on arenaceous quartz in a growth room, and each sample with three times. The growth conditions were as follows: 200 μmol m^{-2} s^{-1}, a light period of 14 h at 20°C, with 10 h dark at 18°C, and humidity ranging from 60 to 80%. The seedlings were irrigated with half-strength Hoagland solution daily after the seeds germinated. Four weeks later, the seedlings were transferred to a cold chamber at −10°C. All leaves were harvested at 0, 0.5, 1, and 2 h after freezing stress. Detailed sample names are as follows: G01-03, G11-13, G21-23, and G31-33 represent three "Gannong NO.3" samples harvested at 0, 0.5, 1, and 2 h, respectively; W01-03, W11-13, W21-23, and W31-33 represent three "WL326GZ" samples harvested at 0, 0.5, 1, and 2 h, respectively. Thus, every sample had three replicates, and the leaves were frozen in liquid nitrogen and stored at −80°C until use.

Enzyme Extraction and Assays

To investigate the contents of four types of antioxidant activities (superoxide-SOD, peroxidase-POD, ascorbate peroxidase-APX, and catalase-CAT), we used 0.1 g fresh leaves which were homogenized in 1.5 ml potassium phosphate buffer (50 mM, pH 7.0) and centrifuged for 20 min at 12,000 r $min^{-1}$with 4°C for obtaining the upper supernatant to detect the antioxidant activities. The SOD activity was measured according to method reported by Giannopolitis et al. (1977), POD activity was determined using the guaiacol oxidation (Chance and Maehly, 1955), and CAT and APX activities were determined based on decomposition of H_2O_2 (Chance and Maehly, 1955).

RNA Extraction and Illumina Sequencing

Total RNA was isolated from leaf tissues using the RNAprep Pure Plant Kit (TIANGEN Biotech, Beijing, China) according to the manufacturer's instructions and then treated with DNase I (NEB) to remove DNA. Quantification was performed using a Qubit RNA Assay Kit and a Qubit 2.0 Fluorimeter (Life Technologies, CA, United States). To ensure the integrity of the RNA, we used an RNA Nano 6000 Assay Kit of the Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, United States) to library construction and sequencing.

Twenty-four libraries were constructed using the NEBNext Ultra RNA Library Prep Kit for Illumina (NEB, United States). Sequences were conducted by the Biomarker Technology Company (Beijing, China) on the Illumina Hiseq™ 4,000 platform (Illumina, San Diego, CA, United States). The raw reads were first processed through in-house Perl scripts, and clean reads were obtained by moving adapter ploy-N and low-quality reads. The clean reads were mapped to the ribosome RNA (rRNA) database using the short reads alignment tool Bowtie2 (version 2.2.8), the rRNA mapped reads were removed, and the remaining clean reads were mapped to the *M. sativa* L. reference genome[1] using HISAT2.2.4. For further analysis, all transcripts were annotated in the Nr, Nt, Swiss-Prot, Pfam, KOG/COG, KEGG, and GO databases.

Differential Expression Analysis

Differentially expressed genes (DEGs) were analyzed using pairwise samples (G01_G02_G03 vs. G11_G12_G13, G01_G02_G03 vs. G21_G22_G23, G01_G02_G03 vs. G31_G32_G33, W01_W02_W03 vs. W11_W12_W13, W01_W02_W03 vs. W21_W22_W23, and W01_W02_W03 vs. W31_W32_W33) by DEseq R package (1.18.0). The expression levels were adjusted with FDR < 0.05, and $|\log_2$ (fold change) $| \geq 1$ as a significantly differential expression.

Trend Analysis of DEGs in "Gannong NO.3" and "WL326GZ"

To under the DEG trends, the DEGs of each variety at each sampling time were clustered by similar expression patterns. Briefly, the expression of each DEG was input, normalized, and clustered using Short Time-series Expression Miner software (STEM) with $|\log_2$ (fold change)$| \geq 1$, $p \leq 0.05$, and profile number ≤20. Then, the DEGs of each profile were annotated using the Gene Ontology (GO) and KEGG databases analysis using a $Q \leq 0.05$.

Weighted Gene Co-expression Network Analysis

The co-expression network was performed using the weighted gene co-expression network analysis (WGCNA) in the R software package (Langfelder and Horvath, 2008). A total of 23,977 genes with FPKM values above nine were selected for the WGCNA unsigned analysis. The adjacency matrix was constructed with a soft threshold power of six by pairwise Pearson's correlation coefficients between pairs of genes. Then, a dissimilarity measure based on the topological overlap matrix (TOM) was used to cluster genes into the network modules on their co-expression. The module eigengene E (as the first principal component of a given module) was constructed and used as a weight to calculate the correlation of the module. The networks were visualized using Cytoscape v.3.8.0 to analyze the key modules and hub genes.

[1]https://figshare.com/projects/whole_genome_sequencing_and_assembly_of_Medicago_sativa/66380

RESULTS

Phenotypic Changes and Responses of Antioxidant Enzymes Following Exposure to Freezing Stress

To distinguish the two different freezing tolerance types of alfalfa, we analyzed the antioxidant enzymes and phenotypic changes in the leaves following exposure to freezing stress for 0, 0.5, 1, and 2 h. The alfalfas in "Gannong NO.3" were survived more than in the "WL326GZ" under the direct exposure to −10°C temperatures for 2 h (**Figure 1A**). To further exhibit the phenotypic changes, we recovered the alfalfas at a normal growth temperature for 2 days (**Figure 1B**). Interestingly, 3.33% of the "WL326GZ" were survived, whereas 9.33% in "Gannong NO.3" (**Figure 1B**), indicating a higher tolerance in "Gannong NO.3," comparing to "WL326GZ."

In order to detect the effects on the freezing stress, we analyze the SOD, POD, APX, and CAT activities of the alfalfas. As shown in **Figure 2A**, there was no significant change of the SOD activity in "WL326GZ" under freezing stress from 0 to 1 h exposure and dramatically reduced 70.11% at 2 h, whereas the SOD activity of "Gannong NO.3" was decreased at 0.5 h and then increased 7.51% at 2 h exposure to freezing stress. The change in POD activity was similar between the "WL326GZ" and "Gannong NO.3," both of which decreased as the freezing stress time increased, and naturally increased at 2 h. Importantly, "Gannong NO.3" has a greater extent of POD activity to "WL326GZ" at 2 h of freezing stress (**Figure 2B**). The variation in APX activity was similar to POD, but that of "WL326GZ" increased at 1 h, and that of "Gannong NO.3" increased at 2 h. The APX activity of "Gannong NO.3" was therefore higher than in "WL326GZ" at 2 h (**Figure 2C**). The CAT activity of the two types of alfalfa cultivars increased consistently with the duration of the freezing stress time, and the CAT activity of "Gannong NO.3" was higher than that of "WL326GZ" (**Figure 2D**). Taken together, the antioxidant activities of the sensitive cultivar "WL326GZ" occurred earlier in contrast to the tolerant type of "Gannong NO.3."

Bioinformatic Analysis of Differential Expressed Genes in "Gannong NO.3" and "WL326GZ"

To investigate which resistance-related genes can be beneficial to improve the freezing stress, we used RNA-seq approach to profile the leaves of "Gannong NO.3" and "WL326GZ" at 0, 0.5, 1, and 2 h exposure to −10°C. Our preliminary RNA-seq screen generated 155,133,490,050 raw data points. The raw data from each sample consisted of more than 59,000,000 data points, and all raw data were deposited in NCBI with the accession number PRJNA769225. Importantly, we cleared the raw data with GC < 43%, Q20 > 97%, and Q30 > 94%, and finally obtained 153,834,608,062 data points (**Supplementary Table S1**). To remove ribosome RNA (rRNA), we used Bowtie 2 to map the cleared data points to the (rRNA) database and got the clean reads ranged from 36,671,116 to 45,294,426 (**Supplementary Table S2**). We then mapped the clean reads to the alfalfa reference genome (Chen et al., 2020) and finally identified 121,366 genes from more than 89% of the mapped reads (see **Supplementary Tables S3, S4**).

We then constructed a comparative analysis between freezing stress treated and control leaves of "Gannong NO.3" or "WL326GZ." There were 7,245 DEGs that showed significant upregulation/downregulation, among that, the G0 vs. G2 and W0 vs. W2 displayed 3,892 DEGs and 2,608 DEGs, respectively (**Figure 3**).

The Trend Analysis and GO and KEGG Functional Annotation of DEGs

To cluster the same trend expression genes, 20 profiles were generated using the trend analysis of "Gannong NO.3" and the top 6 profiles of the gene numbers were, 8, 11, 16, 18,

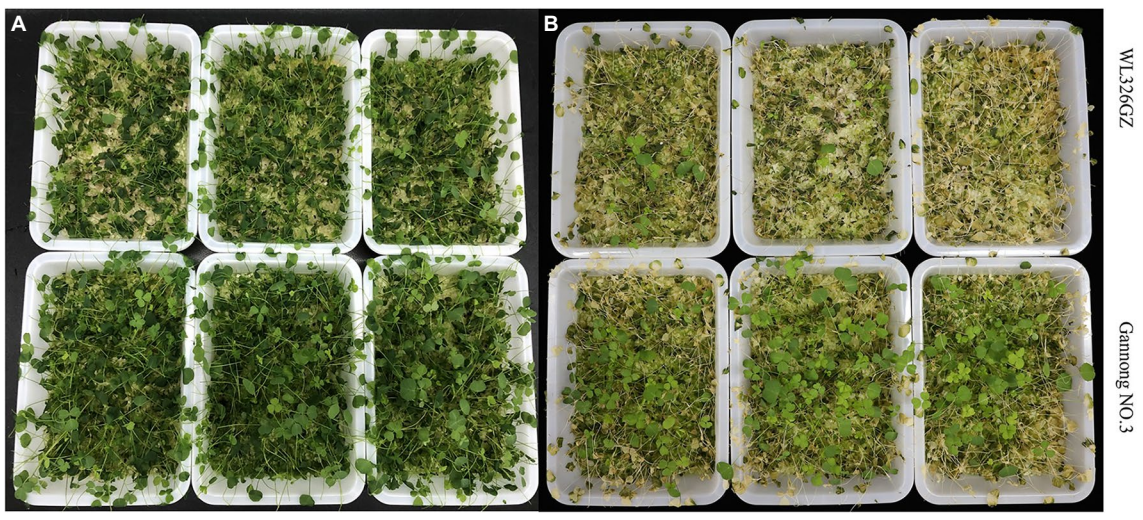


FIGURE 1 | Phenotypes of two alfalfa cultivars after exposure to −10°C. **(A)** The phenotype of "WL326GZ" and "Gannong NO.3" exposure to −10°C for 2 h. **(B)** The phenotype of the **(A)** growth in the normal temperature for 2 days.

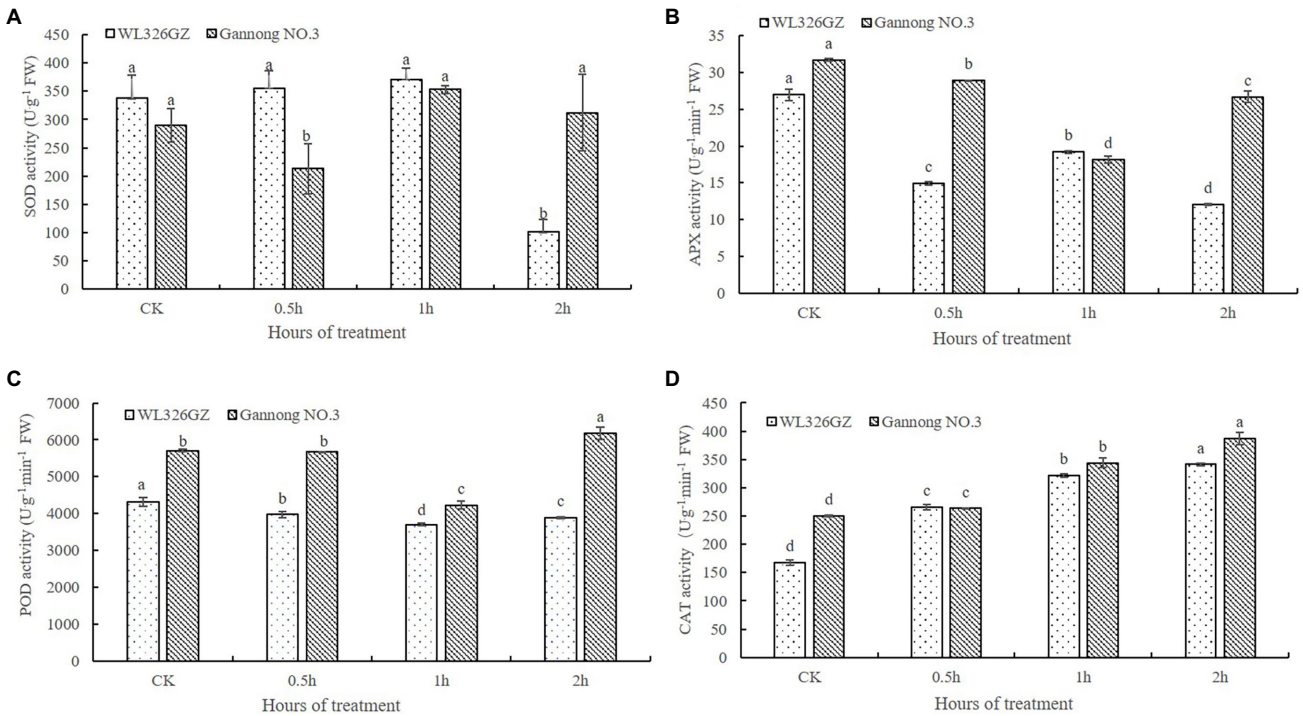


FIGURE 2 | Effect of freezing stress on the activities of SOD **(A)**, APX **(B)**, POD **(C)**, and CAT **(D)** in leaves of two types of alfalfa. The data represents mean ± SD ($n = 3$), and the different letters above the bar represents the values were significant difference in "WL326GZ" or "Gannong NO.3" with different freezing treatment ($p < 0.05$, Duncan's test).

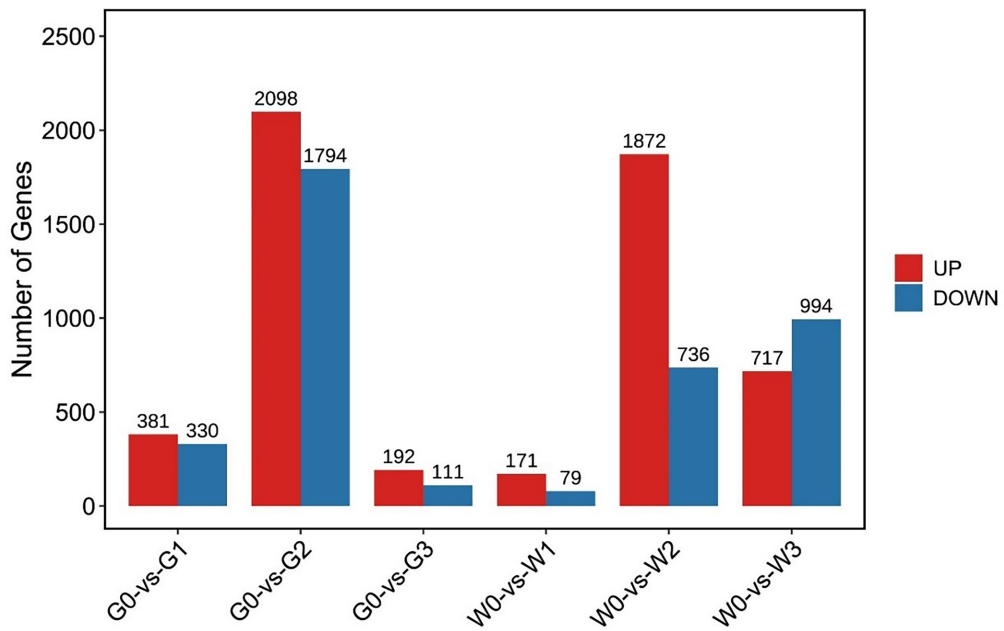


FIGURE 3 | Number of differentially expressed genes, either upregulated and downregulated, in alfalfa leaves following exposure to freezing stress. G0, G1, G2, and G3 represent the samples of "Gannong NO.3" exposed to freezing stress for 0, 0.5, 1, and 2 h, respectively. W0, W1, W2, and W3 represent the samples of "WL326GZ" exposed to freezing stress for 0, 0.5, 1, and 2 h, respectively.

3, and 1 (**Figure 4A**). We then utilized the GO and KEGG pathway to identify the biological functions of the profiles. The top 20 GO enrichment (**Supplementary Table S5**) of the profile 8 showed 571 (81.34%) DEGs enriched in the

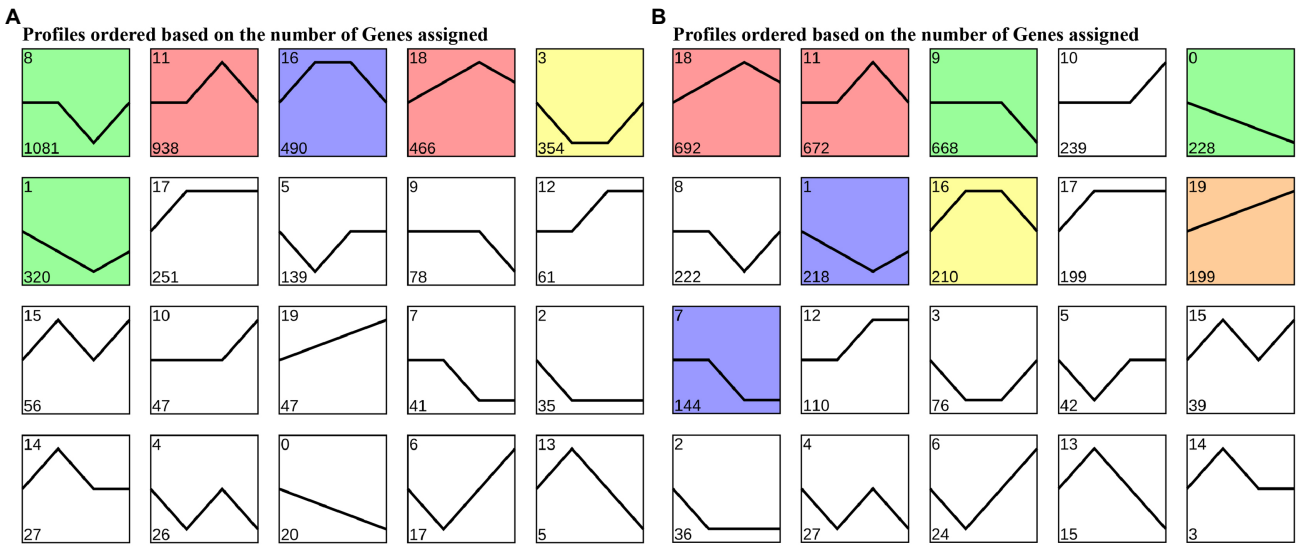


FIGURE 4 | Trends in expression levels of differentially expressed genes in leaves of "Gannong NO.3" **(A)** and "WL326GZ" **(B)** in response to freezing stress.

GO:0044699 (single-organism process) and 535 (78.68%) DEGs enriched in the GO:0003824 (catalytic activity). Interestingly, profile 11 showed the most DEGs (400, 61.73%) enriched in the GO:0050896 (response to stimulus), followed by GO:0042221 (response to chemical: 270, 41.67%). As in profile 16, most DEGs (400, 61.73%) were enriched in GO:0005737 (cytoplasm), followed by GO:0044444 (cytoplasmic part). Moreover, the KEGG pathway (**Supplementary Figure S1**) analyses were performed for most DEGs enriched in the ko01100 (metabolic pathways) and ko01110 (biosynthesis of secondary metabolites) of profiles 8, 11, and 16.

Similar to "Gannong NO.3.," there were 20 profiles were produced for "WL326GZ" by trend analyzed and the top 6 profiles of the DEGs were, 18, 11, 9, 10, 0, and 8 (**Figure 4B**). The results of GO enrichment show that 291 (67.67%) DEGs were enriched in GO:0050896 (response to stimulus), 219 (50.93%) DEGs were enriched in GO:0065007 (biological regulation) of profile 18, 264 (66.17%) DEGs were enriched in GO: 0050896 (response to stimulus), 204 (51.13%) DEGs were enriched in GO:0065007 (biological regulation) of profile 11, 315 (77.40%) DEGs were enriched in GO:0071704 (organic substance metabolic process), and 307 (75.43%) DEGs were enriched in GO:0044237 (cellular metabolic process) of profile 9 (**Supplementary Table S6**). The KEGG pathway (**Supplementary Figure S2**) analysis revealed that the results of profiles 11 and 9 coincided with those of "Gannong NO.3"; most DEGs in profile 18 were enriched in ko01110 (biosynthesis of secondary metabolites) and ko04626 (plant-pathogen interaction).

Construction of Weight Co-expression Network and Identification of Key Modules at Different Times

To construct the gene co-expression network of RNA-seq, we obtained 23,977 genes from the clean data for WGCNA. As shown in **Figure 5A**, 20 modules were generated and each module was independently by analyzed with eigengene adjacency heat map (**Figure 5C**). Moreover, the most genes from WGCNA were enriched in the turquoise module (6,336 genes, **Figure 5B**), followed by the light green module (2,856 genes, **Figure 5B**).

To identify the different reactions between "Gannong NO.3" and "WL326GZ" to the abrupt freezing stress, we obtained two modules with opposite expression patterns in those modules at different times (**Figure 6**). Obviously, the dark red and midnight blue modules showed opposite expressions after exposure to freezing stress.

Function Enrichment Analysis of the Two Key Modules

To analyze the biological function of the genes from dark red and midnight blue modules, we utilized GO enrichment and KEGG pathways to classify the genes. The results of top 20 GO showed that the genes in the dark red (**Supplementary Table S7**) module were enriched in GO:0044699 (single-organism process), GO:0044763 (single-organism cellular process), GO:0044710 (single-organism metabolic process), GO:0050896 (response to stimulus), and GO:0044444 (cytoplasmic part), which coincide with **Supplementary Table S5**. The genes of the midnight blue module were enriched in GO:0005737 (cytoplasm), GO:0044444 (cytoplasmic part), GO:0044710 (single-organism metabolic process), GO:0044281 (small molecule metabolic process), and GO:0009536 (plastid). Importantly, the KEGG pathway (**Supplementary Figure S3**) analysis showed that most genes in the two modules were enriched in ko01100 and ko01110, which is consistent with the results of KEGG pathway in "Gannong NO.3" (**Supplementary Figure S1**).

Validation of the Hub Genes of DEGs in Two Modules of the Key Pathway in KEGG

To validate the hub genes as DEGs, we set the cutoff as $|\log_2(\text{fold change})| \geq 1$ and $Q < 0.05$ to screen the DEGs. There

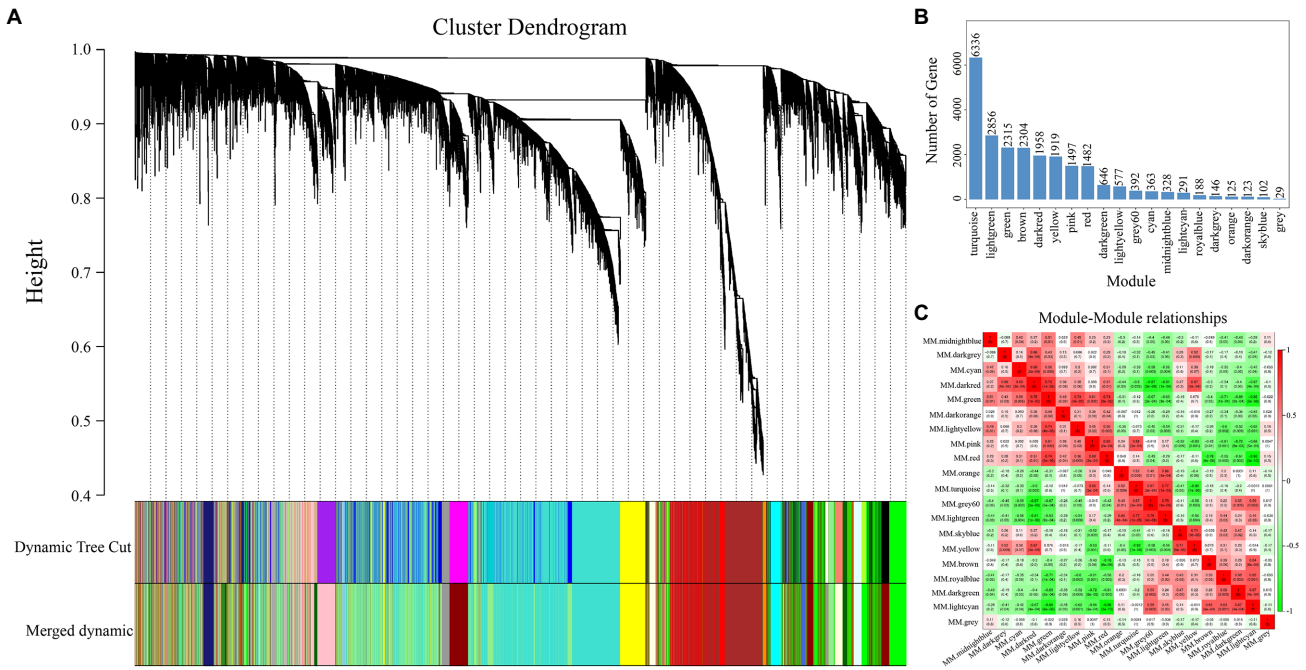


FIGURE 5 | Details of the modules constructed *via* weighted gene co-expression network analysis (WGCNA). **(A)** The cluster dendrogram of the 23,977 genes. **(B)** The number of genes in each module which construct by **(A)**. **(C)** The heat map of correlation among all modules.

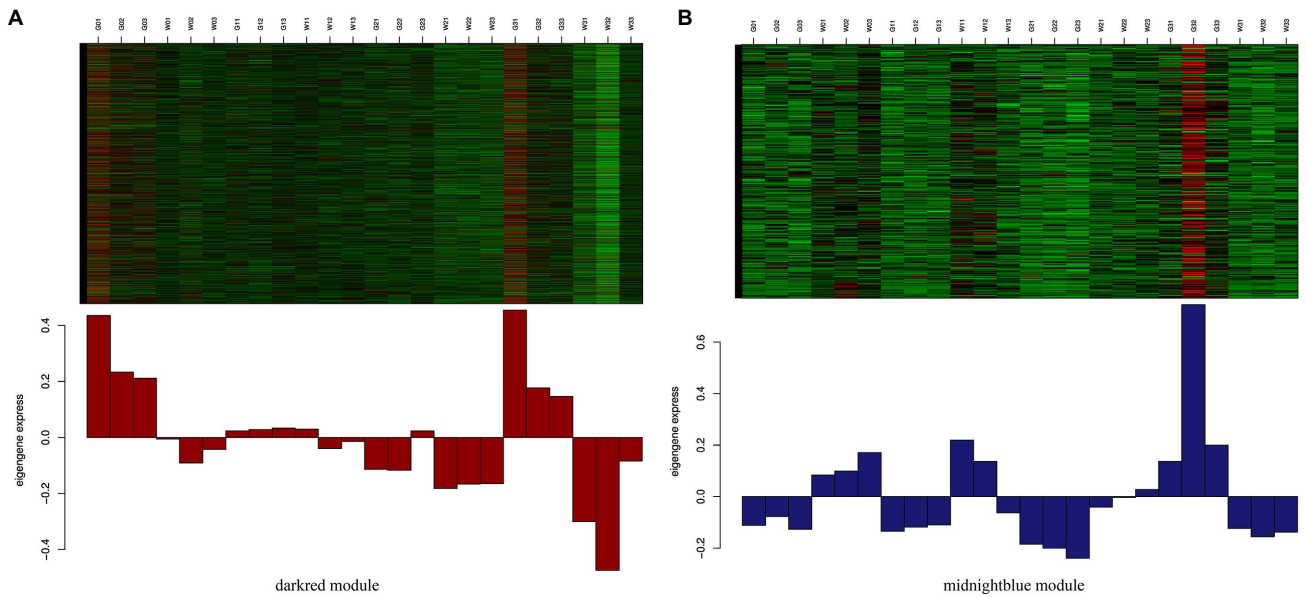


FIGURE 6 | Expression levels of genes from darkred module **(A)** and midnightblue module **(B)**. Those between the darkred and midnightblue modules are most significantly correlated with "Gannong NO.3" and "WL326GZ" following exposure to freezing stress.

were 7,245 DEGs identified in **Figure 7A**, and we overlapped the DEGs and the genes in the two modules. A total of 338 genes were observed between all DEGs and the dark red module, furthermore we chose the top 10 genes with relatively high connectivity values in the module as candidate hub genes (all.kWithin value, **Figure 7B; Table 1**). The results of the hub genes demonstrated that MS.gene011170 was one of the most important genes in the dark red module with highest connectivity value; its gene name is *ABCC8* [ATP-binding cassette (ABC) transporter C family member 8], which belongs

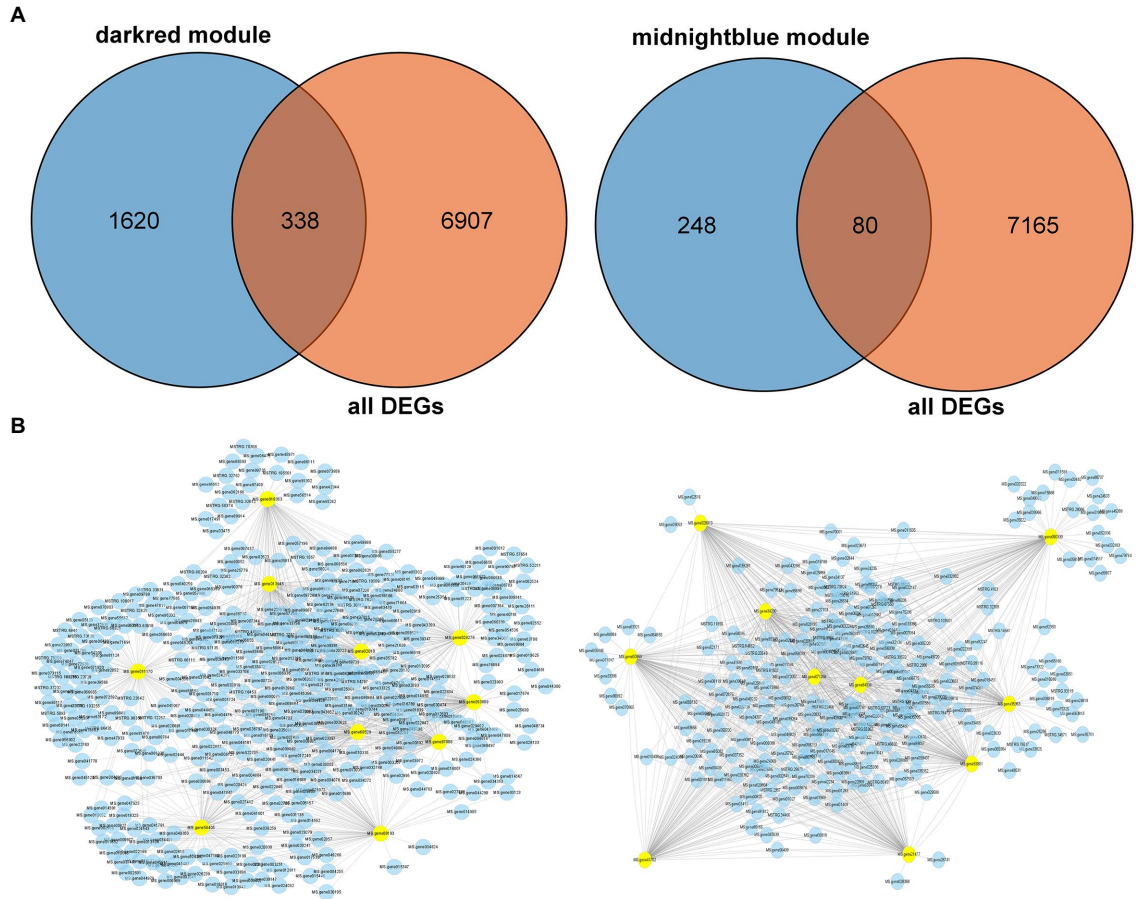


FIGURE 7 | Identification the hub genes of differentially expressed genes (DEGs) in the two key modules. **(A)** Venn diagram between the two modules and all DEGs. **(B)** Correlation networks of hub genes corresponding to the darkred and midnightblue modules, respectively. One hundred genes of edge weight in the top 10 hub genes are visualized using Cytoscape.

to the ABC transporter C subfamily. MS.gene68529 (ABC transporter C family member 3 isoform X2, *ABCC3*), as another important candidate hub gene, also belongs to the ABC transporter C family, which indicates that the ABC transporter C subfamily may play an important role in freezing stress. However, there were only 80 DEGs identified between the genes of entire set and the midnight blue module (**Figure 7A**). And the top 10 genes as candidate hub genes which based on its top 10 of all.kWithin values showed that MS.gene00859 (BGLU47, beta-glucosidase 46) was one of the most critical genes activated during freezing stress (**Figure 7B; Table 1**).

Freezing Stress Induces Numerous DEGs in the ABC Family

The ABC transporter family is one of the largest transporter families in living organisms because of its important physiological functions (Nguyen et al., 2014). In our study, totally of 80 DEGs belonged to the ABC gene family, and they were grouped into seven subfamilies (**Figure 8**). A heat map of 80 DEGs indicates (**Figure 8**) that four genes (5.00%) belong to ABC subfamily A, 11 genes (13.75%) belong to ABC subfamily B, 39 genes (48.75%) belong to ABC subfamily C, one gene (1.25%) belongs to ABC subfamily E, five genes (6.25%) belong to ABC subfamily F, 17 genes (21.25%) belong to ABC subfamily G, and three genes (3.75%) belong to ABC subfamily I. Obviously, ABC subfamily C plays a critical role in freezing stress response.

Furthermore, we also compared the 39 genes of ABC subfamily C to the database of the WGCNA, which indicated that there were 17 genes included in the database (**Table 2**). And analyzed the relatively high connectivity values of 17 genes, the results indicated that the highest values of genes were MS.gene011170 and MS.gene68529 that could be acted as hub genes of ABC subfamily C in all modules (All.kTotal value), which coincides with the results of **Table 1**.

Expression Profiles of the Putative Complex Network of Signal Transduction

Freezing stress could cause osmotic stress in plant cells and rapidly trigger Ca^{2+} signal transduction, which could induce

several Ca^{2+} signal transduction-related genes expressed (Gong et al., 2020). In this study, CaMs/CMLs, CAMATs CBLs, CIPKs, CPKs, and MAPKs were induced during the freezing stress (**Figure 9**). Furthermore, most of the DEGs of the CaMs/CMLs-CAMATs increased over the freezing time and then decreased. And freezing stress induced the expression of most genes in G2 and W2. In contrast, only three (MS.gene036729, MS.gene047644, and MS.gene074100) decreased from G0 to G2 and W0 to W1 as well as increased to G3 and W3. Interestingly, the similar change of CBLs-CIPK for "Gannong NO.3" showed a kick point at G2, but that of "WL326GZ" was at W1 Whereas, there only were three DEGs belonged to CPKs. And all DEGs consistently increased at G2 and W2 and then decreased. Furthermore, CPKs could activated the MAPK cascade expressed to regulate the cold responses in plants. In our study, the changes of the MS.genes 56787, 99208, 011812, 56784, 56785, and 90503 genes belongs to MAPK casacade which were decreased from G0 to G2 and increased at G3; whereas in "WL326GZ," the genes were increased from W0 to W2 and then decreased at W3. And the variation of the other DEGs was similar between "Gannong NO.3" and "WL326GZ." Importantly, the DEGs of "WL326GZ" showed a greater expression than those of "Gannong NO.3." And all the DEGs of CaMs-CAMTAs, CBLs-CIPKs, and the MAPK cascade were induced *via* CBFs/DREB1s (C-repeat binding factors/

TABLE 1 | The details of 10 hub genes from the two key modules.

Gene name	All.kWithin	Description
Darkred module		
MS.gene011170	148.03	ABC transporter C family member 8
MS.gene013945	143.24	4-alpha-glucanotransferase DPE2
MS.gene54405	142.70	Clathrin interactor EPSIN 3
MS.gene87888	138.31	Very-long-chain aldehyde decarbonylase CER3-like isoform X1
MS.gene053889	136.06	Peptide/nitrate transporter
MS.gene09193	130.04	Lipoxygenase 6, chloroplastic
MS.gene024276	123.06	Light-regulated protein 1, chloroplastic
MS.gene018353	120.55	Pathogenesis-related protein PR-1
MS.gene68529	119.15	ABC transporter C family member 3 isoform X2
MS.gene02010	118.06	Ferric reduction oxidase 7, chloroplastic
Midnightblue module		
MS.gene00859	24.24	Beta-glucosidase 46
MS.gene94034	21.13	D-pinitol dehydrogenase
MS.gene41702	19.69	NAC transcription factor 47
MS.gene53801	18.02	B2 protein
MS.gene071258	17.61	Serine protease SPPA, chloroplastic
MS.gene35363	16.41	Histone deacetylase HDT1
MS.gene21477	14.10	Cytochrome P450 family ent-kaurenoic acid oxidase
MS.gene028913	13.53	E3 ubiquitin-protein ligase RMA1H1
MS.gene060330	13.09	Pyrophosphate-energized vacuolar membrane proton pump
MS.gene34230	12.80	Probable aminotransferase TAT2 isoform X1

TABLE 2 | The results of overlapped genes between ABCC subfamily genes and differentially expressed genes (DEGs).

Gene ID	Module	All.kTotal
MS.gene011170	Darkred	719.0367653
MS.gene68529	Darkred	616.0700359
MS.gene006196	Green	585.8495929
MS.gene59189	Green	557.947855
MS.gene00301	Green	541.9193836
MS.gene066239	Lightyellow	541.5092946
MS.gene52055	Lightyellow	487.9046613
MS.gene52408	Green	464.8650029
MS.gene34800	Darkred	452.5614335
MS.gene046329	Lightyellow	390.2598191
MS.gene81377	Turquoise	357.9430749
MS.gene005606	Dark red	305.3770089
MS.gene039515	Lightyellow	294.0799577
MS.gene97154	Lightyellow	269.936822
MS.gene72055	Darkred	144.6371911
MS.gene66210	Darkred	91.06072232
MS.gene002714	Lightgreen	75.9891247

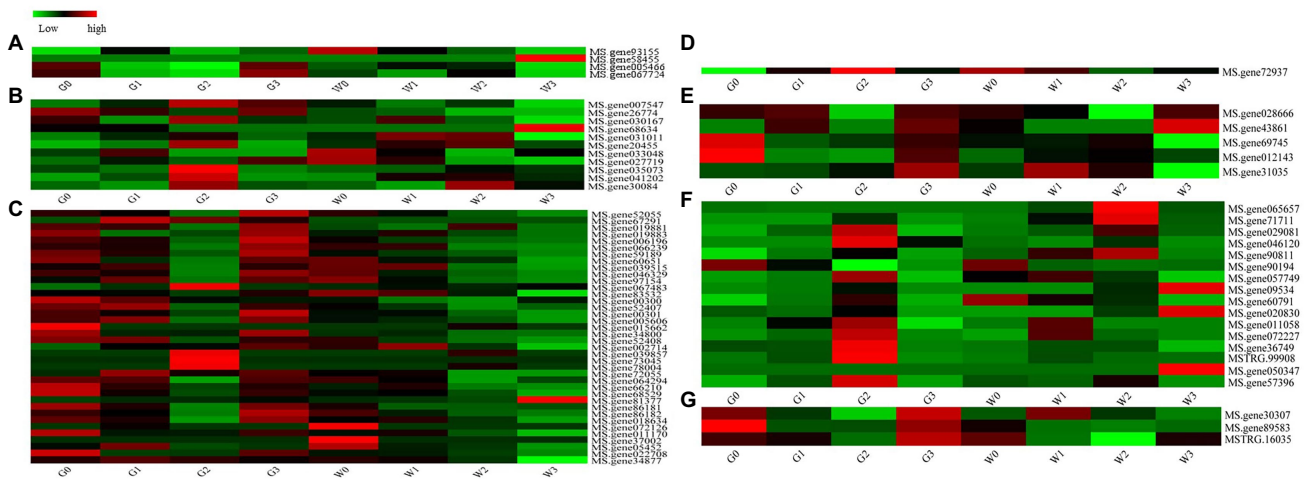


FIGURE 8 | Heat maps of the ATP-binding cassette (ABC) family genes in the alfalfas from RNA-seq. **(A)** ABC Subfamily A, **(B)** ABC Subfamily B, **(C)** ABC Subfamily C, **(D)** ABC Subfamily E, **(E)** ABC Subfamily F, **(F)** ABC Subfamily G, and **(G)** ABC Subfamily I.

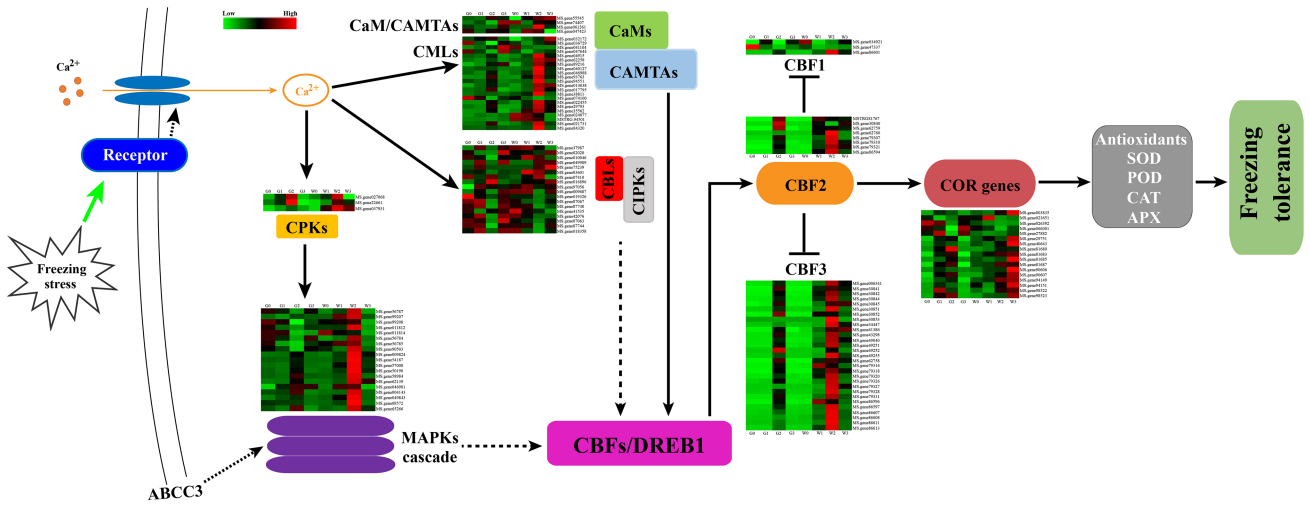


FIGURE 9 | Hypothetical schematic diagram for the mechanism of tolerance to abrupt freezing stress in alfalfa. Freezing stress induced calcium signal transduction through CaMs, CMLs, CBLs, and CPKs. CPKs directly responded to the signal transduction and active MAPKs were involved in the downstream CBFs/DREB1 expression. CaMs and CBLs interacted with CAMTAs and CIPKs, respectively, and regulated the genes of CBFs/DREB1. CBF1, CBF2, and CBF3 trigged an expression of COR genes, which resulted in an increase in the activities of SOD, POD, CAT, and APX. We hypothesize that the hub gene ABCC3 was involved in the MAPK signal pathway and regulated the related genes of CBFs expression. This may have triggered the COR genes and led to an increase in antioxidant enzymes that counteracted abrupt freezing stress in the alfalfa leaves. The dashed arrows indicate that the mechanism is unclear, and CBF2 reduced the expression of CBF1 and CBF3.

dehydrate responsive element binding factors) to confer freezing stress.

Expression of CBFs/DREB1s Related Genes Induced by Abrupt Freezing

The CBF/DREB1 network of genes is important in mitigating freezing stress (Cheng et al., 2007), especially *CBF1/DREB1B*, *CBF2/DREB1C*, and *CBF3/DREB1A* (Ullah et al., 2017). In our study, the expression of *CBF1/DREB1B*, *CBF2/DREB1C*, and *CBF3/DREB1A* genes in alfalfa leaves was higher after freezing stress (**Figure 9**) and was associated with the expression of COR genes. Importantly, the most genes of CBFs/DREB1 of "WL326GZ" were induced earlier than in "Gannong NO.3" leaves.

DISCUSSION

Physiological Responses of Leaves in Two Different Types of Alfalfa to Freezing Stress

Freezing stress induces ROS (e.g., superoxide radicals, hydrogen peroxide, and hydrogen radicals), which damages the cell membrane and increases lipid peroxidation in plants. Plants activate antioxidant defense systems, including SOD, CAT, POD, and APX, to alleviate the damage caused by ROS (Liang et al., 2008). It is demonstrated that the functions of SOD, POD, and APX convert the superoxide radical into hydrogen peroxide, and CAT eliminated hydrogen peroxide (Wang et al., 2009). Recently, several reports have demonstrated that ROS levels are negatively related to stress resistance, and the greater the stress resistance of plants, the stronger the activities of the antioxidant system (Peng et al., 2021). Importantly, our results (**Figure 2**) are in accordance with this conclusion, as the activities of antioxidant in "Gannong NO.3" leaves were higher than those in "WL326GZ" leaves. Similarly, antioxidant enzyme levels of other freezing tolerant plants have consistently been higher than in their freezing-sensitive counterparts for many species, including maize (Hodges et al., 1997), rice (Huang and Guo, 2005), and chickpeas (Kaur et al., 2009). Furthermore, alfalfa cultivar showed similar trends to those of our study (Wang et al., 2009). Therefore, our conclusions suggest that alfalfa acquires freezing tolerance by increasing the activities of antioxidant enzymes in the ROS scavenging system.

Differentially Expressed Genes in Response to Freezing Stress

To explore the changes in abrupt freezing stress-related gene expression in alfalfa cultivars, we transcribed RNA from alfalfa leaves which exposure to −10°C for 0, 0.5, 1, and 2 h. In this study, we obtained 121,366 genes by RNA-seq based on the alfalfa reference genome. Our transcription assembly was more efficient than that of other alfalfa studies (Song et al., 2016; Shu et al., 2017; Xu et al., 2019). Furthermore, 7,245 DEGs were identified from the leaves of "Gannong NO.3" and "WL326GZ" exposure to freezing stress. To compare the differences in response to freezing stress in the two alfalfa cultivars, we analyzed the trends of DEGs in "Gannong NO.3" and "WL326GZ." These results suggest that metabolic pathways and biosynthesis of secondary metabolites may play a critical role in counteracting freezing stress. Moreover, the GO enrichment results showed DEGs have been enriched in response

to stimulus and stress in biological processes and in the binding and catalytic activities during molecular functions, which were been included in the transcriptomics for several other alfalfa cultivars of freezing stress (Shu et al., 2017; Xu et al., 2019).

Role of ABC Transporter Family Genes in Abrupt Freezing Tolerance

The ABC transporter proteins are one of the largest protein families and found in yeast, mammals, and plants (Raichaudhuri, 2016). And several studies have revealed that ABC proteins play a key role in plants involved in diverse processes, including abiotic stress. In plants, eight ABC transporter subfamilies have been classified, subfamilies ABAA-G and ABCI (Wanke and Üner Kolukisaoglu, 2010). In our study, we identified 80 DEGs belonging to the ABC family, which were grouped into seven subfamilies exclude ABCD subfamily (**Figure 8**). The evidence showed that ABCD subfamily transporters are involved in the development of related peroxisomes in *Caenorhabditis elegans*, which have been previously found in insects (Tian et al., 2017); thus, there were no ABCD subfamily genes induced in our study. Interestingly, there were 39 genes (48.75%) belonging to the ABCC subfamily, which suggests that this subfamily may play a critical role in freezing tolerance. Importantly, as Raichaudhuri (2016) reported that the ABCC subfamily is one of the most important subfamilies of ABC transporters based on it can transport many types of glutathione conjugates in vacuolar membranes. And ABCCs have been identified as glutathione-S (GS) conjugate pumps, which can transport endogenous substances and chlorophyll catabolites as well as regulate the ion channel (Wanke and Üner Kolukisaoglu, 2010). Thus, the ABCC subfamily occupies an important position in response to freezing stress.

To analyze the hub genes as DEGs in alfalfa leaves in response to freezing stress, we obtained two hub genes belonging to the ABCC subfamily (**Table 1**), *ABCC8* and *ABCC3*. Multiple *ABCC8* genes have been observed in plants, such as *AtABCC8* (Sánchez-Fernández et al., 2001), *OsABCC8* (Jasinski et al., 2003), *FvABCC8* (Shi et al., 2020), and *TaABCC8* (Bhati et al., 2015). However, only a few studies have reported the function of *ABCC8* genes in plants. Recently, *SmABCC8* was found to be involved in the transportation of substances in leaves (Yan et al., 2021). Moreover, *OsABCC8* has been identified as a homolog of *AtABCC4* in rice, which is associated with the response of guard cell plasma membrane ion channels to abiotic stress (Saha et al., 2015). Therefore, we hypothesize that *ABCC8* has a similar function in alfalfa.

In addition, recent studies have shown that *ZmMRP3* (*ZmABCC3*) participates in the biosynthesis of anthocyanin (Rea, 2007), and overexpression of *AtABCC3* improves tolerance to bleomycin (Li et al., 2021) and SA. Moreover, *ABCC3* is also associated with the MAPK signaling pathway (Guo et al., 2015; Jin et al., 2020), which can be resistant to biotic (Guo et al., 2019), and thus is important for adaptation to environmental changes (Brunetti et al., 2015; Saha et al., 2015; Su et al., 2021). Therefore, the alfalfa *ABCC3* may help regulate downstream genes *via* the MAPK signal pathway to contribute to freezing tolerance (**Figure 9**).

Ca^{2+} Signal Transduction Components Trigger Downstream CBF/DREB1 Gene Expression for Resistance to Freezing Stress

Calcium ions (Ca^{2+}) act as the second messenger involved in a wide range of biological processes in eukaryotic cells and plays a key role in counteracting cold stress. Interestingly, in alfalfa, Orvar et al. (2000) reported that cold stress-induced rigidification at microdomains on the plasma membrane results in a rearrangement of the cytoskeleton; this leads to stretch-sensitive Ca^{2+} channels and elevates the concentration of Ca^{2+}in the cytosol, which stimulates expression of cold-induced genes to resist stress. Similar Ca^{2+} signals were induced in this study. Ca^{2+} signals are relayed or respond to Ca^{2+} sensor proteins (Ogunrinde et al., 2017), including CaMs, CMLs, and CBLs belonging to sensor relays, and CPKs/CDPKs that acted as sensor responders in this study. CaMs/CMLs and CBLs do not have any catalytic activity; their functions need to be combined with biomolecular interactions. For example, CaMs combine with calmodulin-binding transcriptional activators (CAMTAs) and CBL-interacting protein kinases (CIPKs) to relay Ca^{2+} transients for phosphorylation (Viswanathan and Zhu, 2002; Marcec et al., 2019), and CMLs can interact with many types of proteins (Perochon et al., 2011).

We found that the CaMs/CMLs-CAMTAs DEGs increased with prolonged freezing stress and then decreased. Thus, CaMs-CAMTAs may positively regulate freezing tolerance. Our results are consistent with those of other studies (Kim et al., 2013; Yuan et al., 2018; Yang et al., 2020). CAMTAs positively regulate CBFs and CORs (Kim et al., 2013), thereby increasing the activity of the ROS scavenging system, including POD, CAT, and APX, to defend against freezing stress (Peng et al., 2021). However, the gene expression of CBLs-CIPKs is reportedly more complex than that of CaMs-CAMTAs (Zhang et al., 2020). One CBL can function with one CIPK or multiple CIPKs, and the final function of CIPKs depends on interactions with CBLs, which can participate in multiple biological processes (Tang et al., 2020). Thus, they are expressed in different tissues under various stresses, including cold stress (Yuan et al., 2018). Therefore, several genes of CBLs-CIPKs were expressed in this study. The responses of CPKs/CDPKs proteins to the Ca^{2+} sensor play an important role in the Ca^{2+} signals involved in ion channels, metabolic enzymes, and other processes (Yip Delormel and Boudsocq, 2019). In alfalfa, CPKs/CDPKs prevent cold-induced gene expression, preventing plants from adapting to environmental changes (Monroy et al., 1993); this is consistent with our results. The expression of CPKs in cold-sensitive "WL326GZ" increased more rapidly than in cold-resistant "Gannong NO.3," ultimately resulting in a much higher expression in the former. In addition, CDPKs can induce the MAPK pathways in response to cold stress (Sangwan et al., 2002; Lv et al., 2018). MAPK cascades are key in alleviating adverse environmental conditions. For example, many MAPKs are involved in the cascade process (Zhang and Klessig, 2001); MMK4 in alfalfa is induced under cold stress (Jonak et al., 1996); and MPK3 and MPK6 reduce CBF gene expression to negatively regulate cold resistance (Zhao et al., 2017). Similar

to these reports, we found that several MAPKs were induced by abrupt freezing stress. And Sangwan et al. (2002) demonstrated that cold shock-induced MAPKs in alfalfa to resist freezing tolerance, which is consistent with our results. Overall, the components of Ca^{2+} signal transduction triggered downstream CBFs/DREB1 gene expression in alfalfa to ensure resistance to freezing stress (**Figure 9**).

CBF/DREB1 belongs to the APETALA2/ethylene-responsive factor (AP2/ERF) transcription factor family (Yang et al., 2009), which is also critical in mitigating cold stress based on initially induced cold response transcript genes in a transcriptional cascade (Kidokoro et al., 2017); moreover, it plays a key role in facilitating COR and CBF gene expression (Yang et al., 2009). In Arabidopsis, *CBF1* (*DREB1B*), *CBF2* (*DREB1C*), and *CBF3* (*DREB1A*) were rapidly induced under cold stress, and overexpression of *CBF1, CBF2,* and *CBF3* increased freezing tolerance after exposure to freezing shock stress (Park et al., 2020). Moreover, alfalfa CBFs have been identified as key regulators of freezing tolerance (Anower et al., 2016). In this study, the expression levels of CBF2 and CBF3 genes initially increased at 1 h of freezing stress and then decreased (**Figure 9**). This dynamic has been identified in Arabidopsis (Daniel et al., 2003). The expression levels of CBF2 and CBF3 in "WL326GZ" were higher and occurred earlier than those in "Gannong NO.3," possibly because the former is sensitive to abrupt freezing stress. This is accordance with our observed trends for the antioxidant enzyme activities. However, our results are contrary to the cold acclimation mechanism of alfalfa (Chen et al., 2015). This indicates that the abrupt freezing stress response mechanism is different from that of cold acclimation.

In summary, we compared the leaves of the freeze-tolerant genotype "Gannong NO.3" and the freeze-sensitive genotype "WL326GZ" after exposure to abrupt freezing stress and evaluated their physiology and transcriptome. Our results suggest that the ABCC subfamily plays an important role in tolerance to abrupt freezing stress in alfalfa. Moreover, we demonstrated that *ABCC8* and *ABCC3* are the key genes in counteracting freezing stress, and further studies should explore the exact functions of these genes. A simple working module is illustrated in **Figure 9**. The expression levels of the Ca^{2+} signal transduction and CBF/DREB1 related genes demonstrate their critical role in counteracting freezing stress. Additionally, they may be associated with the higher antioxidant enzyme activities (SOD, POD, CAT, and APX activities) in "Gannong NO.3" than in "WL326GZ" (**Figure 9**). This could explain why "WL326GZ" is more sensitive to abrupt changes in temperature than "Gannong NO.3." Moreover, high antioxidant activities early in the stress exposure were associated with high expression levels of genes related to freezing tolerance, even though the duration of the high antioxidant activities was short. Our results provide new insights into the mechanisms of abrupt freezing tolerance in alfalfa.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

JM and SS conceived and designed the research. XW, WK, and FW conducted the experiment. XW and JM analyzed the data and wrote the manuscript. WK, FW, and SS finalized the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Fostering Foundation for the Excellent Ph.D. Dissertation of Gansu Agricultural University (YB2019001), Major Special Science and Technology Projects of Gansu Province: Germplasm Innovation and Breeding of Alfalfa and Oat (19ZD2NA002-3), and Discipline Construction Fund Project of Gansu Agricultural University (GAU-XKJS-2018-002).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.798118/full#supplementary-material

Supplementary Figure S1 | KEGG pathway of the top 3 profiles in the "Gannong NO.3" leaf samples.

Supplementary Figure S2 | KEGG pathway of the top 3 profiles in the "WL326GZ" leaf samples.

Supplementary Figure S3 | Bar plot of the top 20 KEGG pathways in the two key modules.

REFERENCES

Ahmed, N. U., Jung, H. J., Park, J. I., Cho, Y. G., Hur, Y., and Nou, I. S. (2015). Identification and expression analysis of cold and freezing stress responsive genes of Brassica oleracea. *Gene* 554, 215–223. doi: 10.1016/j.gene.2014.10.050

Anower, M. R., Fennell, A., Boe, A., Mott, I. W., Peel, M. D., and Wu, Y. (2016). Physiological and molecular characterisation of lucerne (*Medicago sativa* L.) germplasm with improved seedling freezing tolerance. *Crop Pasture Sci.* 67:655. doi: 10.1071/cp15204

Bhati, K. K., Sharma, S., Aggarwal, S., Kaur, M., Shukla, V., Kaur, J., et al. (2015). Genome-wide identification and expression characterization of ABCC-MRP transporters in hexaploid wheat. *Front. Plant Sci.* 6:488. doi: 10.3389/fpls.2015.00488

Brunetti, P., Zanella, L., De Paolis, A., Di Litta, D., Cecchetti, V., Falasca, G., et al. (2015). Cadmium-inducible expression of the ABC-type transporter

AtABCC3 increases phytochelatin-mediated cadmium tolerance in *Arabidopsis*. *J. Exp. Bot.* 66, 3815–3829. doi: 10.1093/jxb/erv185

Castonguay, Y., Rocher, S., Bertrand, A., and Michaud, J. B. (2020). Identification of transcripts associated with the acquisition of superior freezing tolerance in recurrently-selected populations of alfalfa. *Euphytica* 216:27. doi: 10.1007/s10681-020-2559-2

Chance, B., and Maehly, A. C. (1955). Assay of catalases and peroxidases. *Methods Enzymol.* 2, 764–775. doi: 10.1016/S0076-6879(55)02300-8

Chao, Y., Yuan, J., Guo, T., Xu, L., Mu, Z., and Han, L. (2019). Analysis of transcripts and splice isoforms in *Medicago sativa* L. by single-molecule long-read sequencing. *Plant Mol. Biol.* 99, 219–235. doi: 10.1007/s11103-018-0813-y

Chen, J., Han, G. Q., Shang, C., Li, J. K., Zhang, H. L., Liu, F. Q., et al. (2015). Proteomic analyses reveal differences in cold acclimation mechanisms in freezing-tolerant and freezing-sensitive cultivars of alfalfa. *Front. Plant Sci.* 6:105. doi: 10.3389/fpls.2015.00105

Chen, J., Xue, B., Xia, X., and Yin, W. (2013). A novel calcium-dependent protein kinase gene from Populus euphratica, confers both drought and cold stress tolerance. *Biochem. Biophys. Res. Commun.* 441, 630–636. doi: 10.1016/j.bbrc.2013.10.103

Chen, H., Zeng, Y., Yang, Y., Huang, L., Tang, B., Zhang, H., et al. (2020). Allele-aware chromosome-level genome assembly and efficient transgene-free genome editing for the autotetraploid cultivated alfalfa. Nature. *Communications* 11:2494. doi: 10.1038/s41467-020-16338-x

Cheng, C., Yun, K. Y., Ressom, H. W., Mohanty, B., Bajic, V. B., Jia, Y., et al. (2007). An early response regulatory cluster induced by low temperature and hydrogen peroxide in seedlings of chilling-tolerant japonica rice. *BMC Genomics* 8:175. doi: 10.1186/1471-2164-8-175

Cheong, Y. H., Kim, K. N., Pandey, G. K., Gupta, R., Grant, J. J., and Luan, S. (2003). CBL1, a calcium sensor that differentially regulates salt, drought, and cold responses in *Arabidopsis*. *Plant Cell* 15, 1833–1845. doi: 10.1105/tpc.012393

Cho, K. M., Nguyen, H. T. K., Kim, S. Y., Shin, J. S., Cho, D. H., Hong, S. B., et al. (2016). CML10, a variant of calmodulin, modulates ascorbic acid synthesis. *New Phytol.* 209, 664–678. doi: 10.1111/nph.13612

Daniel, G. Z., Jonathan, T., Vogel, C. D., and Thomashow, M. F. (2003). Cold induction of *Arabidopsis* CBF genes involves multiple ICE (inducer of CBF expression) promoter elements and a cold-regulatory circuit that is desensitized by low temperature. *Plant Physiol.* 133, 910–918. doi: 10.1104/pp.103.027169

Delk, N. A., Johnson, K. A., Chowdhury, N. I., and Braam, J. (2005). *CML24*, regulated in expression by diverse stimuli, encodes a potential Ca^{2+} sensor that functions in responses to abscisic acid, daylength, and ion stress. *Plant Physiol.* 139, 240–253. doi: 10.1104/pp.105.062612

Dubrovina, A. S., Kiselev, K. V., Khristenko, V. S., and Aleynova, O. A. (2015). *VaCPK20,* a calcium-dependent protein kinase gene of wild grapevine Vitis amurensis Rupr., mediates cold and drought stress tolerance. *J. Plant Physiol.* 185, 1–12. doi: 10.1016/j.jplph.2015.05.020

Eremina, M., Rozhon, W., and Poppenberger, B. (2016). Hormonal control of cold stress responses in plants. *Cell. Mol. Life Sci.* 73, 797–810. doi: 10.1007/s00018-015-2089-6

Giannopolitis, C. N., and Ries, S. K. (1977). Superoxide dismutases I. Occurrence in higher plants. *Plant Physiol.* 59, 309–314. doi: 10.1104/pp.59.2.309

Gong, Z. Z., Xiong, L. M., Shi, H. Z., Yang, S. H., Herrera-Estrella, L. R., Xu, G., et al. (2020). Plant abiotic stress response and nutrient use efficiency. *Sci. China Life Sci.* 63, 635–674. doi: 10.1007/s11427-020-1683-x

Guo, Z., Kang, S., Chen, D., Wu, Q., Wang, S., Xie, W., et al. (2015). MAPK signaling pathway alters expression of Midgut ALP and ABCC genes and causes resistance to bacillus thuringiensis Cry1Ac toxin in diamondback moth. *PLoS Genet.* 11:e1005124. doi: 10.1371/journal.pgen.1005124

Guo, Z., Sun, D., Kang, S., Zhou, J., Gong, L., Qin, J., et al. (2019). CRISPR/Cas9-mediated knockout of both the PxABCC2 and PxABCC3 genes confers high-level resistance to *Bacillus thuringiensis* Cry1Ac toxin in the diamondback moth, *Plutella xylostella* (L.). *Insect Biochem. Mol. Biol.* 107, 31–38. doi: 10.1016/j.ibmb.2019.01.009

Hodges, D. M., Andrews, C. J., Johnson, D. A., and Hamilton, R. I. (1997). Antioxidant enzyme responses to chilling stress in differentially sensitive inbred maize lines. *J. Exp. Bot.* 48, 1105–1113. doi: 10.1093/jxb/48.5.1105

Huang, M., and Guo, Z. (2005). Responses of antioxidative system to chilling stress in two rice cultivars differing in sensitivity. *Biol. Plant.* 49, 81–84. doi: 10.1007/s00000-005-1084-3

Jasinski, M., Ducos, E., Martinoia, E., and Boutry, M. (2003). The ATP-binding cassette transporters: structure, function, and gene family comparison between rice and Arabidopsis. *Plant Physiol.* 131, 1169–1177. doi: 10.1104/pp.102.014720

Jin, M., Yang, Y., Shan, Y., Chakrabarty, S., Cheng, Y., Soberón, M., et al. (2020). Two ABC transporters are differentially involved in the toxicity of two bacillus thuringiensis Cry1 toxins to the invasive crop-pest *Spodoptera frugiperda* (J. E. Smith). *Pest Manag. Sci.* 77, 1492–1501. doi: 10.1002/ps.6170

Jonak, C., Kiegerl, S., Ligterink, W., Barker, P. J., Huskisson, N. S., and Hirt, H. (1996). Stress signaling in plants: a mitogen-activated protein kinase pathway is activated by cold and drought. *Plant Biol.* 93, 111274–111279.

Kanchupati, P., Wang, Y., Anower, M. R., Boe, A., and Wu, Y. (2017). The CBF-like gene family in alfalfa: expression analyses and identification of potential functional homologs of *Arabidopsis* CBF3. *Crop Sci.* 57, 2051–2063. doi: 10.2135/cropsci2016.09.0777

Karimzadeh, S. H., Nezami, A., Nabati, J., Oskoueian, E., and Ahmadi-Lahijani, M. J. (2021). The physiological, biochemical, and molecular modifications of chickpea (*Cicer arietinum* L.) seedlings under freezing stress. *J. Plant Growth Regul.* 1–16. doi: 10.1007/s00344-021-10369-4

Kaur, S., Gupta, A. K., Kaur, N., Sandhu, J. S., and Gupta, S. K. (2009). Antioxidative enzymes and sucrose synthase contribute to cold stress tolerance in chickpea. *J. Agron. Crop Sci.* 195, 393–397. doi: 10.1111/j.1439-037X.2009.00383.x

Kidokoro, S., Yoneda, K., Takasaki, H., Takahashi, F., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2017). Different cold-signaling pathways function in the responses to rapid and gradual decreases in temperature. *Plant Cell* 29, 760–774. doi: 10.1105/tpc.16.00669

Kim, Y., Park, S., Gilmour, S. J., and Thomashow, M. F. (2013). Roles of CAMTA transcription factors and salicylic acid in configuring the low-temperature transcriptome and freezing tolerance of *Arabidopsis*. *Plant J.* 75, 364–376. doi: 10.1111/tpj.12205

Kolukisaoglu, Ü., Weinl, S., Blazevic, D., Batistic, O., and Kudla, J. (2004). Calcium sensors and their interacting protein kinases: genomics of the arabidopsis and rice CBL-CIPK signaling networks. *Plant Physiol.* 134, 43–58. doi: 10.1104/pp.103.033068

Komatsu, S., Yang, G., Khan, M., Onodera, H., Toki, S., and Yamaguchi, M. (2007). Over expression of calcium-dependent protein kinase 13 and calreticulin interacting protein 1 confers cold tolerance on rice plants. *Mol. Gen. Genomics.* 277, 713–723. doi: 10.1007/s00438-007-0220-6

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559. doi: 10.1186/1471-2105-9-559

Lee, J., Lim, Y. P., Han, C. T., Nou, I. S., and Hur, Y. (2013). Genome-wide expression profiles of contrasting inbred lines of Chinese cabbage, Chiifu and Kenshin, under temperature stress. *Genes Genomics* 35, 273–288. doi: 10.1007/s13258-013-0088-2

Li, D., Song, S., Xia, X., and Yin, W. (2012). Two CBL genes from Populus euphratica confer multiple stress tolerance in transgenic triploid white poplar. *Plant Cell Tissue Organ Cult.* 109, 477–489. doi: 10.1007/s11240-011-0112-7

Li, T., Xu, S., Wu, C., Yan, S., and Wang, L. (2021). Loss of an ABC transporter in Arabidopsis thaliana confers hypersensitivity to the anti-cancer drug bleomycin. *DNA Repair* 106:103174. doi: 10.1016/j.dnarep.2021.103174

Li, H., Ye, K., Shi, Y., Cheng, J., Zhang, X., and Yang, S. (2017). BZR1 positively regulates freezing tolerance via CBF-dependent and CBF-independent pathways in *Arabidopsis*. *Mol. Plant* 10, 545–559. doi: 10.1016/j.molp.2017.01.004

Liang, Y., Zhu, J., Li, Z., Chu, G., Ding, Y., Zhang, J., et al. (2008). Role of silicon in enhancing resistance to freezing stress in two contrasting winter wheat cultivars. *Environ. Exp. Bot.* 64, 286–294. doi: 10.1016/j.envexpbot.2008.06.005

Liu, Y., Wu, C., Hu, X., Gao, H., Wang, Y., Luo, H., et al. (2020). Transcriptome profiling reveals the crucial biological pathways involved in cold response in Moso bamboo (*Phyllostachys edulis*). *Tree Physiol.* 40, 538–556. doi: 10.1093/treephys/tpz133

Lv, X., Li, H., Chen, X., Xiang, X., Guo, Z., Yu, J., et al. (2018). The role of calcium-dependent protein kinase in hydrogen peroxide, nitric oxide and ABA-dependent cold acclimation. *J. Exp. Bot.* 69, 4127–4139. doi: 10.1093/jxb/ery212

Ma, Q., Zhou, Q., Chen, C., Cui, Q., Zhao, Y., Wang, K., et al. (2019). Isolation and expression analysis of *CsCML* genes in response to abiotic stresses in the tea plant (*Camellia sinensis*). *Sci. Rep.* 9:8211. doi: 10.1038/s41598-019-44681-7

Mall, T. K., Dweikat, I., Sato, S. J., Neresian, N., Xu, K., Ge, Z., et al. (2011). Expression of the rice CDPK-7 in sorghum: molecular and phenotypic analyses. *Plant Mol. Biol.* 75, 467–479. doi: 10.1007/s11103-011-9741-9

Marcec, M. J., Gilroy, S., Poovaiah, B. W., and Tanaka, K. (2019). Mutual interplay of Ca^{2+} and ROS signaling in plant immune response. *Plant Sci.* 283, 343–354. doi: 10.1016/j.plantsci.2019.03.004

Monroy, A. F., and Dhindsa, R. S. (1995). Low-temperature signal transduction: induction of cold acclimation-specific genes of alfalfa by calcium at 25 degrees C. *Plant Cell* 7, 321–331. doi: 10.2307/386985

Monroy, A. F., Sarhan, F., and Dhindsa, R. S. (1993). Cold induced changes in freezing tolerance, protein phosphorylation, and gene expression: evidence for a role of calcium. *Plant Physiol.* 102, 1227–1235. doi: 10.1104/pp.102.4.1227

Nah, G., Lee, M., Kim, D. S., Rayburn, A. L., Voigt, T., Lee, D. K., et al. (2016). Transcriptome analysis of Spartina pectinata in response to freezing stress. *PLoS One* 11:e0152294. doi: 10.1371/journal.pone.0152294

Nguyen, V. N. T., Moon, S., and Jung, K. H. (2014). Genome-wide expression analysis of rice ABC transporter family across spatio-temporal samples and in response to abiotic stresses. *J. Plant Physiol.* 171, 1276–1288. doi: 10.1016/j.jplph.2014.05.006

Ogunrinde, A., Munro, K., Davidson, A., Ubaid, M., and Snedden, W. A. (2017). *Arabidopsis* calmodulin-like proteins, CML15 and CML16 possess biochemical properties distinct from calmodulin and show non-overlapping tissue expression patterns. *Front. Plant Sci.* 8:2175. doi: 10.3389/fpls.2017.02175

Orvar, B. L., Sangwan, V., Omann, F., and Dhindsa, R. S. (2000). Early steps in cold sensing by plant cells: the role of actin cytoskeleton and membrane fluidity. *Plant J.* 23, 785–794. doi: 10.1046/j.1365-313x.2000.00845.x

Park, S., Shi, A., and Mou, B. (2020). Genome-wide identification and expression analysis of the CBF/DREB1 gene family in lettuce. *Sci. Rep.* 10:5733. doi: 10.1038/s41598-020-62458-1

Peng, T., You, X. S., Guo, L., Zhong, B. L., Mi, L. F., Chen, J. M., et al. (2021). Transcriptome analysis of Chongyi wild mandarin, a wild species more cold-tolerant than *Poncirus trifoliata*, reveals key pathways in response to cold. *Environ. Exp. Bot.* 184:104371. doi: 10.1016/j.envexpbot.2020.104371

Perochon, A., Aldon, D., Galaud, J. P., and Ranty, B. (2011). Calmodulin and calmodulin-like proteins in plant calcium signaling. *Biochimie* 93, 2048–2053. doi: 10.1016/j.biochi.2011.07.012

Raichaudhuri, A. (2016). Arabidopsis thaliana MRP1 (*AtABCC1*) nucleotide binding domain contributes to arsenic stress tolerance with serine triad phosphorylation. *Plant Physiol. Biochem.* 108, 109–120. doi: 10.1016/j.plaphy.2016.07.005

Ray, S., Agarwal, P., Arora, R., Kapoor, S., and Taygi, A. K. (2007). Expression analysis of calcium-dependent protein kinase gene family during reproductive development and abiotic stress conditions in rice (*Oryza sativa* L. spp. indica). *Mol. Gen. Genet.* 278, 493–505. doi: 10.1007/s00438-007-0267-4

Rea, P. A. (2007). Plant ATP-binding cassette transporters. *Annu. Rev. Plant Biol.* 58, 347–375. doi: 10.1146/annurev.arplant.57.032905.105406

Saha, J., Sengupta, A., Gupta, K., and Gupta, B. (2015). Molecular phylogenetic study and expression analysis of ATP-binding cassette transporter gene family in *Oryza sativa* in response to salt stress. *Comput. Biol. Chem.* 54, 18–32. doi: 10.1016/j.compbiolchem.2014.11.0

Sánchez-Fernández, R., Davies, T. G. E., Coleman, J. O. D., and Rea, P. A. (2001). The *Arabidopsis thaliana* ABC protein superfamily, a complete inventory. *J. Biol. Chem.* 276, 30231–30244. doi: 10.1074/jbc.m103104200

Sangwan, V., Orvar, B. L., Beyerly, J., Hirt, H., and Dhindsa, R. S. (2002). Opposite changes in membrane fluidity mimic cold and heat stress activation of distinct plant MAP kinase pathways. *Plant J.* 31, 629–638. doi: 10.1046/j.1365-313x.2002.01384.x

Shi, M., Wang, S., Zhang, Y., Wang, S., Zhao, J., Feng, H., et al. (2020). Genome-wide characterization and expression analysis of ATP-binding cassette (ABC) transporters in strawberry reveal the role of *FvABCC11* in cadmium tolerance. *Sci. Hortic.* 271:109464. doi: 10.1016/j.scienta.2020.109464

Shu, Y., Li, W., Zhao, J., Zhang, S., Xu, H., Liu, Y., et al. (2017). Transcriptome sequencing analysis of alfalfa reveals CBF genes potentially playing important roles in response to freezing stress. *Genet. Mol. Biol.* 40, 824–833. doi: 10.1590/1678-4685-gmb-2017-0053

Song, L., Jiang, L., Chen, Y., Shu, Y., Bai, Y., and Guo, C. (2016). Deep-sequencing transcriptome analysis of field-grown *Medicago sativa* L. crown buds acclimated to freezing stress. *Funct. Integr. Genomics* 16, 495–511. doi: 10.1007/s10142-016-0500-5

Su, L., Xie, Y., He, Z., Zhang, J., Tang, Y., and Zhou, X. (2021). Network response of two cherry tomato (*Lycopersicon esculentum*) cultivars to cadmium stress as revealed by transcriptome analysis. *Ecotoxicol. Environ. Saf.* 222:112473. doi: 10.1016/j.ecoenv.2021.112473

Sun, Q., Huang, R., Zhu, H., Sun, Y., and Guo, Z. (2021). A novel *Medicago truncatula* calmodulin-like protein (MtCML42) regulates cold tolerance and flowering time. *Plant J.* 108, 1069–1082. doi: 10.1111/tpj.15494

Sun, T., Wang, Y., Wang, M., Li, T., Zhou, Y., Wang, X., et al. (2015). Identification and comprehensive analyses of the CBL, and CIPK, gene families in wheat (*Triticum aestivum* L). *Plant Biol.* 15:269. doi: 10.1186/s12870-015-0657-4

Tang, R.-J., Wang, C., Li, K., and Luan, S. (2020). The CBL-CIPK calcium signaling network: unified paradigm from 20 years of discoveries. *Trends Plant Sci.* 25, 604–617. doi: 10.1016/j.tplants.2020.01.009

Tayeh, N., Bahrman, N., Sellier, H., Bluteau, A., Blassiau, C., Fourment, J., et al. (2013). A tandem array of CBF/DREB1 genes is located in a major freezing tolerance QTL region on *Medicago truncatula* chromosome 6. *BMC Genomics* 14:814. doi: 10.1186/1471-2164-14-814

Tian, L., Song, T., He, R., Zeng, Y., Xie, W., Wu, Q., et al. (2017). Genome-wide analysis of ATP-binding cassette (ABC) transporters in the sweetpotato whitefly, *Bemisia tabaci*. *BMC Genomics* 18:330. doi: 10.1186/s12864-017-3706-6

Townley, H. E., and Knight, M. R. (2002). Calmodulin as a potential negative regulator of *Arabidopsis* COR gene expression. *Plant Physiol.* 128, 1169–1172. doi: 10.1104/pp.010814

Tsutsui, T., Kato, W., Asada, Y., Sako, K., Sato, T., Sonoda, Y., et al. (2009). DEAR1, a transcriptional repressor of DREB protein that mediates plant defense and freezing stress responses in *Arabidopsis*. *J. Plant Res.* 122, 633–643. doi: 10.1007/s10265-009-0252-6

Ullah, J. A., Fazal, H., Midrarullah, A. A., and Khaista, R. (2017). Role of CBF/DREB gene expression in abiotic stress tolerance. A review symbiosis list of abbreviations. *Int. J. Hort. Agric.* 2, 1–12.

Viswanathan, C., and Zhu, J. K. (2002). Molecular genetic analysis of cold-regulated gene transcription. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 357, 877–886. doi: 10.1098/rstb.2002.1076

Wang, L., Feng, X., Yao, L., Ding, C., Lei, L., Hao, X., et al. (2020). Characterization of CBL–CIPK signaling complexes and their involvement in cold response in tea plant. *Plant Physiol. Biochem.* 154, 195–203. doi: 10.1016/j.plaphy.2020.06.005

Wang, W. B., Kim, Y. H., Lee, H. S., Deng, X. P., and Kwak, S. S. (2009). Differential antioxidation activities in two alfalfa cultivars under chilling stress. *Plant Biotechnol. Rep.* 3, 301–307. doi: 10.1007/s11816-009-0102-y

Wanke, D., and Üner Kolukisaoglu, H. (2010). An update on the ABCC transporter family in plants: many genes, many proteins, but how many functions? *Plant Biol.* 12, 15–25. doi: 10.1111/j.1438-8677.2010.00380.x

Weckwerth, P., Ehlert, B., and Romeis, T. (2015). ZmCPK1, a calcium-independent kinase member of the Zea mays CDPK gene family, functions as a negative regulator in cold stress signalling. *Plant Cell Environ.* 38, 544–558. doi: 10.1111/pce.12414

Xu, L., Tang, X., Wang, B., Xin, X., and Guo, M. (2019). Comparative transcriptome analysis of five medicago varieties reveals the genetic signals underlying freezing tolerance. *Crop Pasture Sci.* 70:273. doi: 10.1071/CP18165

Yan, L., Zhang, J., Chen, H., and Luo, H. (2021). Genome-wide analysis of ATP-binding cassette transporter provides insight to genes related to bioactive metabolite transportation in Salvia miltiorrhiza. *BMC Genomics* 22:315. doi: 10.1186/s12864-021-07623-0

Yang, T., Chaudhuri, S., Yang, L., Du, L., and Poovaiah, B. W. (2009). A calcium/calmodulin-regulated member of the receptor-like kinase family confers cold tolerance in plants. *J. Biol. Chem.* 285, 7119–7126. doi: 10.1074/jbc.m109.035659

Yang, F., Dong, F., Hu, F., Liu, Y., Chai, J., Zhao, H., et al. (2020). Genome-wide identification and expression analysis of the calmodulin-binding transcription activator (CAMTA) gene family in wheat (*Triticum aestivum* L.). *BMC Genet.* 21:105. doi: 10.1186/s12863-020-00916-5

Yang, T. B., and Poovaiah, B. W. (2003). Calcium/calmodulin-mediated signal network in plants. *Trends Plant Sci.* 8, 505–512. doi: 10.1016/j.tplants.2003.09.004

Yang, S. S., Xu, W., Tesfaye, M., Lamb, J. F., Jung, H. J. G., VandenBosch, K. A., et al. (2010). Transcript profiling of two alfalfa genotypes with contrasting cell wall composition in stems using a cross-species platform: optimizing

analysis by masking biased probes. *BMC Genomics* 11:323. doi: 10.1186/1471-2164-11-323

Yip Delormel, T., and Boudsocq, M. (2019). Properties and functions of calcium-dependent protein kinases and their relatives in *Arabidopsis thaliana*. *New Phytol.* 224, 585–604. doi: 10.1111/nph.16088

Yuan, P., Yang, T., and Poovaiah, B. W. (2018). Calcium signaling-mediated plant response to cold stress. *Int. J. Mol. Sci.* 19:3896. doi: 10.3390/ijms19123896

Zeng, H., Xu, L., Singh, A., Wang, H., Du, L., and Poovaiah, B. W. (2015). Involvement of calmodulin and calmodulin-like proteins in plant responses to abiotic stresses. *Front. Plant Sci.* 6:600. doi: 10.3389/fpls.2015.00600

Zhang, S., and Klessig, D. F. (2001). MAPK cascades in plant defense signaling. *Trends Plant Sci.* 6, 520–527. doi: 10.1016/s1360-1385(01)02103-3

Zhang, F., Li, L., Jiao, Z., Chen, Y., Liu, H., Chen, X., et al. (2016). Characterization of the calcineurin B-like (CBL) gene family in maize and functional analysis of ZmCBL9 under abscisic acid and abiotic stress treatments. *Plant Sci.* 253, 118–129. doi: 10.1016/j.plantsci.2016.09.011

Zhang, X., Li, X., Zhao, R., Zhou, Y., and Jiao, Y. (2020). Evolutionary strategies drive a balance of the interacting gene products for the CBL and CIPK gene families. *New Phytol.* 226, 1506–1516. doi: 10.1111/nph.16445

Zhao, C., Wang, P., Si, T., Hsu, C. C., Wang, L., Zayed, O., et al. (2017). MAP kinase cascades regulate the cold response by modulating ICE1 protein stability. *Dev. Cell* 43, 618.e5–629.e5. doi: 10.1016/j.devcel.2017.09.024

Zhou, Y., Cheng, Y., Yang, Y., Li, X., Supriyo, B., Sun, X., et al. (2016). Overexpression of SpCBL6, a calcineurin B-like protein of *Stipa purpurea*, enhanced cold tolerance and reduced drought tolerance in transgenic *Arabidopsis*. *Mol. Biol. Rep.* 43, 957–966. doi: 10.1007/s11033-016-4036-5

Zhou, Q., Luo, D., Chai, X., Wu, Y., Wang, Y., Nan, Z., et al. (2018). Multiple regulatory networks are activated during cold stress in *Medicago sativa* L. *Int. J. Mol. Sci.* 19:3169. doi: 10.3390/ijms19103169

Zhu, X., Dunand, C., Snedden, W., and Galaud, J. P. (2015). CaM and CML emergence in the green lineage. *Trends Plant Sci.* 20, 483–489. doi: 10.1016/j.tplants.2015.05.010

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 01 April 2022
doi: 10.3389/fpls.2022.791200

Genome-wide Identification of Metal Tolerance Protein Genes in Peanut: Differential Expression in the Root of Two Contrasting Cultivars Under Metal Stresses

*Xueqin Wang, Chaohui Wang, Zheng Zhang and Gangrong Shi**

College of Life Sciences, Huaibei Normal University, Huaibei, China

Metal tolerance proteins (MTP) are $Me^{2+}/H^{+}(K^{+})$ antiporters that play important roles in the transport of divalent cations in plants. However, their functions in peanut are unknown. In the present study, a total of 24 *AhMTP* genes were identified in peanut, which were divided into seven groups belonging to three substrate-specific clusters (Zn-CDFs, Zn/Fe-CDFs, and Mn-CDFs). All *AhMTP* genes underwent whole genome or segmental gene duplication events except *AhMTP12*. Most *AhMTP* members within the same subfamily or group generally have similar gene and protein structural characteristics. However, some genes, such as *AhMTP1.3*, *AhMTP2.4*, and *AhMTP12*, showed wide divergences. Most of *AhMTP* genes preferentially expressed in reproductive tissues, suggesting that these genes might play roles in metal transport during the pod and seed development stages. Excess metal exposure induced expressions for most of *AhMTP* genes in peanut roots depending on cultivars. By contrast, *AhMTP* genes in the root of Fenghua 1 were more sensitive to excess Fe, Cd, and Zn exposure than that of Silihong. Stepwise linear regression analysis showed that the percentage of Fe in shoots significantly and positively correlated with the expression of *AhMTP4.1*, *AhMTP9.1*, and *AhMTPC4.1*, but negatively correlated with that of *AhMTPC2.1* and *AhMTP12*. The expression of *AhMTP1.1* showed a significant and negative correlation with the percentage of Mn in shoots. The percentage of Zn in shoots was significantly and positively correlated with the expression of *AhMTP2.1* but was negatively correlated with that of *AhMTPC2.1*. The differential responses of *AhMTP* genes to metal exposure might be, at least partially, responsible for the different metal translocation from roots to shoots between Fenghua 1 and Silihong.

Keywords: peanut, metal tolerance protein, metal translocation, genome-wide identification, gene expression, cultivar difference

OPEN ACCESS

Edited by:
Rohini Garg,
Shiv Nadar University, India

Reviewed by:
Matthew John Milner,
National Institute of Agricultural Botany (NIAB), United Kingdom
Sanjeev Kumar,
Indian Institute of Technology Guwahati, India

***Correspondence:**
Gangrong Shi
swsgr@126.com

Specialty section:
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

Received: *08 October 2021*
Accepted: *15 March 2022*
Published: *01 April 2022*

Citation:
Wang X, Wang C, Zhang Z and Shi G (2022) Genome-wide Identification of Metal Tolerance Protein Genes in Peanut: Differential Expression in the Root of Two Contrasting Cultivars Under Metal Stresses. Front. Plant Sci. 13:791200. doi: 10.3389/fpls.2022.791200

INTRODUCTION

Some divalent metal ions, such as Zn^{2+}, Fe^{2+}, Mn^{2+}, Cu^{2+}, Co^{2+}, and Ni^{2+}, are essential elements in plants, playing crucial roles in numerous processes, including DNA replication, protein processing, photosynthesis, electron transport in the chloroplasts, and mitochondria. Although deficiency of these metals negatively affects plant growth and development, they can also result

in toxicity at excessive levels (Bhardwaj et al., 2020). Correspondingly, plants have deployed a variety of adaptive mechanisms to precisely maintain metal homeostasis, including controlling the uptake, efflux, mobilization, translocation, trafficking, and storage (Liu et al., 2019). All these processes are mediated by numerous transporters belonging to different protein families including cation diffusion facilitators (CDFs).

The CDFs have been identified to be $Me^{2+}/H^{+}(K^{+})$ antiporters that are involved in the transport of divalent cations such Fe^{2+}, Zn^{2+}, Cd^{2+}, Mn^{2+}, Ni^{2+}, or Co^{2+} in both prokaryotes and eukaryotes (Singh et al., 2015). The CDF family contains three sub-families: (1) Zn-CDF, (2) Fe/Zn-CDFs, and (3) Mn-CDFs (Montanini et al., 2007). Most CDF proteins contain approximately six transmembrane domains (TMDs), a modified signature, and a C-terminal cation efflux domain (Montanini et al., 2007). The six TMDs are normally interconnected by extra-and intra-cellular loops. Among them, one cytosolic loop usually contains a histidine-rich domain, which is predicted to be potential metal binding domains (Montanini et al., 2007). Plant CDFs are usually named metal tolerance proteins (MTPs), while vertebrate CDFs are called solute carrier family 30 (SLC30) or zinc transporter (ZnT).

The MTP proteins in plants can be classified into seven phylogenetic groups: the groups 1 (MTP1–MTP4), 5 (MTP5), and 12 (MTP12) function as Zn-CDFs, the groups 6 (MTP6) and 7 (MTP7) as Fe/Zn-CDFs, and the groups 8 (MTP8) and 9 (MTP9–MTP11) as Mn-CDFs (Gustin et al., 2011). Since the first MTP protein was identified in *Arabidopsis* (ZAT, also named AtMTP1) (van der Zaal et al., 1999), there are large amount of MTP genes were identified in several plant species such as *Arabidopsis* (12 genes), rice (*Oryza sativa*; 10 genes), and wheat (*Triticum aestivum*; 20 genes) (Vatansever et al., 2017). However, most of them have not been functionally characterized in detail. In *Arabidopsis*, both AtMTP1 and AtMTP3 were confirmed to localize in the vacuole, conferring the Zn and/or Co tolerance by the vacuole sequestration of excess Zn^{2+} and/or Co^{2+} (Kobae et al., 2004; Arrivault et al., 2006). AtMTP12 interacts with AtMTP5 to form a functional complex, transporting Zn^{2+} from cytosol to the Golgi apparatus (Fujiwara et al., 2015). As Mn transporters, AtMTP8 and AtMTP11 protect plant cells from Mn toxicity by producing excessive endoplasmic vesicles that phagocytize and excrete Mn^{2+} (Delhaize et al., 2007; Peiter et al., 2007). Rice OsMTP1 localized in vacuole and is involved in the translocation and homeostasis of several divalent metals (Zn^{2+}, Cd^{2+}, Co^{2+}, and Fe^{2+}) (Yuan et al., 2012; Menguer et al., 2013). The five Mn-CDF members (OsMTP8.1, OsMTP8.2, OsMTP9, OsMTP11, and OsMTP11.1) were characterized to participate in the maintenance of Mn homeostasis in rice (Chen et al., 2013; Ueno et al., 2015; Takemoto et al., 2017; Ma et al., 2018; Tsunemitsu et al., 2018).

Abbreviations: CDF, Cation diffusion facilitator; CDS, Coding sequence; Chr, Chromosomes; GRAVY, Grand average of hydropathicity; FPKM, Fragments per kilobase of exon model per million mapped reads; Ka, The number of nonsynonymous substitutions per nonsynonymous site; Ks, The number of synonymous substitutions per synonymous site; MTP, Metal tolerance protein; MW, Molecular weight; pI, Isoelectric point; qRT-PCR, Quantitative real time PCR; TMDs, Transmembrane domains.

Peanut (*Arachis hypogaea* L., $2n = 4x = 40$) is one of the major oilseed crops grown throughout the tropics and subtropics regions. It provides both edible oil and food protein for people all over the world and accounts for 30% of the total oilseed production in China. Like other crops, peanut often encounters various metal stresses during its life history, which not only limits yield, but also threatens human health due to the accumulation of toxic metals (Liu et al., 2017; Yu et al., 2019). We have demonstrated that peanuts show wide cultivar variation in Cd accumulation and tolerance (Su et al., 2013, 2014); however, the underlying mechanism is not fully understood. Recently, the whole-genome sequences of the cultivated peanut (cv. *Tifrunner*) as well as the two wild ancestral species, *A. duranensis* and *A. ipaënsis*, have been released (Bertioli et al., 2016, 2019). Based on these important resources, genome-wide analysis has been conducted on some gene families such as *WRKY* (Song et al., 2016), *Mlo* (Mildew Locus O; Traore et al., 2021), growth-regulating factors (*GRF*s) (Zhao et al., 2019), and monosaccharide transporter (*MST*) (Wan et al., 2020). However, little progress has been made in peanut *MTP* genes.

To fill the knowledge gap, we identified 24 *AhMTP*s from cultivated peanut and characterized the structure and evolutionary relationship of these genes. Furthermore, two peanut cultivars (Fenghua 1 and Silihong) differing in Cd and Fe-deficiency tolerance (Liu et al., 2017; Cao et al., 2019; Yu et al., 2019) were used for evaluating the expression of *AhMTP* genes in response to several divalent metals (Fe^{2+}, Mn^{2+}, Zn^{2+}, and Cd^{2+}). Our findings are expected to provide a perspective on the evolution of *MTP* genes in peanut and were helpful for further functional characterization of *AhMTP* genes, shedding some light on the molecular mechanisms of metal transport and homeostasis in peanut.

MATERIALS AND METHODS

Identification of *MTP* Genes in Peanut

To identify peanut *MTP* genes, the protein sequences of *Arabidopsis* (12 genes) and rice (10 genes) were obtained from phytozome[1] and were used as queries for TBLASTP against the peanut genome on PeanutBase.[2] All retrieved MTP protein sequences were examined with the hmmscan tool,[3] and the candidates containing cation efflux domain (PF01545) were recognized as MTP proteins.

Phylogenetic Analysis

The MTP protein sequences of peanut, *Arabidopsis*, rice, and cucumber (*Cucumis sativus*) were aligned by clustalw in MEGA-X program (v10.2.6). The aligned files were used to construct a phylogenetic tree of the family members using the with neighbor-joining (NJ) method based on the p-distance model with 1,000

[1]https://phytozome-next.jgi.doe.gov
[2]https://peanutbase.org
[3]https://www.ebi.ac.uk/Tools/hmmer/search/hmmscan

bootstrap replicates. The constructed evolutionary tree was displayed and manipulated using an online software iTOL.[4]

Physicochemical Properties and Structure Characteristics of AhMTP Proteins

Physiochemical properties of AhMTP proteins including molecular weight (MW), amino acid number, grand average of hydropathicity (GRAVY), instability, aliphatic index, and isoelectric points (pI) were analyzed using ProtParam tool[5] (Duvaud et al., 2021). TMDs of AhMTP proteins were predicted using TOPCONS[6] (Tsirigos et al., 2015). Subcellular localization of proteins was predicted using Plant-mPLoc[7] (Chou and Shen, 2010). The conserved motifs and domains in AhMTP sequences were examined using the MEME v. 5.3.3[8] and Pfam tool,[9] respectively (Bailey et al., 2006; Mistry et al., 2020). Homology-modeled 3D structures of AhMTP proteins were predicted using the SwissModel[10] (Waterhouse et al., 2018).

Structure, Duplication, and Ka/Ks of *AhMTP* Genes

The exon-intron structure of all *AhMTP* genes was determined using GSDS v. 2.0[11] (Hu et al., 2015). Gene collinearity and Ka/Ks (ratios of the number of nonsynonymous substitutions per nonsynonymous site to the number of synonymous substitutions per synonymous site) were analyzed by One Step MCScanX and simple Ka/Ks calculator (NJ) of TBtools software, respectively (Chen et al., 2020). Diagrams of exon-intron organization and gene duplication event were drawn using TBtools software (Chen et al., 2020).

Cis-acting Regulatory Elements and MicroRNA Target Sites of *AhMTP* Genes

The coding and promoter (upstream 1.0 kb) sequences of *AhMTP* genes were retrieved from PeanutBase.[12] The promoter sequences were used for prediction of cis-acting regulatory elements (CREs) using PlantCARE (Lescot et al., 2002). The coding sequences were used for analyzing miRNA target sites by psRNATarget (Dai et al., 2018).

Expression Profiles of *AhMTP* Genes in Different Peanut Tissues

Expression profiles of *AhMTP* genes from cv. *Tifrunner* were identified using RNA-seq data obtained from PeanutBase (See footnote 12) (Clevenger et al., 2016). Read counts were transformed to FPKM (fragments per kilobase of exon per million aligned fragments), and the heatmap diagram was constructed with $\lg^{(FPKM+1)}$ using TBtools (Chen et al., 2020).

Plant Growth, Metal Determination, and qRT-PCR Analysis

Based on our previous studies (Liu et al., 2017; Cao et al., 2019; Yu et al., 2019), two peanut cultivars differing in Cd and Fe-deficiency tolerance, Fenghua 1 (Cd tolerant but sensitive to Fe deficiency), and Silihong (Cd sensitive but tolerant to iron deficiency) were used for determining relationships between the expression of *AhMTP* genes and metal tolerance and accumulation in peanut plants. The two cultivars are widely cultivated in the main production area of peanut in China, having close genetic relationship but differing in metal accumulation, translocation, and tolerance and therefore, being good model plants to study the mechanisms of metal translocation (Liu et al., 2017; Cao et al., 2019; Yu et al., 2019). After surface sterilized with 5% sodium hypochlorite (1 min), seeds were presoaked in distilled water for 24 h and then, they were sown in sand for germination. Three-day-old uniform seedlings were transferred to polyethylene pots and cultured as previously reported (Su et al., 2014). The 10-day-old seedlings were treated with 0.1 mM $CdCl_2$, 0.5 mM $FeSO_4$, 1 mM $MnSO_4$, or 0.5 mM $ZnSO_4$ in hydroponic cultures, with those without additional metals as the control. The experiment was arranged in a randomized complete design with triplications (pots) for each treatment. Each replication includes three seedlings. During the growing period, pots were randomly arranged and moved daily for minimizing position effects. After 4 days of metal exposure, plants were harvested and fresh root tissues were sampled for qRT-PCR analysis.

The harvested plants were separated into roots and shoots, and then, the roots were rinsed with 20 mM Na_2EDTA for 15 min to remove surface-bound metals. After oven-drying, dry weight (DW) of roots and shoots was weighed, and thereafter, tissues were ground into powder. Root (0.1 g) and shoot (0.5 g) samples were digested with HNO_3–$HClO_4$ (3:1, *v*/*v*) as the method described by Su et al. (2014). Concentrations of Fe, Zn, Cd, and Mn were determined by flame atomic absorbance spectrometry (WFX-210, Beijing Rayleigh Analytical Instrument Company, China). Metal translocation from roots to shoots was indicated as the percentage of metal in the shoot, which was calculated as following equation:

$$\text{Percentage of metal in shoots}(\%) = 100 \times \frac{\text{shoot DW} \times \text{metal conc.in shoots}}{\left(\text{shoot DW} \times \text{metal conc.in shoots} + \text{root DW} \times \text{metal conc.in roots}\right)}$$

To investigate the expression of *AhMTP* genes in responses to excessive metal stress, the first homolog of each *AhMTP* gene as well as *AhMTP12* was selected for qRT-PCR analysis as the method described previously (Cao et al., 2019), with *Ah60S* as the endogenous reference gene. Primers are listed in **Supplementary Table 1**. Three technical replications were carried out for each sample. The relative gene expression was calculated using the $2^{-\Delta\Delta CT}$ method.

[4]https://itol.embl.de/itol.cgi
[5]https://web.expasy.org/protparam/
[6]http://topcons.net/
[7]http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/
[8]https://meme-suite.org/meme/tools/meme
[9]http://pfam.xfam.org/search#tabview=tab1
[10]http://www.swissmodel.expasy.org
[11]http://gsds.gao-lab.org/
[12]https://www.peanutbase.org/gene_expression/atlas

Statistical Analysis

Data were subjected to one-way analysis of variance, and significant variations among means were determined by the Duncan's Multiple Range Test at a probability level of 5%. Stepwise linear regression analysis was performed on the percentage of metal in shoots and expression of *AhMTP* genes. All statistical analysis was conducted using IBM SPSS Statistics v. 22 (IBM, New York, United States).

RESULTS

Summary of the *AhMTP* Gene Family in Peanut

A total of 24 genes were identified in peanut, including three homologous genes of *AhMTP1* (*AhMTP1.1*/*1.2*/*1.3*), four homologous genes of *AhMTP2* (*AhMTP2.1*/*2.2*/*2.3*/*2.4*), two homologous genes of *AhMTP4* (*AhMTP4.1*/*4.2*), two homologous genes of *AhMTP9* (*AhMTP9.1*/*9.2*), two homologous genes of *AhMTP10* (*AhMTP10.1*/*10.2*), four homologous genes of *AhMTP11* (*AhMTP11.1*/*11.2*/*11.3*/*11.4*), two homologous genes of *AhMTPC2* (*AhMTPC2.1*/*C2.2*), two homologous genes of *AhMTPC4* (*AhMTPC4.1*/*C4.2*), two homologous genes of *AhMTPB* (*AhMTPB1*/*B2*), and *AhMTP12* (**Table 1**). The length of *AhMTP* genes varied from 970 bp (*AhMTP1.3*) to 9,270 bp (*AhMTP2.3*), with CDS lengths from 345 bp (*AhMTP1.3*) to 2,604 bp (*AhMTP12*). The amino acid number of AhMTP proteins ranged from 114 (AhMTP1.3) to 867 (AhMTP12), and the molecular weight of AhMTP proteins varied from 12.20 kDa (AhMTP1.3) to 97.04 kDa (AhMTP12). The instability, aliphatic index, and GRAVY of the AhMTPs ranged from 25.32 (AhMTP1.2) to 48.77 (AhMTP1.3), from 91.56 (AhMTP10.1) to 120.25 (AhMTP2.3), and from-0.221 (AhMTP9.1) to 0.474 (AhMTP2.3), respectively. The isoelectric point (pI) ranged from 4.92 (AhMTP11.3) to 8.67 (AhMTPC4.1), with 15 AhMTP members pI <7 and 9 AhMTP members pI >7 (**Table 1**). The number of TMDs showed a wide variation among AhMTP proteins, and most AhMTPs contained 4–6 TMDs (**Table 1**). All AhMTP proteins were predicted to localize to vacuole membranes (**Table 1**).

Phylogenetic Analysis of MTP Gene Families

Phylogenetic relationship of 53 MTPs from peanut, *Arabidopsis*, rice, and cucumber was analyzed with the NJ method. These MTP proteins were divided into seven groups (1, 5, 6, 7, 8, 9, and 12), belonging to three major sub-families (Zn-CDFs, Zn/Fe-CDFs, and Mn-CDFs; **Figure 1**). Of the seven primary groups, group 9 is the largest one containing eight AhMTPs (AhMTP9.1/9.2, AhMTP10.1/10.2, and AhMTP11.1/11.2/11.3/11.4), followed by group 1 (AhMTP1.1/1.2/1.3 and AhMTPB1/B2) and 6 (AhMTP2.1/2.2/2.3/2.4), while group 12 is the smallest group with only one member (AhMTP12). The remaining three groups contained two AhMTP members each (**Figure 1**).

Structure and Duplication of *AhMTP* Genes

To gain insight into the evolution of the *MTP* family in peanut, exon-intron organizations of *AhMTP* genes were examined. As showed in **Figure 2A**, *AhMTP* genes belonging to the same groups showed similar exon-intron organizations, which coincided with the results obtained from the phylogenetic analysis. The group 1 and 12 of Zn-CDFs contained only one exon (with one or without intron), whereas group 5 possessed 10 exons (nine intron). Zn/Fe-CDF genes contained 12–13 exons except *AhMTP2.3*/*2.4*, which harbored seven exons. The group 8 of Mn-CDFs possessed seven exons, while most genes of group 9 contained six exons except *AhMTP9.1* and *AhMTP9.2*, in which five and four exons were harbored, respectively.

AhMTP genes were located unevenly in 16 chromosomes. A total of 13 and 11 *AhMTP* genes were identified from the subgenome A (Chr.01–10) and B (Chr.11–20), respectively (**Figure 2B**). The Chr.06 contained three *AhMTP* genes, the Chr. 02, 04, 07, 12, 14, and 18 had two genes each, each of the nine chromosomes (01, 03, 09, 10, 11, 13, 16, 19, and 20) contained only one gene, while no *AhMTP* gene was identified in the Chr. 05, 08, 15, and 17. All *AhMTP* genes that located in the same chromosome showed greatly large distances (**Figure 2B**).

Collinearity analysis revealed that almost all *AhMTP* genes experienced gene duplication events except *AhMTP12*, resulting in 13 gene pairs (**Figure 2B**). Interestingly, all the 11 *AhMTP* genes of the subgenome B were crossly collineared with corresponding genes of subgenome A, forming 11 gene pairs including *AhMTPC4.1/C4.2*, *AhMTP9.1/9.2*, *AhMTP2.1/2.2*, *AhMTP1.1/1.2*, *AhMTPC2.1/C2.2*, *AhMTP11.3/11.2*, *AhMTP11.4/11.1*, *AhMTPB1/B2*, *AhMTP10.1/10.2*, *AhMTP4.1/4.2*, and *AhMTP2.4/2.3*. These collinear blocks result from whole-genome duplications (WGDs). In addition, *AhMTP1.3* was also identified to collinear with *AhMTP1.1* and *AhMTP1.2*, respectively. The pair of *AhMTP1.2/1.3* might be result from segmental duplication because they are in different chromosomes of the same subgenome. No tandem duplication was detected in *AhMTP* genes. The Ka/Ks ratios of all gene duplication pairs were less than 1 (**Table 2**), indicating that *AhMTP* genes evolved under purifying selection (Hurst, 2002).

Conserved Motifs, Domain Architectures, and Models of AhMTP Proteins

AhMTP proteins contained a total of 15 conserved motifs, among them, eight motifs (3, 4, 5, 6, 10, 11, 13, and 14) were annotated to be the TMD or CTD of cation_efflux domains according to the InterProScan tools (**Figure 3A** and **Supplementary Table 2**). It was observed that conserved motifs were specifically distributed in the members of different cluster or group (**Figure 3A**). Two motifs (6 and 11) were shared by eight AhMTP proteins belonging to the Zn-CDF cluster. Interestingly, 10 members of the Mn-CDF cluster contained a specific motif complex composed of motifs 1, 2, 4, 5, and 9. The six proteins of the Fe/Zn-CDF subfamily widely varied

TABLE 1 | Physicochemical properties and subcellular localization of the 24 metal tolerance proteins (MTPs) identified in peanut.

Gene name	Gene ID	Gene length (bp)	CDS length (bp)	MW (kDa)	aa	Instability	Aliphatic index	GRAVY	PI	TMD	Subcellular localization
AhMTP1.1	arahy.ARQ7QI	3,937	1,287	47.14	428	27.63	99.09	−0.064	5.91	6	Vacuole
AhMTP1.2	arahy.854BZC	3,902	1,293	47.39	430	25.32	98.63	−0.072	5.95	6	Vacuole
AhMTP1.3	arahy.NB6CJE	970	345	12.20	114	48.77	100.18	0.259	5.28	2	Vacuole
AhMTP2.1	arahy.5I5ID6	5,188	1,509	54.78	502	42.25	95.62	−0.090	6.75	5	Vacuole
AhMTP2.2	arahy.L0WASJ	5,250	1,509	54.80	502	43.11	95.24	−0.098	6.67	5	Vacuole
AhMTP2.3	arahy.INQ3MB	9,270	828	30.33	275	26.19	120.25	0.474	8.15	6	Cell/Vacuole membrane
AhMTP2.4	arahy.XK0FMC	3,180	1,059	39.83	352	44.40	99.72	0.170	6.72	3	Vacuole
AhMTP4.1	arahy.N1ZWYI	4,890	1,266	47.34	421	46.25	108.86	0.005	5.87	5	Vacuole
AhMTP4.2	arahy.LH0XEN	5,012	1,266	47.38	421	46.32	108.86	0.000	5.93	5	Vacuole
AhMTP9.1	arahy.1E45SP	8,680	1,245	47.62	414	45.50	93.53	−0.221	8.20	4	Vacuole
AhMTP9.2	arahy.3T6XNU	5,968	801	30.45	266	38.28	104.14	0.056	6.31	4	Vacuole
AhMTP10.1	arahy.DW8A2N	4,985	1,236	46.78	411	43.20	91.56	−0.109	7.24	6	Cell/Vacuole membrane
AhMTP10.2	arahy.JR9DQ2	4,952	1,236	47.01	411	46.26	91.56	−0.129	8.64	6	Cell/Vacuole membrane
AhMTP11.1	arahy.J9LUB0	5,202	1,191	44.99	396	47.43	100.71	−0.009	4.97	6	Vacuole
AhMTP11.2	arahy.C1W7MD	4,156	1,215	46.04	404	45.81	101.36	0.008	5.04	6	Vacuole
AhMTP11.3	arahy.F3QZDX	4,293	1,191	44.97	396	46.11	100.71	−0.011	4.92	6	Cell/Vacuole membrane
AhMTP11.4	arahy.HJ8JHV	5,202	1,191	44.99	396	47.43	100.71	−0.009	4.97	6	Vacuole
AhMTP12	arahy.LDE0EK	2,605	2,604	97.04	867	43.83	97.01	0.033	7.20	16	Vacuole
AhMTPB1	arahy.IEEP0H	1,165	1,164	42.85	387	35.35	109.53	0.199	5.89	6	Vacuole
AhMTPB2	arahy.WGY3ZU	1,168	1,158	42.47	385	35.86	111.38	0.246	5.83	6	Vacuole
AhMTPC2.1	arahy.C1QR5S	5,448	1,143	42.39	380	43.04	98.50	0.149	8.33	5	Vacuole
AhMTPC2.2	arahy.F8DN54	5,482	1,143	42.39	380	43.04	98.50	0.149	8.33	5	Vacuole
AhMTPC4.1	arahy.2A2X48	4,032	1,329	49.01	442	33.62	94.66	0.005	8.67	4	Vacuole
AhMTPC4.2	arahy.660VQ8	4,544	1,326	48.88	441	33.48	94.88	−0.005	7.73	6	Vacuole

MW, aa, GRAVY, PI, and TMD refer to molecular weight, amino acid number, grand average of hydropathicity, isoelectric points, and transmembrane domain, respectively.

in motif distribution and shared two motifs (motif 11 and 15; **Figure 3A**).

Conserved domain analysis revealed that all AhMTP proteins except AhMTP2.4 contained the typical domain of MTP, cation efflux domain (**Figure 3B**). Additionally, a dimerization domain of zinc transporters, ZT-dimer, was also identified in all members of Mn-CDF cluster and three members (AhMTP2.1/2.2/2.4) of group 6 belonging to Fe/Zn-CDFs (**Figure 3B**).

To better understand the functions of AhMTP proteins, their sequences were modeled using the SwissModel. As showed in **Table 3**, the sequence identity was ranged from 16.25 to 42.86%, the value of GMQE ranged from 0.19 to 0.53, and QMEANDisCo global score ranged from 0.41 to 0.70. These data suggested a high quality of a 3D protein structure models of AhMTP proteins, which were presented as **Supplementary Figure 1**. The 3D-model prediction revealed that the three subfamily of AhMTPs differed from each other in the protein structure (**Table 3**). All members of Zn-CDFs are best modeled with the template, 6xpd.1 (Cryo-EM structure of human ZnT8 double mutant—D110N and D224N, determined in outward-facing conformation). The best model for all Mn-CDF members was 3h90.1 (Structural basis for the autoregulation of the zinc transporter YiiP). Zn/Fe-CDFs were commendably modeled to 5vrf.1 (Cryo-EM structure of the Zinc transporter YiiP from helical crystals) except AhMTP2.4, of which the best model was 3h90.1.

CREs and the MicroRNA Target Sites of *AhMTP* Genes

To explore the probable post-transcriptional regulation of the *AhMTP* genes, their CREs and microRNA target sites were predicted. A total of 1,548 CREs were identified in the promoter of *AhMTP* genes, and most of them were identified to associate with gene transcription (1,246), light response (155), phytohormonal response (83), and abiotic stress (40) (**Table 4** and **Supplementary Table 3**). The promoter of all *AhMTP* genes harbored gene transcription (CAAT and TATA-box) and light-responsive elements whose number ranged from 12 to 110 and from 1 to 16, respectively. Most of *AhMTP* genes have phytohormone responsive elements in the promoters except *AhMTP1.1*, *AhMTP1.2*, *AhMTP2.4*,

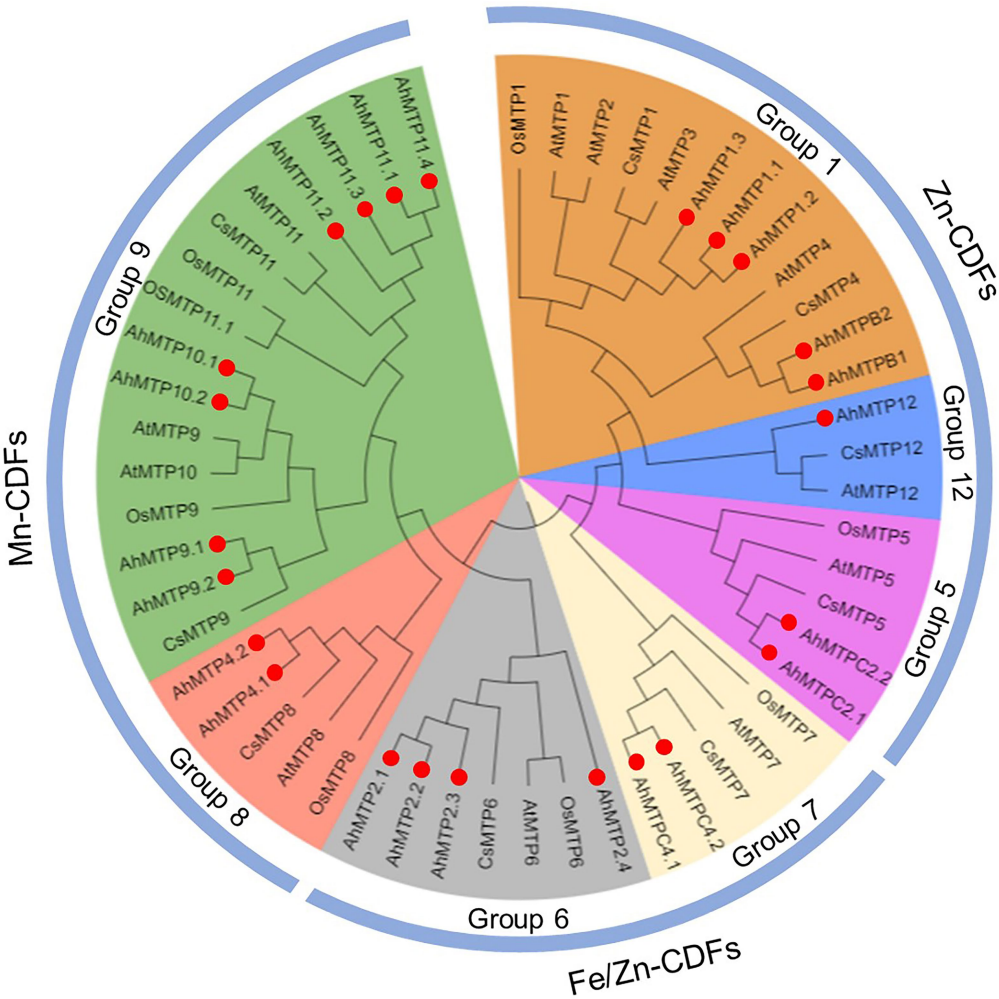


FIGURE 1 | Phylogenetic relationships of MTP proteins in peanut and other plant species. Red solid circles represent the 24 AhMTP proteins of peanut.

AhMTPC2.1, and *AhMTPC4.1*. Nineteen *AhMTP* genes were identified to contain biotic stress elements, including ARE (*AhMTP1.1/1.2/1.3*, *AhMTP2.1/2.2*, *AhMTPC2.1/C2.2*, *AhMTP11.1/11.3/11.4*, *AhMTP9.2*, *AhMTP10.1/10.2*, *AhMTP12*), LTR (*AhMTP1.1/1.2*, *AhMTP9.2*, *AhMTP10.1/10.2*, and *AhMTPC4.1/C4.2*), MBS (*AhMTP1.1/1.2*, *AhMTPC2.1*, *AhMTP12*, and *AhMTP4.1*), and TC-rich repeats (*AhMTPC2.1*, *AhMTP1.3*, and *AhMTP2.2*).

A total of five miRNAs were identified, with eight genes belonging to three groups as target genes (**Table 5**). The UPE varied from 14.877 (*ahy-miR156a*/*AhMTPC2.1*) to 21.179 (*ahy-miR3519*/*AhMTP11.2*). Two gene pairs of group 9, *AhMTP9.1/9.2* and *AhMTP11.2/11.3*, were predicted to be targets of *ahy-miR167-5p* and *ahy-miR3519*, respectively. The two members of group 5, *AhMTPC2.1* and *AhMTPC2.2*, were potential target genes of *ahy-miR156a* and *ahy-miR156c*. Possible target sites of *ahy-miR3511-3p* were identified in two genes of group 8, *AhMTP4.1* and *AhMTP4.2*. All identified miRNA targeted genes were inhibited by the corresponding miRNA in a cleavage manner.

Expression Profiles of *AhMTP* Genes in Different Tissues of Peanut

RNA-seq data showed that most of the *AhMTP* genes expressed in peanut tissues except *AhMTPB1*, *AhMTPB2*, and *AhMTP12*, which did not express in all the 22 tissues (**Supplementary Table 4**). To better understand the gene expression profiles, a hierarchical cluster analysis was carried out. As presented in **Figure 4**, 24 *AhMTP* genes were divided into two clusters: cluster I and II. Cluster I includes 12 genes with low expression levels. Among them, six genes showed very low expression (*AhMTP2.3*/*2.4* and *AhMTP11.4*) or did not express (*AhMTPB1*/*B2* and *AhMTP12*) in any tissues. The other six genes exhibit tissue-specific expressions with low levels. *AhMTP11.1* mainly expressed in peg tip to fruit, seed, and pericarp; *AhMTP11.3* expressed in all tissues but relatively higher in shoot tips and roots, and *AhMTP1.3* mainly expressed in fruit.

Cluster II consists of six pair of *AhMTP* genes including *AhMTP1.1*/*1.2*, *AhMTP2.1*/*2.2*, *AhMTP4.1*/*4.2*, *AhMTPC2.1*/*C2.2*, *AhMTPC4.1*/*C4.2*, and *AhMTP10.1*/*10.2*, with high expression levels in approximately all tissues. Each pair of genes showed a

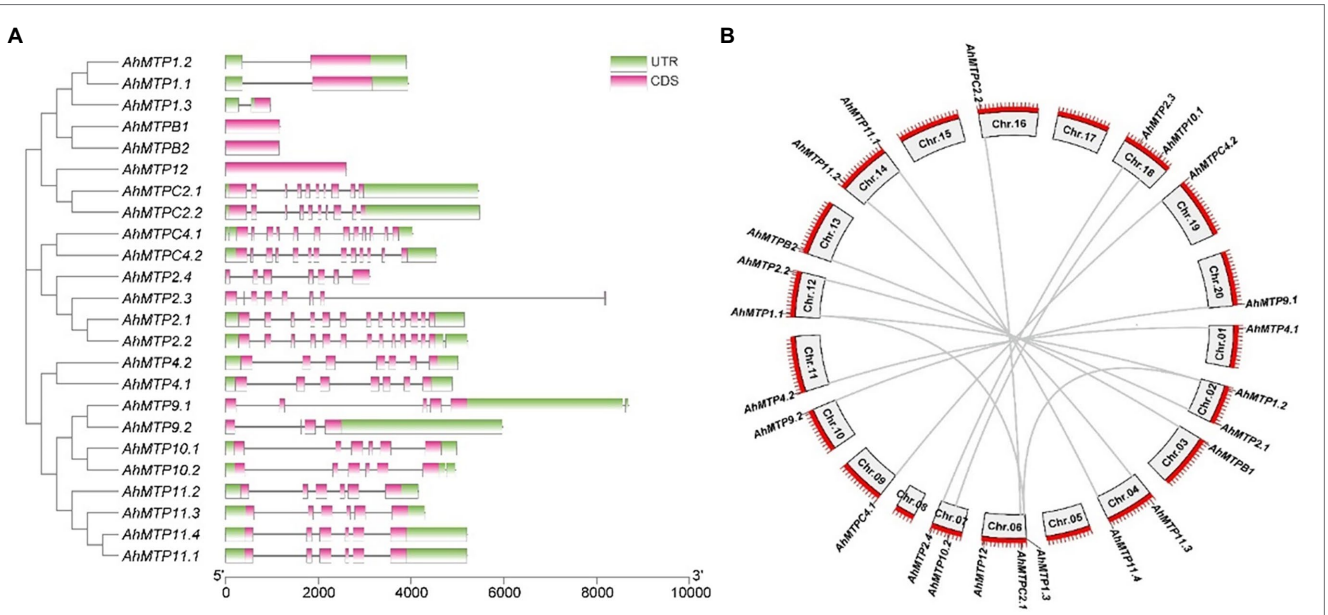


FIGURE 2 | Phylogenetic relationships, exon-intron organization, and collinearity analysis of AhMTP genes from peanut. **(A)** Phylogenetic tree and exon-intron structure of *AhMTP* genes. **(B)** Chromosomal locations and duplications of *AhMTP* genes.

TABLE 2 | Ka/Ks analysis of all gene duplication pairs for *AhMTP* genes.

Duplicated pair	Duplicate type	Ka	Ks	Ka/Ks	Positive selection
AhMTPC4.1/C4.2	Whole-genome	0.005	0.023	0.231	No
AhMTP9.2/9.1	Whole-genome	0.137	0.210	0.650	No
AhMTP2.1/2.2	Whole-genome	0.003	0.023	0.113	No
AhMTP1.2/1.3	Segmental	0.107	0.711	0.150	No
AhMTP1.2/1.1	Segmental	0.008	0.058	0.141	No
AhMTPC2.1/C2.2	Whole-genome	0.000	0.041	0.000	No
AhMTP11.3/11.2	Whole-genome	0.007	0.013	0.563	No
AhMTP11.4/11.1	Whole-genome	0.000	0.000	NaN	No
AhMTPB1/B2	Whole-genome	0.008	0.028	0.286	No
AhMTP10.2/10.1	Whole-genome	0.005	0.026	0.206	No
AhMTP4.1/4.2	Whole-genome	0.003	0.037	0.086	No
AhMTP1.3/1.1	Whole-genome	0.116	0.682	0.170	No
AhMTP2.4/2.3	Whole-genome	0.480	0.558	0.859	No

similar expression pattern. *AhMTP1.1*/*1.2* is specifically and highly expressed in shoot tips and seed. *AhMTP4.1*/*4.2* highly expressed in roots, nodules, stamens, peg tips, fruit, peg tip to fruit, pericarp, and seed. *AhMTP10.1*/*10.2* highly expressed in leaves, perianth, and pistil. The highest expression of *AhMTPC2.1*/C*2.2* was in seed.

Differential Expression of *AhMTP* Genes in the Root of Two Peanut Cultivars

To investigate the expression of *AhMTP* genes in responses to excessive metal stress, the first homolog of each *AhMTP* gene as well as *AhMTP12* was selected for qRT-PCR analysis. The two cultivars differ from each other in the expression of *AhMTP* genes as well as their responses to excessive metal stress (**Figure 5**). Under normal nutrition condition, Fenghua 1 showed higher expressions of *AhMTPB1*, *AhMTP9.1*, *AhMTP12*, *AhMTPC2.1*, and *AhMTPC4.1* than Silihong, while the expression of *AhMTP1.1*, *AhMTP2.1*, and *AhMTP4.1* was higher in Silihong than in Fenghua 1. Cd enhanced the expression of *AhMTPB1*, *AhMTP12*, *AhMTPC2.1*, *AhMTP10.1*, and *AhMTP11.1* for both cultivars, while other genes showed different responses to Cd exposure between the two cultivars. By contrast, Cd induced increase of gene expression was more pronounced in Fenghua 1 than that in Silihong. Although Cd induced expressions for most *AhMTP* genes, some genes such as *AhMTP9.1* and *AhMTPC4.1* in Fenghua 1 as well as *AhMTP2.1* and *AhMTP4.1* in Silihong were observed to be repressed by Cd, while *AhMTP9.1* and *AhMTP1.1* in Silihong remained unaffected.

Excess Fe upregulated the expression of *AhMTP12*, *AhMTPC2.1*, *AhMTPC4.1*, *AhMTP2.1*, *AhMTP10.1*, and *AhMTP11.1* for both cultivars (**Figure 5**). The induction was more pronounced in Fenghua 1 than that in Silihong for most genes. *AhMTP1.1* and *AhMTP4.1* were significantly induced by Fe stress in Fenghua 1, whereas in Silihong, expressions of these genes were

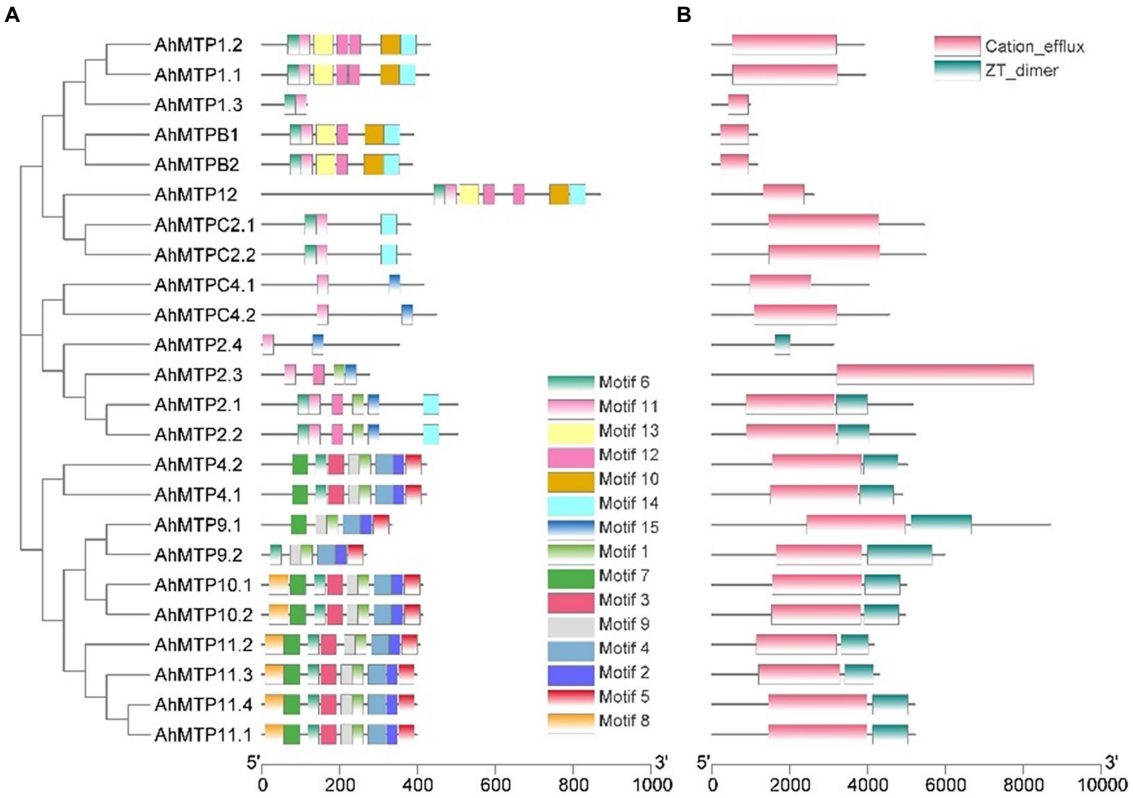


FIGURE 3 | Distributions of the conserved motifs **(A)** and domains **(B)** in AhMTP proteins from peanut.

unchanged. *AhMTP9.1* and *AhMTPB1* were induced by Fe stress in Silihong but were unchanged or downregulated in Fenghua 1 (**Figure 5**).

Excess Mn increased the expression of *AhMTP12*, *AhMTPC4.1*, *AhMTPC2.1*, and *AhMTP10.1* for both cultivars (**Figure 5**). The expression of *AhMTPB1* and *AhMTP9.1* was increased by Mn in Silihong; however, it was repressed in Fenghua 1. Conversely, the expression of *AhMTP1.1*, *AhMTP2.1*, and *AhMTP4.1* was repressed by excess Mn in Silihong, while in Fenghua 1, they were unaffected. Mn exposure downregulated the expression of *AhMTPB1* in Fenghua but not in Silihong.

Excess Zn induced the expression of *AhMTP11.1*, *AhMTP12*, *AhMTPC2.1*, and *AhMTPC4.1* but repressed that of *AhMTP9.1* for both cultivars, while *AhMTP4.1* was unaffected (**Figure 5**). The remaining four *AhMTP* genes exhibit cultivar-specific patterns in response to Zn stress. *AhMTP1.1* and *AhMTP2.1* were induced by Zn in Fenghua 1 but were unaffected or repressed in Silihong. Zn dramatically increased the expression of *AhMTPB1* and *AhMTP10.1* in Silihong but did not affect that in Fenghua 1 (**Figure 5**).

Differential Metal Translocation in Plants of Two Peanut Cultivars

To investigate differences of metal translocation between the two peanuts, the percentage of metal in shoots was considered as an indicator for the translocation capability of metals from roots to shoots. Generally, Silihong exhibited higher capacity for the translocation of Cd and Mn from roots to shoots, while Fenghua 1 showed a higher capability of Fe translocation (**Figures 6A–C**). There are interactive effects between cultivar and metal treatment on the percentage of Mn ($F = 30.34$, $p = 0.000$) and Zn ($F = 65.97$, $p = 0.000$) in shoots, indicating that cultivar differences in the translocation of Mn and Zn from roots to shoots were dependent on metal exposure (**Figures 6C,D**). Exposure of excess Cd, Fe, Zn, and Mn significantly reduced Fe translocation from roots to shoots for both cultivars, which was more pronounced in Fe and Mn treatments. With regard to Mn, its translocation in Silihong was increased by all metal treatments, whereas in Fenghua 1, it was increased by Cd and Mn treatments but decreased by the Zn treatment. Exposure of excess Cd and Fe did not affect the translocation of Zn for both cultivars; however, it was decreased by the Zn treatment. Mn exposure dramatically reduced Zn translocation in Silihong but unaffected that in Fenghua 1. A significant and negative correlation was observed between the percentage of Fe and Mn in shoots ($r = -0.498$, $n = 30$, $p = 0.005$).

To determine whether *AhMTP* genes were involved in metal translocation in peanut plants, a stepwise linear regression analysis was performed on the percentage of metal in shoots and expression of *AhMTP* genes ($-\Delta\Delta CT$). As showed in **Table 6**, the percentage of Fe in shoots significantly and positively correlated with the expression of *AhMTP4.1*, *AhMTP9.1,* and *AhMTPC4.1* but negatively correlated with that

TABLE 3 | The best templates of peanut AhMTP proteins selected from the SwissModel template library for building 3D structure models.

Protein name	Template	Sequence identity (%)	Coverage	GMQE	QMEANDisCo Global	Description
MTP1.2	6xpd.1	36.86	A44-430	0.49	0.61 ± 0.05	Zinc transporter 8
MTP1.1	6xpd.1	36.86	A44-428	0.49	0.61 ± 0.05	Zinc transporter 8
MTP1.3	6xpd.1	42.86	A37-113	0.47	0.70 ± 0.11	Zinc transporter 8
MTPB1	6xpd.1	32.18	A53-387	0.53	0.62 ± 0.05	Zinc transporter 8
MTPB2	6xpd.1	32.53	A53-385	0.52	0.61 ± 0.05	Zinc transporter 8
MTPC2.1	6xpd.1	16.25	A94-379	0.44	0.54 ± 0.05	Zinc transporter 8
MTPC2.2	6xpd.1	16.25	A94-379	0.44	0.54 ± 0.05	Zinc transporter 8
MTP12	6xpd.1	24.83	A421-866	0.19	0.49 ± 0.05	Zinc transporter 8
MTPC4.1	5vrf.1	17.36	A96-408	0.33	0.41 ± 0.05	Cadmium and zinc efflux pump FieF
MTPC4.2	5vrf.1	17.48	A95-440	0.36	0.43 ± 0.05	Cadmium and zinc efflux pump FieF
MTP2.1	5vrf.1	27.62	A75-385	0.37	0.55 ± 0.05	Cadmium and zinc efflux pump FieF
MTP2.2	5vrf.1	27.62	A75-385	0.37	0.54 ± 0.05	Cadmium and zinc efflux pump FieF
MTP2.3	5vrf.1	24.86	A41-259	0.39	0.50 ± 0.06	Cadmium and zinc efflux pump FieF
MTP2.4	3h90.1	20.00	A2-246	0.33	0.41 ± 0.05	Ferrous-iron efflux pump fieF
MTP11.4	3h90.1	19.71	A106-386	0.40	0.49 ± 0.05	Ferrous-iron efflux pump fieF
MTP11.1	3h90.1	19.71	A106-386	0.40	0.49 ± 0.05	Ferrous-iron efflux pump fieF
MTP11.3	3h90.1	19.27	A106-387	0.40	0.50 ± 0.05	Ferrous-iron efflux pump fieF
MTP11.2	3h90.1	18.91	A106-395	0.39	0.48 ± 0.05	Ferrous-iron efflux pump fieF
MTP9.1	3h90.1	18.06	A89-322	0.38	0.47 ± 0.05	Ferrous-iron efflux pump fieF
MTP9.2	3h90.1	18.62	A8-256	0.49	0.45 ± 0.05	Ferrous-iron efflux pump fieF
MTP10.1	3h90.1	18.91	A122-402	0.38	0.49 ± 0.05	Ferrous-iron efflux pump fieF
MTP10.2	3h90.1	19.27	A122-402	0.38	0.49 ± 0.05	Ferrous-iron efflux pump fieF
MTP4.2	3h90.1	21.74	A124-405	0.36	0.50 ± 0.05	Ferrous-iron efflux pump fieF
MTP4.1	3h90.1	21.74	A124-405	0.36	0.49 ± 0.05	Ferrous-iron efflux pump fieF

of *AhMTPC2.1* and *AhMTP12*. The expression of *AhMTP1.1* showed a significant and negative correlation with the percentage of Mn in shoots. The percentage of Zn in shoots was significantly and positively correlated with the expression of *AhMTP2.1* but was negatively correlated with that of *AhMTPC2.1*.

DISCUSSION

Genome-wide identification of MTP proteins has been extensively performed in diverse plant species, including *Arabidopsis* (Montanini et al., 2007), sweet orange (*Citrus sinensis*) (Fu et al., 2017), wheat (Vatansever et al., 2017), turnip (*Brassica rapa* var. *rapa*) (Li et al., 2018), tobacco (*Nicotiana tabacum*) (Liu et al., 2019), grape (*Vitis vinifera*) (Shirazi et al., 2019), and *Populus trichocarpa* (Gao et al., 2020). However, there is little information about MTP family in peanut that limits understanding the molecular mechanisms underlying the regulation of metal homeostasis. Herein, 24 putative *AhMTP* genes were identified in cultivated peanut, which were divided into seven groups (1, 5, 6, 7, 8, 9, and 12), belonging to three major substrate-specific groups (Zn-CDFs, Zn/Fe-CDFs, and Mn-CDFs) (**Figure 1**). Our findings concurred with the results obtained from *Arabidopsis* (Montanini et al., 2007), turnip (Li et al., 2018), tobacco (Liu et al., 2019), and *P. trichocarpa* (Gao et al., 2020), suggesting that AhMTPs may have similar functions to their homologs in these plant species.

In agreement with previous studies (Vatansever et al., 2017; Liu et al., 2019; Shirazi et al., 2019; Gao et al., 2020), most of AhMTP proteins were predicted to localize to vacuole membranes. All AhMTP proteins except AhMTP2.4 contained the typical domain of MTP, cation efflux domain. These results indicated that AhMTPs might be metal transporters playing key roles in the homeostasis or detoxification of divalent metals. Although AhMTP2.4 does not contain cation efflux domain, collinearity analysis revealed that it is a whole-genome duplicated gene of *AhMTP2.3* (**Figure 2B**). Moreover, phylogenetic analysis

TABLE 4 | Frequency and function of cis-regulatory elements (CREs) in the promoter regions of *AhMTP* genes.

Gene name	Gene transcription	Abiotic stress	Biotic stress	Tissue expression	Secondary metabolism	Phytohormonal response	Light response	Circadian control	Site-binding	Cell-cycle regulation
AhMTP1.1	46	4	0	0	0	0	2	0	0	1
AhMTP1.2	44	4	0	0	0	0	2	0	0	0
AhMTP1.3	37	2	0	0	0	7	8	0	1	0
AhMTP2.1	61	1	0	0	1	5	10	0	0	0
AhMTP2.2	68	2	0	0	1	7	9	0	0	0
AhMTP2.3	29	0	0	0	0	3	15	1	1	1
AhMTP2.4	41	0	1	1	0	0	7	0	0	0
AhMTP4.1	46	1	1	1	1	2	8	0	0	0
AhMTP4.2	110	0	0	0	0	2	3	1	0	0
AhMTP9.1	86	0	0	0	1	6	16	0	1	0
AhMTP9.2	75	0	0	0	0	1	6	0	0	0
AhMTP10.1	49	2	0	1	0	4	7	0	0	0
AhMTP10.2	45	3	0	0	1	7	8	0	0	0
AhMTP11.1	53	3	0	0	0	9	9	0	1	0
AhMTP11.2	25	0	0	1	0	4	4	0	2	0
AhMTP11.3	56	1	0	0	0	3	3	0	1	0
AhMTP11.4	53	3	0	0	0	9	9	0	1	0
AhMTP12	12	2	0	0	0	4	6	0	0	1
AhMTPB1	55	0	0	0	0	5	3	0	0	0
AhMTPB2	56	0	0	0	0	3	6	0	0	0
AhMTPC2.1	42	6	0	0	0	0	1	0	0	0
AhMTPC2.2	43	3	0	0	0	1	4	0	0	0
AhMTPC4.1	59	1	0	0	0	0	5	0	0	0
AhMTPC4.2	55	2	0	0	0	1	4	0	0	0

TABLE 5 | Prediction of miRNAs for *AhMTP* transcripts.

miRNA	Target	Expectation	UPE	miRNA aligned	Target aligned	Inhibition
ahy-miR167-5p	*AhMTP9.1*	4.5	16.384	UGAAGCUGCCA	GGGAUUAUGGU	Cleavage
ahy-miR167-5p	*AhMTP9.2*	4.5	16.874	GCAUGAUCUU UGAAGCUGCCA	GGCAGCAUCU GGGAUUAUGGU	Cleavage
ahy-miR3519	*AhMTP11.2*	4.5	21.179	GCAUGAUCUU UCAAUCAAUGA	GGCAGCAUCU GCAAAUUAUGU	Cleavage
ahy-miR3519	*AhMTP11.3*	4.5	19.995	CAGCAUUUCA UCAAUCAAUGA	CGAUGAUUGG GCAAAUUAUGU	Cleavage
ahy-miR156a	*AhMTPC2.1*	5	14.877	CAGCAUUUCA UGACAGAAGAG	CGAUGAUUGG AUGCCCUGUUU	Cleavage
ahy-miR156a	*AhMTPC2.2*	5	15.246	AGAGAGCAC UGACAGAAGAG	CUUUUGUUU AUGCCCUGUUU	Cleavage
ahy-miR156c	*AhMTPC2.1*	5	15.19	AGAGAGCAC UUGACAGAAGA	CUUUUGUUU AUGCCCUGUUU	Cleavage
ahy-miR156c	*AhMTPC2.2*	5	15.344	GAGAGAGCAC UUGACAGAAGA	CUUUUGUUUA AUGCCCUGUUU	Cleavage
ahy-miR3511-3p	*AhMTP4.1*	5	20.147	GAGAGAGCAC UGUUACUAUGG	CUUUUGUUUA GAAGUCGAUGC	Cleavage
ahy-miR3511-3p	*AhMTP4.2*	5	16.76	CAUCUGGUAA UGUUACUAUGG CAUCUGGUAA	CAUAGUGGCA GAAGUUGAUGC CAUAGUGGCA	Cleavage

indicated that AhMTP2.4 shows a close relationship with other members of group 6 (**Figure 1**). Therefore, it could be regarded as an AhMTP member.

A total of 10 AhMTPs were identified to be Mn-CDF members, which were further divided into two groups: group 8 and 9. These proteins share the following characteristics: (1) containing both the cation efflux domain and ZT-dimer; (2) shared a specific motif complex composed of motifs 1, 2, 4, 5, and 9, among them motifs 4 and 5 were annotated to be the TMD and CTD of cation efflux, respectively; (3) containing four to six TMDs; (4) best modeled with 3h90.1 (Structural basis for the autoregulation of the zinc transporter YiiP); and (5) localizing to vacuolar or plasma membranes. YiiP is a dimeric Zn^{2+}/H^{+} antiporter from *Escherichia coli* that selectively

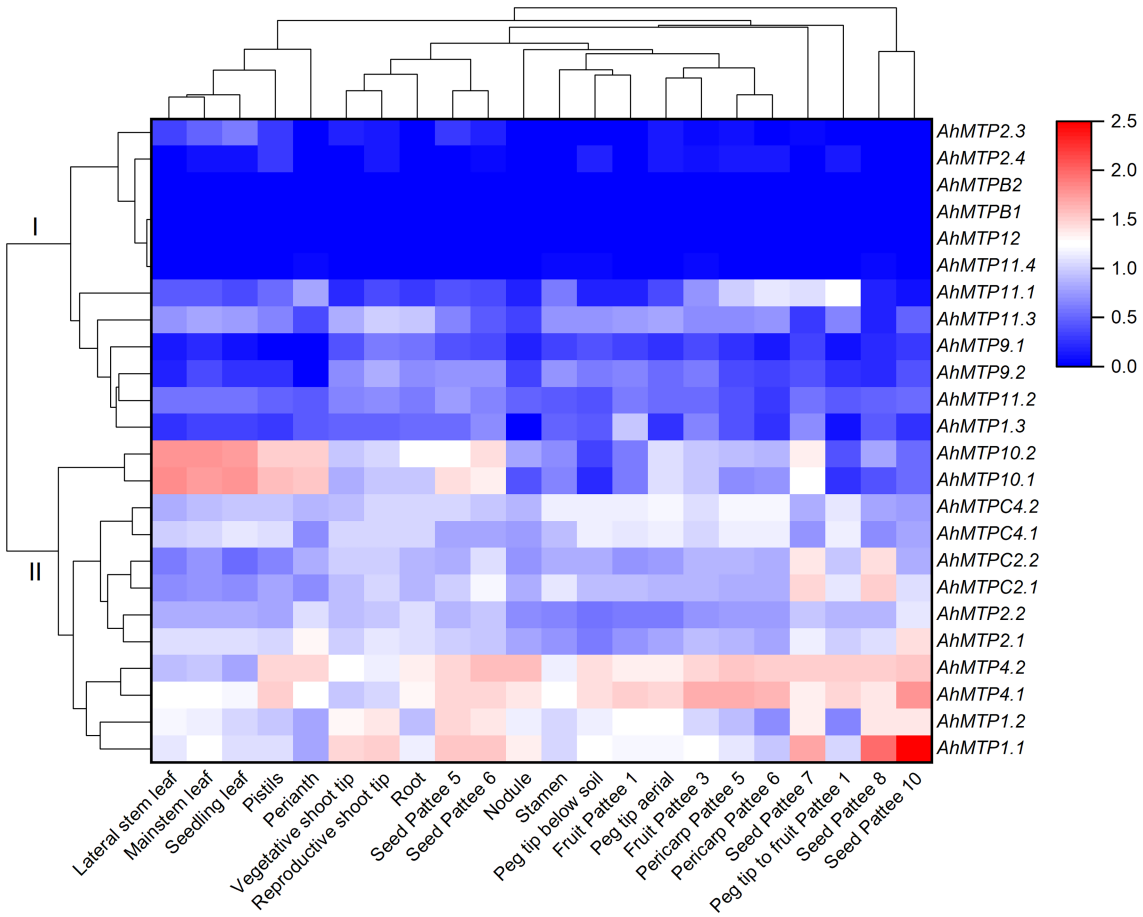


FIGURE 4 | The expression profiles of *AhMTP* genes in different tissues of peanut. Gene expression is expressed in $lg^{(FPKM+1)}$ and *Z* score normalized. Pattee 1, 3, 5, 6, 7, 8, and 10 represent different developmental stages of peanut pods according to Pattee' s classification (Clevenger et al., 2016).

binds Zn^{2+} or Cd^{2+} to the active site and transports the bound ions across the bacterial membrane using the proton motive force (Lu et al., 2009; Lopez-Redondo et al., 2018). The presence of ZT-dimer and specific motif complex in Mn-CDFs suggested that these proteins might serve as metal transporters by forming homodimers or heterodimers functional complexes (Liu et al., 2019). These structural features indicate that Mn-CDFs might be multiple metal transporter in peanut plants.

Some proteins of group 8 have been proven to be vacuolar-or Golgi-localized Mn transporters (Delhaize et al., 2003; Chen et al., 2013; Pedas et al., 2014). Meanwhile, several members of group 9, including AtMTP11 from *Arabidopsis* (Delhaize et al., 2007; Peiter et al., 2007), OsMTP11 from rice (Tsunemitsu et al., 2018), and BmMTP10 and BmMTP11 from *Beta vulgaris* spp. *maritima* (Erbasol et al., 2013), were also characterized as the Golgi- and/or pre-vacuolar-localized transporters that confer Mn tolerance *via* intra-cellular Mn compartmentalization.

A total of eight Zn-CDF members were identified in peanut, which were further divided into three groups: group 1 (AhMTP1.1/1.2/1.3 and AhMTPB1/B2), 5 (AhMTPC2.1/C2.2), and 12 (AhMTP12). Unlike the Mn-CDF subfamily, Zn-CDF members share characteristics as follows: (1) containing cation efflux domain but not ZT-dimer; (2) containing motifs 6 and 11, which were annotated as cation efflux TMD; (3) containing five to six TMDs except AhMTP1.3 and AhMTP12, which contained 2 and 16 TMDs respectively; (4) best modeled with 6xpd.1 (Cryo-EM structure of human ZnT8 double mutant—D110N and D224N, determined in outward-facing conformation); and (5) localizing to vacuolar membranes. ZnT8 is a Zn^{2+}/H^{+} antiporter that plays essential roles in regulating Zn^{2+} accumulation in the insulin secretory granules of pancreatic β cells (Xue et al., 2020). The structural features suggest that Zn-CDF members might serve as Zn transporters in peanut plants. Two Zn-CDF members from *Arabidopsis*, AtMTP1 and AtMTP3, belonging to group 1, have been showed to be Zn transporters that are localized to vacuolar membrane and mediate the vacuole sequestration of Zn (Kobae et al., 2004; Arrivault et al., 2006).

Interestingly, AhMTP12 contained 16 TMDs, which is higher than the homologous gene from other plant species, including BrrMTP12 from turnip (Li et al., 2018), CitMTP12 (12 TMDs) from sweet orange (Fu et al., 2017), PtrMTP12 (12 TMDs) from *P. trichocarpa* (Gao et al., 2020), VvMT12 (12 TMDs) from grape (Shirazi et al., 2019), and NtMTP12.1 (10 TMDs) and NtMTP12.2 (eight TMDs) from tobacco

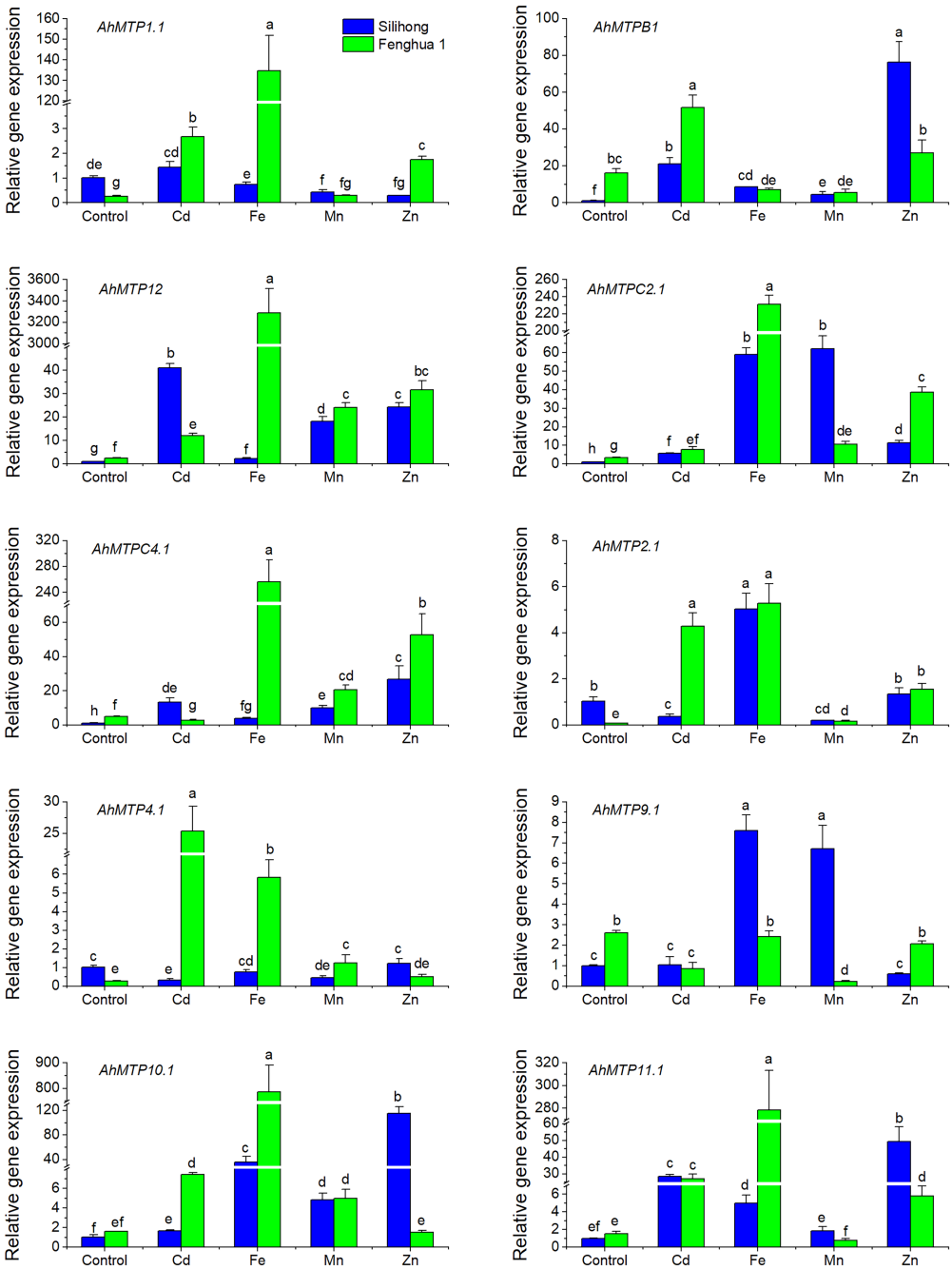


FIGURE 5 | Expression levels of 10 *AhMTP* genes in the root of two peanut cultivars exposed to excess metals. Data (means ± SE, $n = 3$) shared the same letter(s) above the error bars are not significantly different at the 0.05 level by Duncan multiple range test.

(Liu et al., 2019). It was also observed that AhMTP12 had the largest protein size (867 amino acids) and MW (97.04 kDa). The results were in accordance with the characteristics of MTP12 in other plants (Li et al., 2018; Liu et al., 2019; Gao et al., 2020). Moreover, collinearity analysis showed that *AhMTP12* is the only gene in the *AhMTP* family that has not experienced gene duplication events. The current results indicate that *AhMTP12* might be distinctive both in the evolutionary process and physiological functions in peanut plants. In *Arabidopsis*, AtMTP12 has been shown to forms a functional complex with AtMTP5t1, one of the splicing variants of AtMTP5, transporting Zn into the Golgi (Fujiwara et al., 2015).

Another distinctive gene of Zn-CDFs is *MTP1.3*, which has the shortest sequence of the gene and protein. Only two TMDs and two motifs (6 and 11) were harbored in the MTP1.3

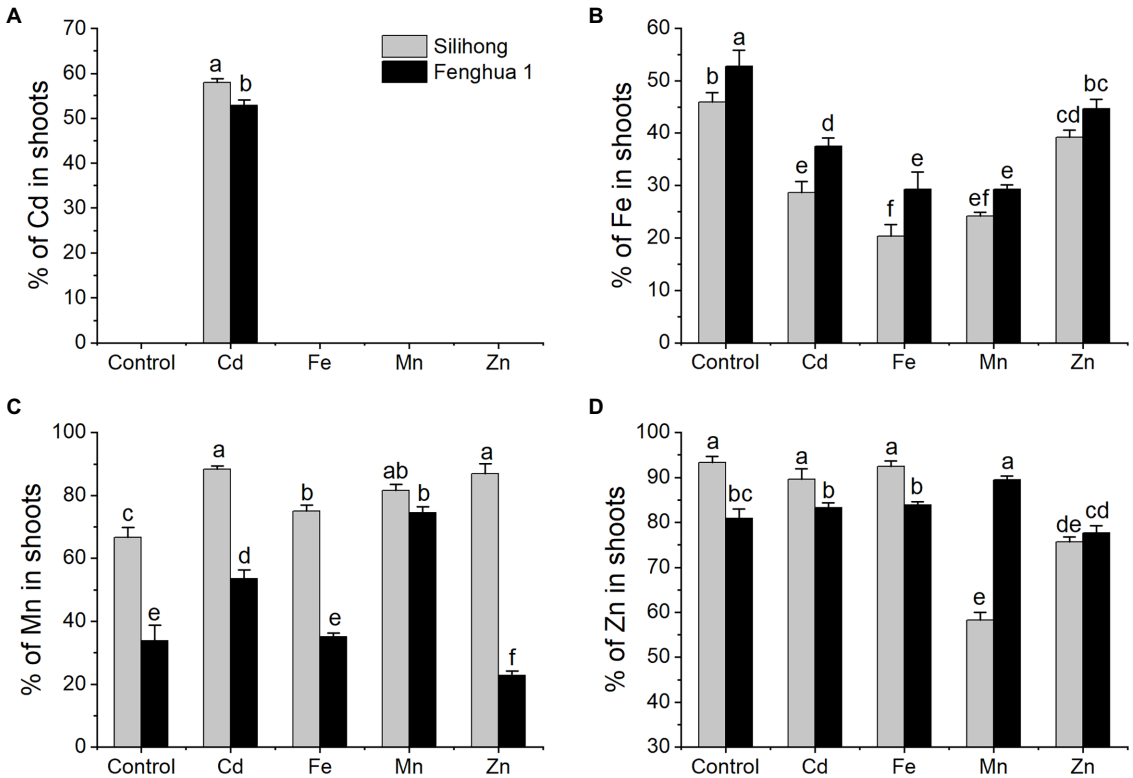


FIGURE 6 | Percentages of Cd **(A)**, Fe **(B)**, Mn **(C)**, and Zn **(D)** in the shoot of two peanut cultivars exposed to excess metals. Data (means ± SE, $n = 3$) shared the same letter(s) above the error bars are not significantly different at the 0.05 level by Duncan multiple range test.

TABLE 6 | Stepwise linear regression analysis of the percentage of metal in shoots and expression of *AhMTP* genes ($n = 30$).

Gene expression (−ΔΔCT)	% of Fe in shoots			% of Mn in shoots			% of Zn in shoots		
	β	*T*	*R*	*β*	*T*	*R*	*β*	*T*	*R*
AhMTP1.1	0.086	0.380	0.079	−0.445*	−2.627	−0.445	0.261	1.240	0.236
AhMTP2.1	−0.110	−0.678	−0.140	0.208	0.968	0.183	0.509**	2.978	0.497
AhMTP4.1	0.657***	4.168	0.648	0.098	0.487	0.093	−0.156	−0.755	−0.147
AhMTP9.1	0.810***	4.613	0.686	−0.145	−0.846	−0.161	−0.185	−1.049	−0.202
AhMTP10.1	−0.242	−1.515	−0.301	0.366	1.882	0.341	0.086	0.371	0.072
AhMTP11.1	−0.111	−0.632	−0.131	0.434	1.897	0.343	−0.096	−0.462	−0.090
AhMTP12	−0.976**	−3.196	−0.546	0.261	1.010	0.191	0.125	0.583	0.114
AhMTPB1	0.004	0.032	0.007	−0.118	−0.691	−0.132	−0.212	−1.366	−0.259
AhMTPC2.1	−1.751***	−7.252	−0.829	0.071	0.350	0.067	−0.579**	−3.385	−0.546
AhMTPC4.1	1.905***	5.590	0.752	−0.081	−0.398	−0.076	0.191	0.843	0.163

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

protein. Collinearity analysis revealed that *AhMTP1.3* and *AhMTP1.2* might be resulted from segmental duplication within the same subgenome.

The six Zn/Fe-CDF members were divided into two significantly different groups: Group 6 (AhMTP2.1/2.2/2.3/2.4) and 7 (AhMTPC4.1/C4.2). The two members of group 7 are WGD genes, which have the following common characteristics: (1) containing cation efflux domain but not ZT-dimer; (2) only containing two motifs (11 and 15), of which motif 11 were annotated as cation efflux TMD; and (3) localizing to vacuolar membranes. In contrast, the group 6 is composed of two pairs of WGD genes (*AhMTP2.1/2.2* and *AhMTP2.3/2.4*). The pair of *AhMTP2.1/2.2* is similar in all structural characteristics both at protein and gene levels. However, there are significant differences in the structure of the gene and protein between *AhMTP2.3* and *AhMTP2.4*. For instances, while AhMTP2.3 contained six TMDs, AhMTP2.4 only contained three TMDs. Unlike AhMTP2.1/2.2 that contained both cation efflux domain and ZT-dimer, AhMTP2.3 contained cation efflux domain but not ZT-dimer, whereas AhMTP2.4 contained ZT-dimer but not cation

efflux domain. A 3D-model prediction revealed that all Zn/Fe-CDFs were commendably modeled to 5vrf.1 except AhMTP2.4, which preferentially modeled with 3h90.1. The large divergence in the protein structure suggested that AhMTP2.3 and AhMTP2.4 might have distinctive physiological functions in peanut plants.

Gene duplication as a major source of novel genes contributes to the acquirement of novel functions (Panchy et al., 2016). Peanut is an allotetraploid species whose genome contains essentially complete sets of chromosomes (AABB) from two diploid ancestral species, *A. duranensis* (AA) and *A. ipaensis* (BB) (Bertioli et al., 2019). Expectedly, our results indicate that almost all *AhMTP* genes experienced gene duplication events except *AhMTP12*, resulting in 13 gene pairs (**Figure 2B**). Among them, 11 gene pairs were evolved from WGDs, and *AhMTP1.3* and *AhMTP1.2* might be segmentally duplicated genes. Gene pairs of duplication were consistent with the results obtained from phylogenetic analysis.

Duplicated genes, if they survive, would diverge in both the regulatory and coding regions during evolution, resulting in distinct functions (Xu et al., 2012). Thus, exon-intron structure can provide additional evidence to support phylogenetic analysis of gene families (Zhang et al., 2012). In the present study, we found that *AhMTP* genes within the same phylogenetic group show highly similar exon-intron structures, despite there are large variations among groups even in the same subfamily (**Figure 2A**). While most of the WGD gene pairs shared a similar exon-intron structure, divergences were also found within three gene pairs (*AhMTPC4.1/C4.2*, *AhMTP2.3/2.4*, and *AhMTP9.1/9.2*). Divergences of duplicate genes in exon-intron structure mainly resulted from three mechanisms: exon-intron gain/loss, exonization/pseudoexonization, and insertion/deletion (Xu et al., 2012). Although all the gene pairs experienced exon-intron gain/loss events, exonization/pseudoexonization also occurred in the evolutionary process of *AhMTP2.3/2.4* and *AhMTP9.1/9.2*, which might contribute to functionally distinct paralogues.

It is noteworthy that six Zn-CDF genes belonging to group 1 (*AhMTP1.1/1.2/1.3* and *AhMTPB1/B2*) and 12 (*AhMTP12*) have one exon without intron in CDS, which is called single exon genes (SEG) (Sakharkar et al., 2004). The presence of SEG in multi-cellular eukaryotic genomes intriguing since such kind of genes are archetypal of prokaryotes (Sakharkar et al., 2004). SEGs can be divided into two groups: (i) UTR intron-containing SEGs (uiSEGs) that have introns in their untranslated region (UTR) and (ii) intronless genes (IGs) lacking introns in the entire gene (Jorquera et al., 2018). Herein, *AhMTP1.3* having an intron in the UTR belong to uiSEGs while *AhMTP1.1/1.2*, *AhMTPB1/B2,* and *AhMTP12* are IGs. SEGs have been reported in MTP family from tobacco (Liu et al., 2019) and tomato (*Solanum lycopersicum*) (El-Sappah et al., 2021).

Gene expression was generally regulated at two different levels: transcriptional and post-transcriptional. CRE play essential roles in the transcriptional regulation of gene expression by interacting with specific transcription factors and RNA polymerase. In the current study, a large number of CAAT-box and TATA-box were detected in the promoters of *AhMTP* genes; this is expectedly because the two elements are involved in the regulation of expression frequency and initiation of transcription (Laloum et al., 2013). Besides, the wide distribution of light-responsive elements, phytohormonal responsive elements, and abiotic stress in the promoters indicate that expression of *AhMTP* genes could be regulated by many factors in multiple pathways. MYB-binding sites (MBS) were proposed to be involved in metal tolerance by driving *MTP1* expression in *Arabidopsis helleri* (Fasani et al., 2017). The location of MBS in the promoters of *AhMTP1.1/1.2*, *AhMTPC2.1*, *AhMTP12*, and *AhMTP4.1* suggests that these genes might be regulated at the MYB-binding sites.

MicroRNAs are generally believed to downregulate the expression of target genes by cleaving mRNA or inhibiting the translation of target genes (Bartel, 2009). Four pairs of *AhMTP* genes (*AhMTP9.1/9.2*, *AhMTPC2.1/C2.2*, *AhMTP4.1/4.2,* and *AhMTP11.2/11.3*) were predicted as targets of five miRNAs including *ahy-miR167-5p*, *ahy-miR156a*, *ahy-miR156c*, *ahy-miR3511-3p*, and *ahy-miR3519*. Although the function of *miR3511-3p* and *miR3519* is unknown, the regulation of *miR167* and *miR156* on their targets has been extensively studied. The *miR167* has been reported to target the mRNAs encoding the *ARF6*, *ARF8*, and *IAR3*, regulating auxin signaling and homeostasis in *Arabidopsis* (Wu et al., 2006; Kinoshita et al., 2012; Yao et al., 2019). Another study demonstrated that *BnNRAMP1b* is a target of *miR167* in *Brassica napus* (Meng et al., 2017). *MiR156a* and *miR156c* play dominant roles in regulating abiotic stress resistance through a *miR156-SPL* regulatory pathway (Cui et al., 2014; Wang et al., 2021). The expression of *miR167* and *miR156* was downregulated by Ca deficiency in peanut embryos (Chen et al., 2019). Similar results were reported in the roots and shoots of high-Fe rice line under Fe deficiency (Agarwal et al., 2015). Together, it can be inferred that *AhMTP9.1/9.2* and *AhMTPC2.1/C2.2* might be post-transcriptionally repressed by *ahy-miR167-5p* and *ahy-miR156*(a/c), respectively, which deserves to be experimentally tested.

Tissue expression profiles indicate that half of *AhMTP* genes showed low expression in the 22 tissues, among them, six genes (*AhMTP2.3/2.4*, *AhMTPB1/B2*, *AhMTP11.4,* and *AhMTP12*) were not or rarely expressed in any tissues. The reduced gene expression of these genes might be beneficial for maintenance of their biological functions and avoiding gene loss during evolution processes (Qian et al., 2010; Liu et al., 2019). With respect for the remaining 18 *AhMTP* genes, most of them preferentially expressed in reproductive tissues such as stamens, peg tips, fruit, peg tip to fruit, pericarp, and seed. These genes might play roles in the development of pods or seeds by mediating the transport of divalent metals. High expression levels of *AhMTP10.1/10.2* in leaves, perianth, and pistil indicate they might be involved in metal homeostasis in leaves and flower.

Due to most of AhMTP proteins were predicted to localize to vacuole membranes, we speculated that changing expressions of *AhMTP* genes in roots might alter the root-to-shoot translocation of divalent metals. To test this hypothesis, the expression of *AhMTP* genes in roots of two peanut cultivars in response to metal stresses and its relationship to metal translocation was investigated. *AhMTP* genes showed wide differences in response to various metal stresses depending on cultivars and metal exposure. Generally, 10 *AhMTP* genes tested in the current study were sensitive to excess metal exposure and most of them showed higher expressions under

metal stress (particularly in Fe, Cd, and Zn treatments). In agreement with previous studies (Liu et al., 2017; Yu et al., 2019), we found that Silihong showed higher capacity for the translocation of Cd from roots to shoots than Fenghua 1. Cd translocation was also reported to be decreased by increasing Cd concentration in rhizosphere (Yu et al., 2019). Most of *AhMTP* genes were induced by Cd in a cultivar-dependent manner, and the increase of gene expression was more pronounced in Silihong than in Fenghua 1. These results indicate that *AhMTP* genes, particularly *AhMTP1.1*, *AhMTPB1*, *AhMTPC2.1*, and *AhMTP10.1*, might be responsible for the cultivar difference of Cd translocation in peanuts.

Stepwise linear regression analysis revealed Fe translocation from roots to shoots might be associate to *AhMTP4.1*, *AhMTP9.1*, *AhMTPC4.1*, *AhMTPC2.1*, and *AhMTP12*. The upregulation of *AhMTPC2.1* and *AhMTP12* might contribute to the reduction of Fe translocation *via* sequestration of Fe in vacuoles or other vesicles under excess metal exposure. More sensitive response of *AhMTP4.1* and *AhMTPC4.1* to metal exposure in the roots might be responsible for the higher translocation of Fe in plants of Fenghua 1.

The significant and negative correlation of the expression of *AhMTP1.1* and the percentage of Mn in shoots indicates that *AhMTP1.1* might be involved in Mn translocation. The downregulation of *AhMTP1.1* in the root under Fe, Mn, and Zn stresses might contribute to increased Mn translocation in Silihong. Higher Mn translocation in Fenghua 1 under metal exposure might result from higher expression of *AhMTP1.1* in the root. The translocation of Zn was not affected by Cd and Fe exposure, while excess Zn reduced its translocation in Silihong, and excess Mn largely decreased Zn translocation in Silihong but increased that in Fenghua 1. Stepwise linear regression analysis revealed *AhMTP2.1* and *AhMTPC2.1* might be associate with Zn translocation in peanut plants. Thus, the reverse change of percentages of Zn in shoots between Silihong and Fenghua 1 under Mn exposure might also be in part a consequence of the reverse response of *AhMTP2.1* and *AhMTPC2.1* to Mn exposure in the root of the two cultivars.

CONCLUSION

A total of 24 *AhMTP* genes were identified in peanut, which were divided into seven groups belonging to three substrate-specific clusters (Zn-CDFs, Zn/Fe-CDFs, and Mn-CDFs). All *AhMTP* genes underwent whole genome or segmental gene duplication events except *AhMTP12*. Most *AhMTP* members within the same subfamily or group generally have similar gene and protein structural characteristics. However, some genes, such as *AhMTP1.3*, *AhMTP2.4*, and *AhMTP12*, showed wide divergences. Most of *AhMTP* genes preferentially expressed in reproductive tissues, suggesting that these genes might play roles in metal transport during the pod and seed development stages. Excess metal exposure induced expressions for most of *AhMTP* genes in peanut roots depending on cultivars. By contrast, *AhMTP* genes in the root of Fenghua 1 were more sensitive to excess Fe, Cd, and Zn exposure than that of Silihong. The differential responses of *AhMTP* genes to metal exposure might be, at least partially, responsible for the different metal translocation from roots to shoots between Fenghua 1 and Silihong. The findings provide clues to further characterize the functions of AhMTP proteins in the uptake and translocation of metal ions in peanut plants, which is great of importance for screening or breeding cultivars for peanut safe production in heavy metal contaminated soil.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: NCBI's GEO database (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE71357), accession number: GSE71357.

AUTHOR CONTRIBUTIONS

XW and CW carried out most of the experimental work with assistance from GS and ZZ. GS and ZZ were responsible for experimental design. GS and XW carried out data analyses and wrote and revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grants from the Natural Science Foundation of Anhui Province (grant number 2108085MC83), the Natural Science Foundation for Colleges and Universities of Anhui Province (grant numbers KJ2020ZD83 and KJ2019A0587), and the Innovation Team of Scientific Research Platform of Anhui Province (grant number KJ2015TD001).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.791200/full#supplementary-material

REFERENCES

Agarwal, S., Mangrauthia, S. K., and Sarla, N. (2015). Expression profiling of iron deficiency responsive microRNAs and gene targets in rice seedlings of *Madhukar* × *Swarna* recombinant inbred lines with contrasting levels of iron in seeds. *Plant Soil* 396, 137–150. doi: 10.1007/s11104-015-2561-y

Arrivault, S., Senger, T., and Krämer, U. (2006). The *Arabidopsis* metal tolerance protein AtMTP3 maintains metal homeostasis by mediating Zn exclusion from the shoot under Fe deficiency and Zn oversupply. *Plant J.* 46, 861–879. doi: 10.1111/j.1365-313X.2006.02746.x

Bailey, T. L., Williams, N., Misleh, C., and Li, W. W. (2006). MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res.* 34(Suppl_2), W369–W373. doi: 10.1093/nar/gkl198

Bartel, D. P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell* 136, 215–233. doi: 10.1016/j.cell.2009.01.002

Bertioli, D. J., Cannon, S. B., Froenicke, L., Huang, G., Farmer, A. D., Cannon, E. K., et al. (2016). The genome sequences of *Arachis duranensis* and *Arachis ipaensis*, the diploid ancestors of cultivated peanut. *Nat. Genet.* 48, 438–446. doi: 10.1038/ng.3517

Bertioli, D. J., Jenkins, J., Clevenger, J., Dudchenko, O., Gao, D., Seijo, G., et al. (2019). The genome sequence of segmental allotetraploid peanut *Arachis hypogaea*. *Nat. Genet.* 51, 877–884. doi: 10.1038/s41588-019-0405-z

Bhardwaj, E., Shukla, R., and Das, S. (2020). "Plant roots and mineral nutrition: an overview of molecular basis of uptake and regulation, and strategies to improve nutrient use efficiency (NUE)," in *Plant Stress Biology: Strategies and Trends*. eds. B. Giri and M. P. Sharma (Singapore: Springer Singapore), 131–184.

Cao, Q., Xv, C., Jiang, Q., Wang, L., and Shi, G. (2019). Comparative transcriptome analysis reveals key genes responsible for the homeostasis of iron and other divalent metals in peanut roots under iron deficiency. *Plant Soil* 445, 513–531. doi: 10.1007/s11104-019-04313-5

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 13, 1194–1202. doi: 10.1016/j.molp.2020.06.009

Chen, Z., Fujii, Y., Yamaji, N., Masuda, S., Takemoto, Y., Kamiya, T., et al. (2013). Mn tolerance in rice is mediated by MTP8.1, a member of the cation diffusion facilitator family. *J. Exp. Bot.* 64, 4375–4387. doi: 10.1093/jxb/ert243

Chen, H., Yang, Q., Chen, K., Zhao, S., Zhang, C., Pan, R., et al. (2019). Integrated microRNA and transcriptome profiling reveals a miRNA-mediated regulatory network of embryo abortion under calcium deficiency in peanut (*Arachis hypogaea* L.). *BMC Genomics* 20:392. doi: 10.1186/s12864-019-5770-6

Chou, K. C., and Shen, H. B. (2010). Plant-mPLoc: a top-down strategy to augment the power for predicting plant protein subcellular localization. *PLoS One* 5:e11335. doi: 10.1371/journal.pone.0011335

Clevenger, J., Chu, Y., Scheffler, B., and Ozias-Akins, P. (2016). A developmental transcriptome map for allotetraploid *Arachis hypogaea*. *Front. Plant Sci.* 7:1446. doi: 10.3389/fpls.2016.01446

Cui, L. G., Shan, J. X., Shi, M., Gao, J. P., and Lin, H. X. (2014). The *miR156-SPL9-DFR* pathway coordinates the relationship between development and abiotic stress tolerance in plants. *Plant J.* 80, 1108–1117. doi: 10.1111/tpj.12712

Dai, X., Zhuang, Z., and Zhao, P. X. (2018). psRNATarget: a plant small RNA target analysis server. *Nucleic Acids Res.* 46, W49–w54. doi: 10.1093/nar/gky316

Delhaize, E., Gruber, B. D., Pittman, J. K., White, R. G., Leung, H., Miao, Y., et al. (2007). A role for the AtMTP11 gene of *Arabidopsis* in manganese transport and tolerance. *Plant J.* 51, 198–210. doi: 10.1111/j.1365-313X.2007.03138.x

Delhaize, E., Kataoka, T., Hebb, D. M., White, R. G., and Ryan, P. R. (2003). Genes encoding proteins of the cation diffusion facilitator family that confer manganese tolerance. *Plant Cell* 15, 1131–1142. doi: 10.1105/tpc.009134

Duvaud, S., Gabella, C., Lisacek, F., Stockinger, H., Ioannidis, V., and Durinx, C. (2021). Expasy, the Swiss bioinformatics resource portal, as designed by its users. *Nucleic Acids Res.* 49, W216–W227. doi: 10.1093/nar/gkab225

El-Sappah, A. H., Elrys, A. S., Desoky, E.-S. M., Zhao, X., Bingwen, W., El-Sappah, H. H., et al. (2021). Comprehensive genome wide identification and expression analysis of MTP gene family in tomato (*Solanum lycopersicum*) under multiple heavy metal stress. *Saudi J. Biol. Sci.* 28, 6946–6956. doi: 10.1016/j.sjbs.2021.07.073

Erbasol, I., Bozdag, G. O., Koc, A., Pedas, P., and Karakaya, H. C. (2013). Characterization of two genes encoding metal tolerance proteins from *Beta vulgaris* subspecies *maritima* that confers manganese tolerance in yeast. *Biometals* 26, 795–804. doi: 10.1007/s10534-013-9658-7

Fasani, E., DalCorso, G., Varotto, C., Li, M., Visioli, G., Mattarozzi, M., et al. (2017). The *MTP1* promoters from *Arabidopsis halleri* reveal cis-regulating elements for the evolution of metal tolerance. *New Phytol.* 214, 1614–1630. doi: 10.1111/nph.14529

Fu, X. Z., Tong, Y. H., Zhou, X., Ling, L. L., Chun, C. P., Cao, L., et al. (2017). Genome-wide identification of sweet orange (*Citrus sinensis*) metal tolerance proteins and analysis of their expression patterns under zinc, manganese, copper, and cadmium toxicity. *Gene* 629, 1–8. doi: 10.1016/j.gene.2017.07.072

Fujiwara, T., Kawachi, M., Sato, Y., Mori, H., Kutsuna, N., Hasezawa, S., et al. (2015). A high molecular mass zinc transporter MTP12 forms a functional heteromeric complex with MTP5 in the Golgi in *Arabidopsis thaliana*. *FEBS J.* 282, 1965–1979. doi: 10.1111/febs.13252

Gao, Y., Yang, F., Liu, J., Xie, W., Zhang, L., Chen, Z., et al. (2020). Genome-wide identification of metal tolerance protein genes in *Populus trichocarpa* and their roles in response to various heavy metal stresses. *Int. J. Mol. Sci.* 21:1680. doi: 10.3390/ijms21051680

Gustin, J. L., Zanis, M. J., and Salt, D. E. (2011). Structure and evolution of the plant cation diffusion facilitator family of ion transporters. *BMC Evol. Biol.* 11:76. doi: 10.1186/1471-2148-11-76

Hu, B., Jin, J., Guo, A. Y., Zhang, H., Luo, J., and Gao, G. (2015). GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics* 31, 1296–1297. doi: 10.1093/bioinformatics/btu817

Hurst, L. D. (2002). The Ka/Ks ratio: diagnosing the form of sequence evolution. *Trends Genet.* 18, 486–487. doi: 10.1016/S0168-9525(02)02722-1

Jorquera, R., González, C., Clausen, P., Petersen, B., and Holmes, D. S. (2018). Improved ontology for eukaryotic single-exon coding sequences in biological databases. *Database* 2018:bay089. doi: 10.1093/database/bay089

Kinoshita, N., Wang, H., Kasahara, H., Liu, J., Macpherson, C., Machida, Y., et al. (2012). IAA-Ala Resistant3, an evolutionarily conserved target of *miR167*, mediates *Arabidopsis* root architecture changes during high osmotic stress. *Plant Cell* 24, 3590–3602. doi: 10.1105/tpc.112.097006

Kobae, Y., Uemura, T., Sato, M. H., Ohnishi, M., Mimura, T., Nakagawa, T., et al. (2004). Zinc transporter of *Arabidopsis thaliana* AtMTP1 is localized to vacuolar membranes and implicated in zinc homeostasis. *Plant Cell Physiol.* 45, 1749–1758. doi: 10.1093/pcp/pci015

Laloum, T., De Mita, S., Gamas, P., Baudin, M., and Niebel, A. (2013). CCAAT-box binding transcription factors in plants: Y so many? *Trends Plant Sci.* 18, 157–166. doi: 10.1016/j.tplants.2012.07.004

Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., et al. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* 30, 325–327. doi: 10.1093/nar/30.1.325

Li, X., Wu, Y., Li, B., He, W., Yang, Y., and Yang, Y. (2018). Genome-wide identification and expression analysis of the cation diffusion facilitator gene family in turnip under diverse metal ion stresses. *Front. Genet.* 9:103. doi: 10.3389/fgene.2018.00103

Liu, J., Gao, Y., Tang, Y., Wang, D., Chen, X., Yao, Y., et al. (2019). Genome-wide identification, comprehensive gene feature, evolution, and expression analysis of plant metal tolerance proteins in tobacco under heavy metal toxicity. *Front. Genet.* 10:345. doi: 10.3389/fgene.2019.00345

Liu, C., Yu, R., and Shi, G. (2017). Effects of drought on the accumulation and redistribution of cadmium in peanuts at different developmental stages. *Arch. Agron. Soil Sci.* 63, 1049–1057. doi: 10.1080/03650340.2016.1271120

Lopez-Redondo, M. L., Coudray, N., Zhang, Z., Alexopoulos, J., and Stokes, D. L. (2018). Structural basis for the alternating access mechanism of the cation diffusion facilitator YiiP. *Proc. Natl. Acad. Sci. U. S. A.* 115, 3042–3047. doi: 10.1073/pnas.1715051115

Lu, M., Chai, J., and Fu, D. (2009). Structural basis for autoregulation of the zinc transporter YiiP. *Nat. Struct. Mol. Biol.* 16, 1063–1067. doi: 10.1038/nsmb.1662

Ma, G., Li, J., Li, J., Li, Y., Gu, D., Chen, C., et al. (2018). OsMTP11, a trans-Golgi network localized transporter, is involved in manganese tolerance in rice. *Plant Sci.* 274, 59–69. doi: 10.1016/j.plantsci.2018.05.011

Meng, J. G., Zhang, X. D., Tan, S. K., Zhao, K. X., and Yang, Z. M. (2017). Genome-wide identification of Cd-responsive NRAMP transporter genes and analyzing expression of *NRAMP 1* mediated by *miR167* in *Brassica napus*. *Biometals* 30, 917–931. doi: 10.1007/s10534-017-0057-3

Menguer, P. K., Farthing, E., Peaston, K. A., Ricachenevsky, F. K., Fett, J. P., and Williams, L. E. (2013). Functional analysis of the rice vacuolar zinc transporter OsMTP1. *J. Exp. Bot.* 64, 2871–2883. doi: 10.1093/jxb/ert136

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar Gustavo, A., Sonnhammer, E. L. L., et al. (2020). Pfam: the protein families database in 2021. *Nucleic Acids Res.* 49, D412–D419. doi: 10.1093/nar/gkaa913

Montanini, B., Blaudez, D., Jeandroz, S., Sanders, D., and Chalot, M. (2007). Phylogenetic and functional analysis of the cation diffusion facilitator (CDF) family: improved signature and prediction of substrate specificity. *BMC Genomics* 8:107. doi: 10.1186/1471-2164-8-107

Panchy, N., Lehti-Shiu, M., and Shiu, S. H. (2016). Evolution of gene duplication in plants. *Plant Physiol.* 171, 2294–2316. doi: 10.1104/pp.16.00523

Pedas, P., Schiller Stokholm, M., Hegelund, J. N., Ladegård, A. H., Schjoerring, J. K., and Husted, S. (2014). Golgi localized barley MTP8 proteins facilitate Mn transport. *PLoS One* 9:e113759. doi: 10.1371/journal.pone.0113759

Peiter, E., Montanini, B., Gobert, A., Pedas, P., Husted, S., Maathuis, F. J. M., et al. (2007). A secretory pathway-localized cation diffusion facilitator confers plant manganese tolerance. *Proc. Natl. Acad. Sci. U. S. A.* 104, 8532–8537. doi: 10.1073/pnas.0609507104

Qian, W., Liao, B. Y., Chang, A. Y. F., and Zhang, J. (2010). Maintenance of duplicate genes and their functional redundancy by reduced expression. *Trends Genet.* 26, 425–430. doi: 10.1016/j.tig.2010.07.002

Sakharkar, M. K., Chow, V. T. K., Chaturvedi, I., Mathura, V. S., Shapshak, P., and Kangueane, P. (2004). A report on single exon genes (SEG) in eukaryotes. *Front. Biosci.* 9, 3262–3267. doi: 10.2741/1478

Shirazi, Z., Abedi, A., Kordrostami, M., Burritt, D. J., and Hossain, M. A. (2019). Genome-wide identification and characterization of the metal tolerance protein (MTP) family in grape (*Vitis vinifera* L.). *3 Biotech.* 9:199. doi: 10.1007/s13205-019-1728-2

Singh, S., Parihar, P., Singh, R., Singh, V. P., and Prasad, S. M. (2015). Heavy metal tolerance in plants: role of transcriptomics, proteomics, metabolomics, and ionomics. *Front. Plant Sci.* 6:1143. doi: 10.3389/fpls.2015.01143

Song, H., Wang, P., Lin, J. Y., Zhao, C., Bi, Y., and Wang, X. (2016). Genome-wide identification and characterization of *WRKY* gene family in peanut. *Front. Plant Sci.* 7:534. doi: 10.3389/fpls.2016.00534

Su, Y., Liu, J., Lu, Z., Wang, X., Zhang, Z., and Shi, G. (2014). Effects of iron deficiency on subcellular distribution and chemical forms of cadmium in peanut roots in relation to its translocation. *Environ. Exp. Bot.* 97, 40–48. doi: 10.1016/j.envexpbot.2013.10.001

Su, Y., Wang, X., Liu, C., and Shi, G. (2013). Variation in cadmium accumulation and translocation among peanut cultivars as affected by iron deficiency. *Plant Soil* 363, 201–213. doi: 10.1007/s11104-012-1310-8

Takemoto, Y., Tsunemitsu, Y., Fujii-Kashino, M., Mitani-Ueno, N., Yamaji, N., Ma, J. F., et al. (2017). The tonoplast-localized transporter MTP8.2 contributes to manganese detoxification in the shoots and roots of *Oryza sativa* L. *Plant Cell Physiol.* 58, 1573–1582. doi: 10.1093/pcp/pcx082

Traore, S. M., Han, S., Binagwa, P., Xu, W., Chen, X., Liu, F., et al. (2021). Genome-wide identification of mlo genes in the cultivated peanut (*Arachis hypogaea* L.). *Euphytica* 217:61. doi: 10.1007/s10681-021-02792-1

Tsirigos, K. D., Peters, C., Shu, N., Käll, L., and Elofsson, A. (2015). The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. *Nucleic Acids Res.* 43, W401–W407. doi: 10.1093/nar/gkv485

Tsunemitsu, Y., Genga, M., Okada, T., Yamaji, N., Ma, J. F., Miyazaki, A., et al. (2018). A member of cation diffusion facilitator family, MTP11, is required for manganese tolerance and high fertility in rice. *Planta* 248, 231–241. doi: 10.1007/s00425-018-2890-1

Ueno, D., Sasaki, A., Yamaji, N., Miyaji, T., Fujii, Y., Takemoto, Y., et al. (2015). A polarly localized transporter for efficient manganese uptake in rice. *Nat. Plants* 1:15170. doi: 10.1038/nplants.2015.170

van der Zaal, B. J., Neuteboom, L. W., Pinas, J. E., Chardonnens, A. N., Schat, H., Verkleij, J. A., et al. (1999). Overexpression of a novel *Arabidopsis* gene related to putative zinc-transporter genes from animals can lead to enhanced zinc resistance and accumulation. *Plant Physiol.* 119, 1047–1056. doi: 10.1104/pp.119.3.1047

Vatansever, R., Filiz, E., and Eroglu, S. (2017). Genome-wide exploration of metal tolerance protein (*MTP*) genes in common wheat (*Triticum aestivum*): insights into metal homeostasis and biofortification. *Biometals* 30, 217–235. doi: 10.1007/s10534-017-9997-x

Wan, L., Ren, W., Miao, H., Zhang, J., and Fang, J. (2020). Genome-wide identification, expression, and association analysis of the monosaccharide transporter (MST) gene family in peanut (*Arachis hypogaea* L.). *3 Biotech.* 10:130. doi: 10.1007/s13205-020-2123-8

Wang, K., Liu, Y., Teng, F., Cen, H., Yan, J., Lin, S., et al. (2021). Heterogeneous expression of *Osa-MIR156bc* increases abiotic stress resistance and forage quality of alfalfa. *Crop J.* 9, 1135–1144. doi: 10.1016/j.cj.2020.11.009

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46, W296–w303. doi: 10.1093/nar/gky427

Wu, M. F., Tian, Q., and Reed, J. W. (2006). *Arabidopsis microRNA167* controls patterns of *ARF6* and *ARF8* expression, and regulates both female and male reproduction. *Development* 133, 4211–4218. doi: 10.1242/dev.02602

Xu, G., Guo, C., Shan, H., and Kong, H. (2012). Divergence of duplicate genes in exon–intron structure. *Proc. Natl. Acad. Sci. U. S. A.* 109, 1187–1192. doi: 10.1073/pnas.1109047109

Xue, J., Xie, T., Zeng, W., Jiang, Y., and Bai, X. C. (2020). Cryo-EM structures of human ZnT8 in both outward- and inward-facing conformations. *elife* 9:e58823. doi: 10.7554/eLife.58823

Yao, X., Chen, J., Zhou, J., Yu, H., Ge, C., Zhang, M., et al. (2019). An essential role for *miRNA167* in maternal control of embryonic and seed development. *Plant Physiol.* 180, 453–464. doi: 10.1104/pp.19.00127

Yu, R., Jiang, Q., Xv, C., Li, L., Bu, S., and Shi, G. (2019). Comparative proteomics analysis of peanut roots reveals differential mechanisms of cadmium detoxification and translocation between two cultivars differing in cadmium accumulation. *BMC Plant Biol.* 19:137. doi: 10.1186/s12870-019-1739-5

Yuan, L., Yang, S., Liu, B., Zhang, M., and Wu, K. (2012). Molecular characterization of a rice metal tolerance protein, OsMTP1. *Plant Cell Rep.* 31, 67–79. doi: 10.1007/s00299-011-1140-9

Zhang, Y., Gao, M., Singer, S. D., Fei, Z., Wang, H., and Wang, X. (2012). Genome-wide identification and analysis of the *TIFY* gene family in grape. *PLoS One* 7:e44465. doi: 10.1371/journal.pone.0044465

Zhao, K., Li, K., Ning, L., He, J., Ma, X., Li, Z., et al. (2019). Genome-wide analysis of the growth-regulating factor family in peanut (*Arachis hypogaea* L.). *Int. J. Mol. Sci.* 20:4120. doi: 10.3390/ijms20174120

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 01 April 2022
doi: 10.3389/fpls.2022.847853

MicroRNA and Degradome Profiling Uncover Defense Response of *Fraxinus velutina* Torr. to Salt Stress

Jian Ning Liu[1†], Xinmei Ma[1†], Liping Yan[2†], Qiang Liang[1,3,4], Hongcheng Fang[1,3,4], Changxi Wang[1], Yuhui Dong[1,3,4], Zejia Chai[1], Rui Zhou[1], Yan Bao[1], Lichang Wang[1], Shasha Gai[1], Xinya Lang[1], Ke Qiang Yang[1,3,4], Rong Chen[5]* and Dejun Wu[2]**

[1] College of Forestry, Shandong Agricultural University, Tai'an, China, [2] Shandong Provincial Academy of Forestry, Jinan, China, [3] Shandong Taishan Forest Ecosystem Research Station, Shandong Agricultural University, Tai'an, China, [4] State Forestry and Grassland Administration Key Laboratory of Silviculture in Downstream Areas of the Yellow River, Shandong Agricultural University, Tai'an, China, [5] Culaishan Forest Farm, Tai'an, China

OPEN ACCESS

Edited by:
Rohini Garg,
Shiv Nadar University, India

Reviewed by:
Charu Lata,
National Institute of Science Communication and Information Resources (CSIR), India
Youxiong Que,
Fujian Agriculture and Forestry University, China

***Correspondence:**
Ke Qiang Yang
yangwere@126.com
Dejun Wu
sdlky412x@163.com
Rong Chen
cr13854800805@163.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

Received: 03 January 2022
Accepted: 23 February 2022
Published: 01 April 2022

Citation:
Liu JN, Ma X, Yan L, Liang Q, Fang H, Wang C, Dong Y, Chai Z, Zhou R, Bao Y, Wang L, Gai S, Lang X, Yang KQ, Chen R and Wu D (2022) MicroRNA and Degradome Profiling Uncover Defense Response of Fraxinus velutina Torr. to Salt Stress. Front. Plant Sci. 13:847853. doi: 10.3389/fpls.2022.847853

Soil salinization is a major environmental problem that seriously threatens the sustainable development of regional ecosystems and local economies. *Fraxinus velutina* Torr. is an excellent salt-tolerant tree species, which is widely planted in the saline-alkaline soils in China. A growing body of evidence shows that microRNAs (miRNAs) play important roles in the defense response of plants to salt stress; however, how miRNAs in *F. velutina* exert anti-salt stress remains unclear. We previously identified two contrasting *F. velutina* cuttings clones, salt-tolerant (R7) and salt-sensitive (S4) and found that R7 exhibits higher salt tolerance than S4. To identify salt-responsive miRNAs and their target genes, the leaves and roots of R7 and S4 exposed to salt stress were subjected to miRNA and degradome sequencing analysis. The results showed that compared with S4, R7 showed 89 and 138 differentially expressed miRNAs in leaves and roots, respectively. Specifically, in R7 leaves, miR164d, miR171b/c, miR396a, and miR160g targeting *NAC1*, *SCL22*, *GRF1*, and *ARF18*, respectively, were involved in salt tolerance. In R7 roots, miR396a, miR156a/b, miR8175, miR319a/d, and miR393a targeting *TGA2.3*, *SBP14*, *GR-RBP*, *TCP2/4*, and *TIR1*, respectively, participated in salt stress responses. Taken together, the findings presented here revealed the key regulatory network of miRNAs in R7 responding to salt stress, thereby providing new insights into improving salt tolerance of *F. velutina* through miRNA manipulation.

Keywords: ***Fraxinus velutina*** **Torr., salt stress, microRNA, degradome, defense response**

INTRODUCTION

Soil salinization is a major environmental problem. It is estimated that by the year 2050, more than half of global arable land will be saline contamination (Butcher et al., 2016). Salinized soils hinder the growth and development of plants, resulting in the loss of biomass production, and even the deterioration of regional ecosystem (Polle and Chen, 2015; Hossain and Dietz, 2016; Ondrasek and Rengel, 2021). What's more serious is that increasing soil salinization is now threatening sustainable development of local economies (Chen and Mueller, 2018). How to protect and restore the fragile ecosystem in salinized areas has become an urgent global issue. The selection and cultivation of

naturally salt-tolerant plants are currently considered as an economically feasible strategy for the problem (Litalien and Zeeb, 2020). *Fraxinus velutina* Torr. is an excellent salt-tolerant tree species, which is widely planted in the saline-alkaline soils in Yellow River Delta, China (Mao et al., 2017). However, the mechanisms underlying salt tolerance of *F. velutina* remain largely unclear.

MicroRNAs (miRNAs) are a class of endogenous small non-coding RNAs with 18–25 nucleotides (nt), playing key regulatory roles in the defense response of plants to salt stress by regulating their target genes at a post-transcriptional level (Kumar et al., 2018; Song S.et al., 2019; Xu et al., 2021). The overexpression of miR528 in rice (*Oryza sativa*) can increase the contents of ascorbic and abscisic acid, and decrease reactive oxygen species (ROS) accumulation, thereby enhancing rice salt tolerance (Wang et al., 2021). Constitutive expression of rice miR528 and miR396 in creeping bentgrass modulates the growth and development, and enhances the response to salinity stress (Yuan et al., 2015, 2019). In wheat (*Triticum aestivum*), miRNA408 has been found to act as a crucial mediator in the tolerance to Pi deprivation and salt stress through modulating multiple stresses related to physiological processes (Bai et al., 2018). In maize (*Zea mays*), it is evidenced that *miR169/NF-YA* module is a key regulatory mediator in the response to salt stress in leaves and roots (Luan et al., 2014, 2015). In *Arabidopsis*, it is confirmed that miRNA393-associated regulatory modules can enhance the salt stress resistance by mediating several biological processes, including auxin signaling, redox-related components, osmoregulation and increased Na^+ exclusion (Iglesias et al., 2014; Chen Z.et al., 2015; Denver and Ullah, 2019). The overexpressed miR414c in cotton (*Gossypium hirsutum*) can negatively regulate iron superoxide dismutase gene, thereby enhancing plant tolerance to salinity stress (Wang et al., 2019). A *miR156/SPL* regulatory module increases tolerance to salinity stress via up-regulating *MdWRKY100* in apple (*Malus domestica*) (Ma et al., 2021). However, little information is available about the regulatory role of miRNAs in *F. velutina* responding to salt stress.

As important regulatory molecules, plant miRNAs bind to their target mRNAs and regulate gene expression by direct cleavage and degradation of their targets (Djami-Tchatchou et al., 2017). To date, multiple approaches to study miRNAs, their target genes, and regulatory networks have been established (Sun et al., 2014). Of these approaches, the most commonly used is direct miRNA cloning and/or deep sequencing. Various experimental approaches like RLM-RACE and degradome sequencing have been developed to identify miRNA targets in plants (Shamimuzzaman and Vodkin, 2012; Wang C. et al., 2013). Degradome sequencing is a powerful approach that integrates a modified RLM-RACE and deep sequencing, which has been used to confirm miRNA targets globally in plants. Recently, the combination of miRNA and degradome sequencing has been widely used to identify the salt-responsive miRNAs and their target genes involved in regulating plant tolerance to salt stress (Yu et al., 2016; Kumar et al., 2017; Cervera-Seco et al., 2019; Zhang Y.et al., 2020; Xu et al., 2021).

In our previous study, we identified two contrasting cutting clones R7 (salt-tolerant) and S4 (salt-sensitive) of *F. velutina* in which R7 exhibits higher salt tolerance than S4. Meanwhile, we performed a comparative transcriptome analysis between R7 and S4, and identified some key genes and signaling pathways underlying high salt tolerance of R7 (Ma et al., 2022). In this study, to identify the salt-responsive miRNAs and their target genes involved in salt stress tolerance, an integration of miRNA and degradome sequencing analysis was performed on the leaves and roots of R7 and S4 with or without salt treatment. Our work revealed the key regulatory network of miRNAs in R7 responding to salt stress, thereby providing new insights into improving salt tolerance of *F. velutina* through miRNA manipulation.

MATERIALS AND METHODS

Plant Materials and Culture Conditions

The salt-tolerant R7 and salt-sensitive S4 cutting clones of *F. velutina* were collected from the Experimental Base of Afforestation on Saline-Alkali Soil of Shandong Provincial Academy of Forestry, Shouguang city, China (118°42′9.18″ E, 37°9′38.94″ N). As previously described (Ma et al., 2022), 1-year-old cutting clones of both R7 and S4 were used in this study. After gently being removed rhizosphere soils, the clones were first pre-cultured in distilled water without any nutrient for 2 weeks, and then were transferred to a plastic container with 6 L of half-strength Hoagland's solution for 4 weeks; the solution was refreshed every 7 days. All the experimental cutting clones were cultured in a growth incubator (LICHEN, Shanghai, China) with 25/20°C (day/night temperature), and 16 h light (1,200 μmol m^{-2} s^{-1})/8 h dark.

Salt Treatment and Sampling

After 6 weeks of acclimatization, the healthy clones with uniform size were selected and treated with half-strength Hoagland's solution containing 250 mM NaCl for 12 h; the seedlings without NaCl treatment were considered as the control. After treatment, the leaves and roots of R7 and S4 were collected and stored at −80°C for further analysis. The symbols of leaves and roots samples were letters 'L' and 'R,' respectively. For instance, the samples R7SL and R7CL represented the leaves of R7 cutting clones with or without salt treatment, respectively.

Small RNA and Degradome Library Construction and Sequencing

Total of 24 samples (roots and leaves of R7 and S4, and two treatments with three biological replicates) were used small RNA and degradome library construction and sequencing. Total RNAs including small and large-size of tested samples were extracted using E.Z.N.A. Micro RNA Kit (Omega Bio-Tek, Norcross, GA, United States) following the manufacturer's instructions. The quality and quantity of the isolated RNA were evaluated using a Bioanalyzer 2100 instrument (Agilent, Santa Clara, CA, United States) and an RNA 6000 Nano LabChip Kit (Agilent, Santa Clara, CA, United States), ensuring that the RNA integrity number value was higher than 8.0. For small RNA sequencing, the library of each sample was constructed using

TruSeq Small RNA Library Preparation Kit (Illumina, San Diego, CA, United States), according to the manufacturer's protocol. For degradome sequencing, two libraries from equally pooled samples of S7 and S4 were, respectively, constructed as previously described (German et al., 2008, 2009). Both small RNA and degradome libraries were sequenced on an Illumina Hiseq 4000 instrument (LC-Bio, Hangzhou, China) at single-end (50 bp).

Identification of miRNAs

The raw reads were first trimmed using Trimmomatic v.0.39 (Bolger et al., 2014) to remove the junk reads, adapters, and low-complexity sequences. The validated reads were subjected to remove 3′ adaptor (5′-TGGAATTCTCGGGTGCCAAGG-3′) and perform length filter using Cutadapt v.3.5 (Kechin et al., 2017). The sequences with 18–25 nt were then subjected to exclude mRNA and other non-coding RNAs (tRNAs, rRNAs, snoRNAs, and snRNAs) via aligning the sequences, respectively to *F. velutina* mRNA sequences[1](Kelly et al., 2020) and Rfam v.14.6 database[2] using the BLAST tool (BLASTN) (Altschul et al., 1990). The remaining unique sequences were subsequently aligned to the available plant miRNAs sequences stored in miRBase v.22.1[3] to identify the conserved miRNAs in *F. velutina* using bowtie v1.3.0 (Langmead et al., 2009) with a maximum allowed mismatch ≤ 2.

The mature miRNA sequences obtained were mapped to the *F. velutina* reference genome sequence (see Text Footnote 1) by bowtie software, and then the corresponding flanking sequences of these miRNAs were extracted and used to predict the hairpin RNA structures using UNAfold v3.8 (Markham and Zuker, 2008). The potential miRNA precursors must meet the criteria as following: (1) the miRNAs were located on the arms (3′ or 5′) of stem-loop hairpin structure; (2) the miRNAs cannot contain large loops or breaks; (3) a maximum mismatches ≤ 6 were allowed between the miRNAs and their opposite sequences; (4) the predicted miRNA precursor structures must have minimal free folding energy index (>0.8) and negative minimal folding free energy, to differentiate them from other RNAs (Zhang et al., 2006; Yin et al., 2008).

The remaining small RNA sequences were used to identify novel miRNAs. After mapped to the *F. velutina* reference genome sequence, the flanking sequences of small RNAs were retrieved to predict secondary structures as the above mentioned. Only the small RNAs exhibiting a perfect stem-loop structure and meeting the criteria for plant miRNAs were identified as candidate novel miRNAs (Meyers et al., 2008).

Degradome Sequencing Data Analysis and miRNA Targets Identification

After trimmed by Trimmomatic software, the validated degradome sequencing reads were subjected to predict the putative miRNA cleaved targets using the CleaveLand pipeline v.3.0 (Addo-Quaye et al., 2009). In brief, the sequencing reads were mapped to the *F. velutina* mRNA sequences downloaded from Ash Tree Genomes Database (see Text Footnote 1), and the perfect matching alignments were retained. The resulting tags with a 35–36 nt extended signature by adding 15 nt of upstream sequence were subsequently aligned to the identified mature miRNAs in this study, with a maximum allowed mismatch ≤ 5. Alignments where the degradome tag position coincided with the 10th or 11th nt of a given miRNA were kept and scored (Allen et al., 2005). According to the previously described, targets were classified into I, II, or III (Addo-Quaye et al., 2008). In addition, t-plots showing the distribution of signatures of miRNA cleaved targets were built using R package.

Identification of Salt-Responsive miRNAs and Their Targets

The expression levels of the miRNAs were determined and normalized by transcripts per kilobase million. The differentially expressed miRNAs (DEmiRs) between each group were identified using R package DEseq2 (Love et al., 2014), with $|\log_2$ fold change$| \geq 1$ and P-value ≤ 0.05.

We previously performed a comparative transcriptome analysis between R7 and S4, and identified some key genes and signaling pathways underlying high salt tolerance of R7 (Ma et al., 2022). To narrow the targets of salt-responsive miRNAs, our previous transcriptome data were used. For each comparison, the miRNA-target pairs containing DEmiRs and differentially expressed genes (DEGs) were selected for further analysis, and for each selected pair, the expression patterns of DEmiR were opposite to that of DEGs.

Validation of the Identified miRNAs by Quantitative Real-Time PCR

To verify the results of small RNA sequencing, 15 DEmiRs including 12 conserved (ptc-miR160a, ptc-miR160g, mtr-miR164d, stu-MIR167d-p3_2ss6TC19CT, bna-MIR169c-p5_2ss12GC17TG, gma-miR169j-5p, mtr-MIR2592bj-p3_2ss12TC19AT, mtr-MIR2592bj-p3_2ss12TC19AT, mtr-miR393a_L+1, hbr-miR396a_R-1_2ss19TC20CT, hbr-miR396a_R-1_2ss19TC20CT, and gma-miR403a_R-1) and 3 novels (PC-3p-46517_176, PC-3p-56802_134, and PC-5p-55954_137) were randomly selected to perform quantitative real-time PCR (qRT-PCR) analysis (**Supplementary Table 1**). Total RNA isolated from leaves and roots using the E.Z.N.A. Micro RNA Kit (Omega Bio-Tek, Norcross, GA, United States) according to the manufacturer's protocol. The primers were designed using Primer 5.0 software and synthesized by Sangon Biotech Co., Ltd. (Shanhai, China). The reverse transcription of miRNA was performed using Mir-X miRNA First-Strand Synthesis and TB Green qRT-PCR User Manual (Takara, Dalian, China) according to the manufacturer's protocol.

The qRT-PCR was performed with the TB Green Premix Ex Taq II kit (Takara, Dalian, China) on a CFX Connect Real-Time instrument (Bio-Rad, Hercules, CA, United States). Each sample had three independent replicates. The 5.8s rRNA was used as an internal reference gene. The relative expression level

[1] http://ashgenome.org/

[2] http://rfam.xfam.org/

[3] http://www.mirbase.org/

of each miRNA was calculated based on the $2^{-\Delta\ \Delta\ Ct}$ method (Livak and Schmittgen, 2001).

Statistical Analyses

Statistical data were presented as mean ± standard deviation (SD). Student's *t*-test was used to compare the differences between two groups. A *P*-value ≤ 0.05 was considered to be significant difference. GraphPad Prism v.9.0 (GraphPad Software Inc., La Jolla, CA, United States) was used to perform statistical analysis.

RESULTS

Identification of Conserved and Novel miRNAs in *F. velutina*

To identify miRNAs in *F. velutina* responding to salt stress, 24 small RNA sequencing libraries from leaf and root samples of R7 and S4 under salt treatment and control conditions were constructed and sequenced. In total, an average of 11.92 million raw reads for each sample were obtained. After removing low quality reads and adaptor sequences, approximately 2.13 million unique validated reads were obtained for each sample (**Supplementary Table 2**). After further excluding mRNA, tRNAs, rRNAs, snoRNAs, and snRNAs, an average of 2.12 million unique reads with length of 18–25 nt were generated for each sample (**Supplementary Table 3**).

To identify the conserved and novel miRNAs in *F. velutina*, the filtered unique reads were aligned to the miRBase v.22.1 database and *F. velutina* reference genome. A total of 987 miRNAs including 560 conserved and 427 novel miRNAs were identified. The majority of miRNAs were distributed between 20 and 24 nt, with 24 nt exhibiting the highest abundance (**Figure 1A**). On the basis of sequence similarity, these miRNAs were further classified into 50 miRNA families, with MIR156, MIR159, MIR166, and MIR396 presenting relatively high abundance (**Supplementary Figure 1A** and **Supplementary Table 4**).

Identification of Specifically Expressed miRNAs Between Salt-Tolerant R7 and Salt-Sensitive S4

Based on the normalized expression levels, the correlation analysis between the expression levels within each sample was analyzed. The results showed that the correlation coefficient, γ^2 between three biological replicates in each group was 0.878, indicating the replicates were highly consistent (**Supplementary Figure 1B**). There were 188 and 47 specifically expressed miRNAs

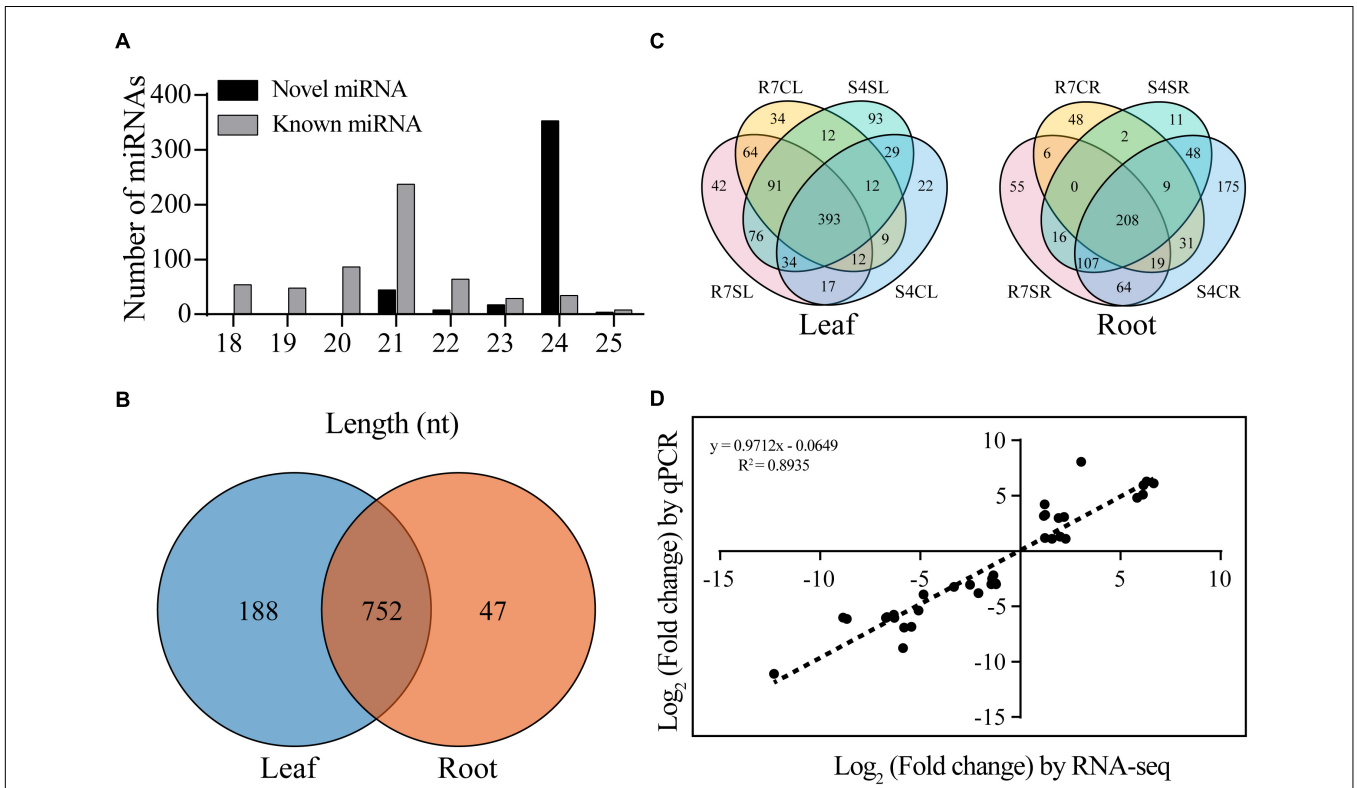


FIGURE 1 | Identification of miRNAs from *Fraxinus velutina* tolerant (R7) and sensitive (S4) leaves and roots. **(A)** Length distribution of all miRNAs identified in this study. **(B)** The distribution of all miRNAs between leaf and root samples. **(C)** The distribution of expressed miRNAs of leaves and roots in R7 and S4 with or without salt treatment. **(D)** The correlation between small RNA sequencing (*x*-axis) and qRT-PCR (*y*-axis) results. R7SL, R7 leaves with salt treatment; R7CL, R7 leaves without salt treatment; R7SR, R7 roots with salt treatment; R7CR, R7 roots without salt treatment; S4SL, S4 leaves with salt treatment; S4CL, S4 leaves without salt treatment; S4SR, S4 roots with salt treatment; S4CR, S4 roots without salt treatment.

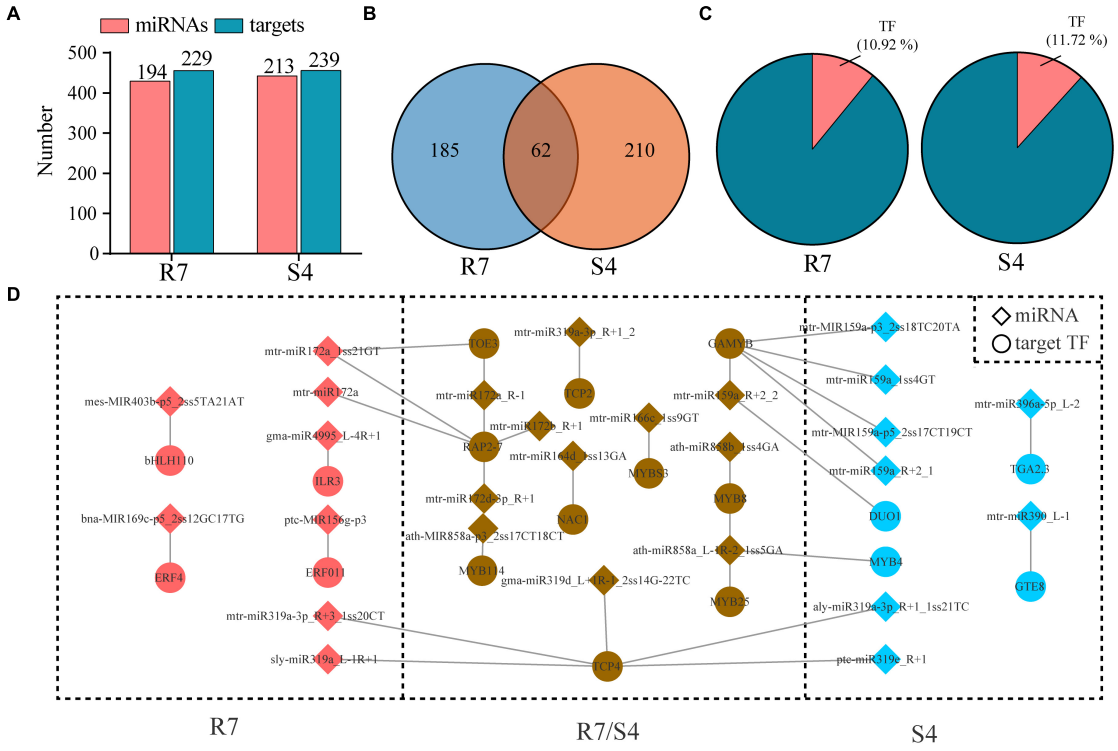


FIGURE 2 | Prediction of miRNA targets based on degradome sequencing data. **(A)** Number of miRNAs and target genes in all miRNA-target pairs in *Fraxinus velutina* tolerant (R7) and sensitive (S4). **(B)** The distribution of miRNA-target pairs between R7 and S4. **(C)** The percentage of transcript factors (TFs) in all targets from R7 and S4. **(D)** Network of all miRNA-TF pairs in R7 and S4.

in leaf and root, respectively, and 752 shared miRNAs in both tissues (**Figure 1B**). There were 796 and 800 miRNAs were expressed in R7 and S4 leaves, respectively; 42 and 93 miRNAs were exclusively expressed in R7 and S4 leaves under salt stress, respectively (**Figure 1C**). In roots, there were 565 and 690 miRNAs expressed in R7 and S4, respectively; 55 and 11 miRNAs were exclusively expressed in R7 and S4 roots under salt treatment, respectively (**Figure 1C**).

To confirm the small RNA sequencing data, 15 DEmiRs were randomly selected to conduct qRT-PCR analysis. The results showed that these selected miRNAs exhibited the same expression patterns with that of small RNA sequencing data (**Supplementary Figure 2**). In addition, based on the $\log_2$ fold change of each comparison, the correlation analysis between qRT-PCR results and sequencing data was performed. The results revealed a high correlation coefficient (R^2 = 0.8935) between sequencing data and qRT-PCR results, demonstrating that the sequencing data are accurate and reliable (**Figure 1D**).

Identification of Target Genes for miRNAs

To investigate biological function of these miRNAs, degradome sequencing from R7 and S4 samples were performed to identify the putative target genes for the miRNAs. After trimmed and polished, 9719583 and 13021448 validated reads were generated from R7 and S4, respectively. The reads were further mapped to the *F. velutina* reference genome, and the results revealed that more than 99% of the reads were perfectly aligned back to the reference (**Supplementary Table 5**). After analyzing by CleaveL and, 247 miRNA-target pairs, including 194 miRNAs targeting 229 genes were identified in R7; 273 pairs, including 213 miRNAs targeting 239 genes were in S4 (**Figure 2A** and **Supplementary Table 6**). Among the miRNA-target pairs, 185 and 210 pairs were exclusively identified in R7 and S4, respectively; 62 pairs were overlapped in both R7 and S4 (**Figure 2B**). Among the targets, we found that there were 25 (10.92%) and 28 (11.72%) genes encoding transcription factors (TF) targeted by the miRNAs in R7 and S4, respectively (**Figure 2C**). For example, an ethylene-responsive transcription factor 4 (*ERF4*) was targeted by bna-MIR169c-p5_2ss12GC17TG in R7, and a transcription factor TGA2.3 (*TGA2.3*) was regulated by mtr-miR390_L-1 in S4. Additionally, we also found multiple shared miRNA-TF pairs in both plants, such as mtr-miR166c_1ss9GT-MYBS3 and mtr-miR164d_1ss13GA-NAC1 (**Figure 2D**).

Identification of Salt Stress-Responsive miRNAs

To identify miRNAs involved in response to salt stress, the differentially expressed miRNAs (DEmiRs) was analyzed using DEseq2. The results revealed 141 (40 up- and 101

down-regulated miRNAs) and 192 (102 up- and 90 down-regulated miRNAs) DEmiRs in R7, and S4 leaves, respectively (**Figure 3A**). In roots, 262 DEmiRs including 124 up- and 138 down-regulated miRNAs were identified in R7, and 249 DEmiRs including 68 up- and 181 down-regulated miRNAs were in S4 (**Figure 3A** and **Supplementary Table 7**). Under salt stress, there were 66 and 172 specific DEmiRs in leaf and root samples, respectively, with 215 shared DEmiRs in both tissues (**Figure 3B** and **Supplementary Table 7**). Among these shared DEmiRs, 27 DEmiRs presented in all comparisons, in which only 8 DEmiRs showed the consistent or opposite expression patterns between these two tissues in two clones after salt treatment (**Figure 3C**). The distribution analysis of leaf DEmiRs revealed that 24 DEmiRs exhibited opposite expression patterns between R7 and S4 after salt treatment (**Figure 3D**). Among the root DEmiRs, 32 DEmiRs presented opposed regulations between R7 and S4 after salt stress (**Figure 3E**).

Analysis of Salt Stress-Responsive miRNA Targets

To explore regulatory roles of DEmiRs in *F. velutina* response to salt stress, our previous transcriptome data were used to narrow the miRNA-target pairs. The results showed that 65 DEmiR-DEG pairs were identified among the comparisons (**Supplementary Table 8**). To further investigate the mechanisms underlying enhanced salt tolerance in R7, the specific DEmiRs in R7 leaves (89 DEmiRs) and roots (138 DEmiRs) were further analyzed. In the leaves, 8 specific DEmiR-DEG pairs, including 8 DEmiRs

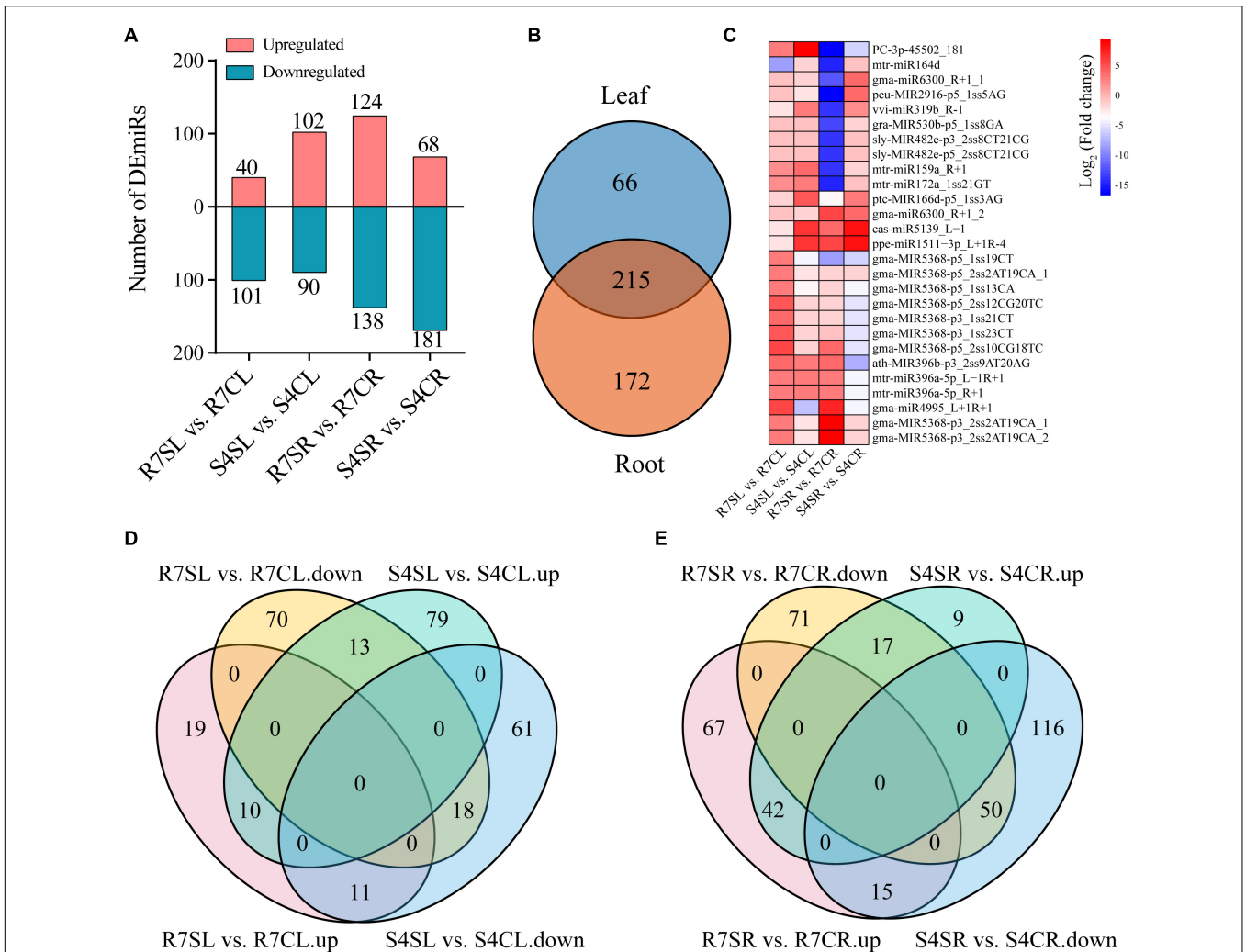


FIGURE 3 | Identification of salt stress-responsive miRNAs. **(A)** Statistics of up- and downregulated differentially expressed miRNAs (DEmiRs) in leaves and roots. The red and green represented up- and down-regulated DEmiRs, respectively. **(B)** The distribution of DEmiRs between leaves and roots. **(C)** Expression patterns of the shared DEmiRs in all comparisons. From red to blue indicated the fold change from high to low. **(D)** The distribution of up- and downregulated DEmiR in R7 and S4 leaves. **(E)** The distribution of up- and down-regulated DEmiR in R7 and S4 roots. R7SL, R7 leaves with salt treatment; R7CL, R7 leaves without salt treatment; R7SR, R7 roots with salt treatment; R7CR, R7 roots without salt treatment; S4SL, S4 leaves with salt treatment; S4CL, S4 leaves without salt treatment; S4SR, S4 roots with salt treatment; S4CR, S4 roots without salt treatment.

TABLE 1 | The specific DEmiR-DEG modules in *Fraxinus velutina* 'R7' leaf under salt stress.

miRNA name	miRNA sequence	R7SL/R7CL	Target transcript	R7SL/R7CL	Annotation	Gene symbol
mtr-miR164d.1	TGGAGAAGCAGGGCACATGCT	−8.86	FRAX13_000377180.1_R0	2.25	NAC domain-containing protein 21/22-like	NAC021
mtr-miR164d.2	TGGAGAAGCAGGACACATGCT	−5.42	FRAX13_000009950.1_R1	2.5	NAC1 transcription factor family protein	NAC1
ptc-miR160g	TGCCTGGCTCCCTGGATGCCA	−5.42	FRAX13_000237750.1_R0	2.33	Auxin response factor 18-like	ARF18
mes-MIR171b-p3	TGATTGAGCCGTGCCAATATC	−3.09	FRAX13_000246880.2_R0	2.95	Scarecrow-like protein 22	SCL22
ptc-MIR171c-p3	TTGAGCCGCGCCAATATCACT	−2.96	FRAX13_000246880.2_R0	2.95	Scarecrow-like protein 22	SCL22
sly-MIR482e-p3	TTTCCTATTCCTCCCATACCGA	−1.28	FRAX13_000069400.1_R0	2.02	Apoptotic ATPase	–
PC-3p-189214	AAGATTGCCCACTGTGGACAGGAG	4.91	FRAX13_000157350.3_R0	−2.59	Fructose-bisphosphate aldolase 1	FBA1
mtr-miR396a-5p	TTCCACAGCTTTCTTGAACTTTT	1.6	FRAX13_000056290.1_R0	−6.63	Growth-regulating factor 1-like	GRF1

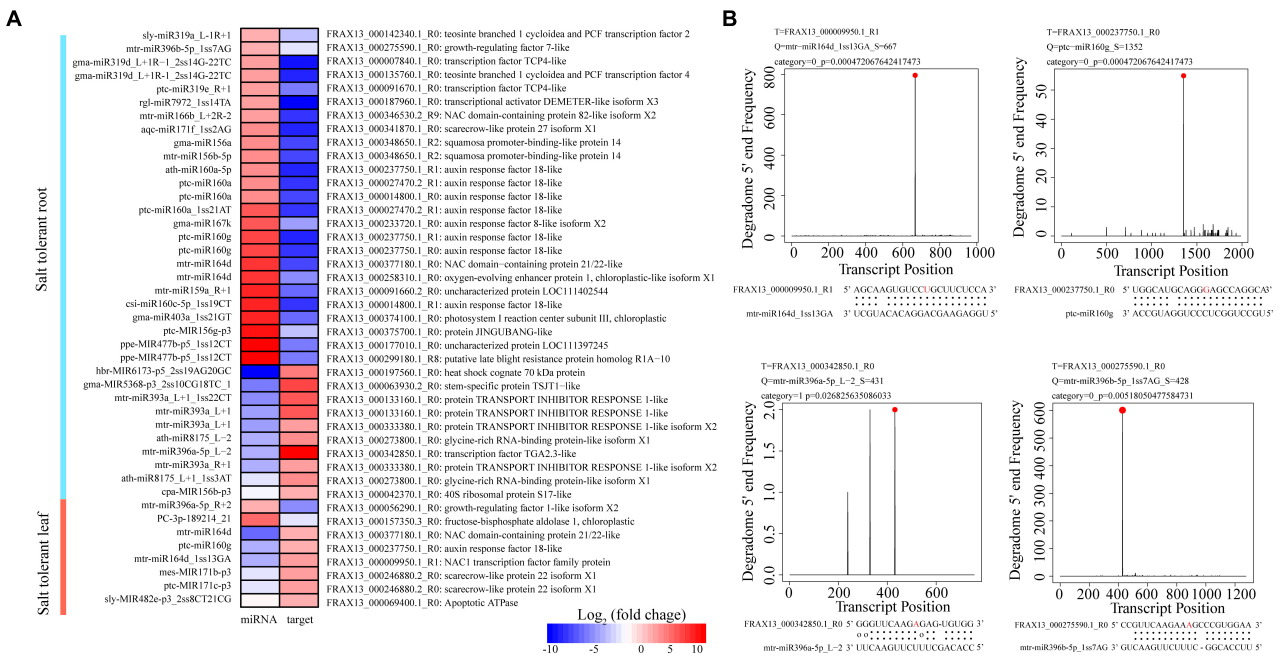


FIGURE 4 | Analysis of salt stress-responsive miRNAs in *F. velutina* tolerant 'R7' leaves and roots. **(A)** Expression patterns of miRNAs and targets from specific DEmiR-DEG pairs in R7 leaves and roots. From red to blue represented the fold change from high to low. **(B)** Target plots (t-plots) of the selected DEmiR-DEG pairs in R7 leaves and roots confirmed by degradome sequencing. The red circle and letter represented slice site.

targeting 7 DEGs were identified (**Table 1**). Among these pairs, multiple DEGs were related to the plant response to salt stress, such as mtr-miR164d_1ss13GA targeted NAC1 transcription factor (*NAC1*), ptc-miR160g targeted *ARF18* (**Figures 4A,B**). In the roots, there were 35 specific DEmiR-DEG pairs including 29 DEmiRs targeting 29 DEGs (**Table 2**). Among these pairs, multiple DEGs were associated with the plant response to salt stress, such as mtr-miR396a-5p_L-2 targeted *TGA2.3* and mtr-miR396b-5p_1ss7AG targeted growth-regulating factor 7 (*GRF7*) (**Figures 4A,B**).

DISCUSSION

In our previous study, we identified two *F. velutina* cuttings clones, salt-tolerant R7 and salt-sensitive S4 and found that R7 exhibits higher salt tolerance than S4. Meanwhile, we performed a comparative transcriptome analysis between R7 and S4, and identified several crucial genes and signaling pathways involved in high salt tolerance of R7 (Ma et al., 2022). In the present study, utilizing an integration of miRNA, mRNA, and degradome sequencing data analysis;

TABLE 2 | The specific DEmiR-DEG modules in *Fraxinus velutina* 'R7' root under salt stress.

miRNA name	miRNA sequence	R7SR/R7CR	Target transcript	R7SR/R7CR	Annotation	Gene symbol
aqc-miR171f	TGATTGAGCCGTGCCAATATC	3.28	FRAX13_000341870.1_R0	-12.02	Scarecrow protein 27	SCL27
ath-miR160a-5p	TGCCTGGCTCCCTGTATGCCA	3.96	FRAX13_000237750.1_R1	-11.55	Auxin response factor 18	ARF18
ath-miR8175.1	CGTTCCCCGGCAACGGCGCCA	−2.8	FRAX13_000273800.1_R0	3.55	Glycine-rich RNA-binding protein	GR-RBP
ath-miR8175.2	TCCCCGGCAACGGCGCCA	−5.62	FRAX13_000273800.1_R0	3.55	Glycine-rich RNA-binding protein	GR-RBP
cpa-MIR156b-p3	GCTCACTTCTCTTTCTGTCAGC	−1.91	FRAX13_000042370.1_R0	2.35	40S ribosomal protein S17	RibS17
csi-miR160c-5p	TGCCTGGCTCCCTGTATGTTT	8.4	FRAX13_000014800.1_R1	-10.82	Auxin response factor 18	ARF18
gma-miR156a	TGACAGAAGAGAGTGAGCAC	3.52	FRAX13_000348650.1_R2	-9.94	Squamosa promoter-binding protein 14	SBP14
gma-miR167k	TGAAGCTGCCAGCCTGATCTTA	6.25	FRAX13_000233720.1_R0	-6.15	Auxin response factor 8	ARF18
gma-miR319d	TTGGACTGAAGGGAGCTCCTC	2.38	FRAX13_000007840.1_R0	-13.08	Transcription factor TCP4	TCP4
gma-miR319d	TTGGACTGAAGGGAGCTCCTC	2.38	FRAX13_000135760.1_R0	-11.57	Transcription factor TCP4	TCP4
gma-miR403a	TTAGATTCACGCACAAACTTT	8.58	FRAX13_000374100.1_R0	-8.67	Photosystem I reaction center subunit III	PsaF
gma-MIR5368-p3	TGGGATTGGGTTTGGGCC	−7.99	FRAX13_000063930.2_R0	7.12	Stem-specific protein TSJT1	TSJT1
hbr-MIR6173-p5	GATACCCCAGTAGTCCTAGCC	−13.69	FRAX13_000197560.1_R0	4.33	Heat shock cognate 70 kDa protein	HSP70
mtr-miR156b-5p	TGACAGAAGAGAGTGAGCAC	3.52	FRAX13_000348650.1_R2	-9.94	Squamosa promoter-binding protein 14	SBP14
mtr-miR159a	TTTGGATTGAAGGGAGCTCTAA	8.11	FRAX13_000091660.2_R0	-8.75	Uncharacterized protein	-
mtr-miR164d	TGGAGAAGCAGGGCACATGCT	7.94	FRAX13_000377180.1_R0	-10.59	NAC domain-containing protein 21/22	NAC021
mtr-miR164d	TGGAGAAGCAGGGCACATGCT	7.94	FRAX13_000258310.1_R0	-6.67	Oxygen-evolving enhancer protein 1	PSBO1
mtr-miR166b	TCTCGGACCAGGCTTCATTCC	2.92	FRAX13_000346530.2_R9	-11.21	NAC domain-containing protein 82	NAC082
mtr-miR393a.1	TTCCAAAGGGATCGCATTGATC	−6.32	FRAX13_000133160.1_R0	6.07	TRANSPORT INHIBITOR RESPONSE 1	TIR1
mtr-miR393a.1	TTCCAAAGGGATCGCATTGATC	−6.32	FRAX13_000333380.1_R0	3.12	TRANSPORT INHIBITOR RESPONSE 1	TIR1
mtr-miR393a.2	TTCCAAAGGGATCGCATTGATT	−7.31	FRAX13_000133160.1_R0	6.07	TRANSPORT INHIBITOR RESPONSE 1	TIR1
mtr-miR393a.3	TCCAAAGGGATCGCATTGATCT	−5.01	FRAX13_000333380.1_R0	3.12	TRANSPORT INHIBITOR RESPONSE 1	TIR1
mtr-miR396a-5p.1	CCACAGCTTTCTTGAACTT	−5.52	FRAX13_000342850.1_R0	9.76	Transcription factor TGA2.3	TGA2.3
mtr-miR396b-5p.2	TTCCACGGCTTTCTTGAACTG	1.91	FRAX13_000275590.1_R0	-2.54	Growth-regulating factor 7	GRF14
ppe-MIR477b-p5	CCTCAAGGGCTTCCAATATTCC	10.53	FRAX13_000177010.1_R0	-7.99	Uncharacterized protein	-
ppe-MIR477b-p5	CCTCAAGGGCTTCCAATATTCC	10.53	FRAX13_000299180.1_R8	-7.76	Putative late blight resistance proteinhomolog R1A-10	R1A-10
ptc-MIR156g-p3	GCTCTCTAGTCTTCTGTCATC	9.29	FRAX13_000375700.1_R0	-4.34	Protein JINGUBANG	JGB
ptc-miR160a.1	TGCCTGGCTCCCTGTATGCCA	3.96	FRAX13_000027470.2_R1	-10.75	Auxin response factor 18	ARF18
ptc-miR160a.1	TGCCTGGCTCCCTGTATGCCA	3.96	FRAX13_000014800.1_R0	-10.49	Auxin response factor 18	ARF18
ptc-miR160a.2	TGCCTGGCTCCCTGTATGCCT	6.16	FRAX13_000027470.2_R1	-10.75	Auxin response factor 18	ARF18
ptc-miR160g	TGCCTGGCTCCCTGGATGCCA	6.66	FRAX13_000237750.1_R1	-11.55	Auxin response factor 18	ARF18
ptc-miR160g	TGCCTGGCTCCCTGGATGCCA	6.66	FRAX13_000237750.1_R0	-10.76	Auxin response factor 18	ARF18
ptc-miR319e	TTGGACTGAAGGGAGCTCCTC	2.38	FRAX13_000091670.1_R0	-8.21	Transcription factor TCP4	TCP4
rgl-miR7972	TTGTCAGGCTTGTAATTCTCC	2.67	FRAX13_000187960.1_R0	-13.96	Transcriptional activator DEMETER	DEM
sly-miR319a	TTGGACTGAAGGGAGCTCCT	1.68	FRAX13_000142340.1_R0	-4.74	Transcription factor TCP2	TCP2

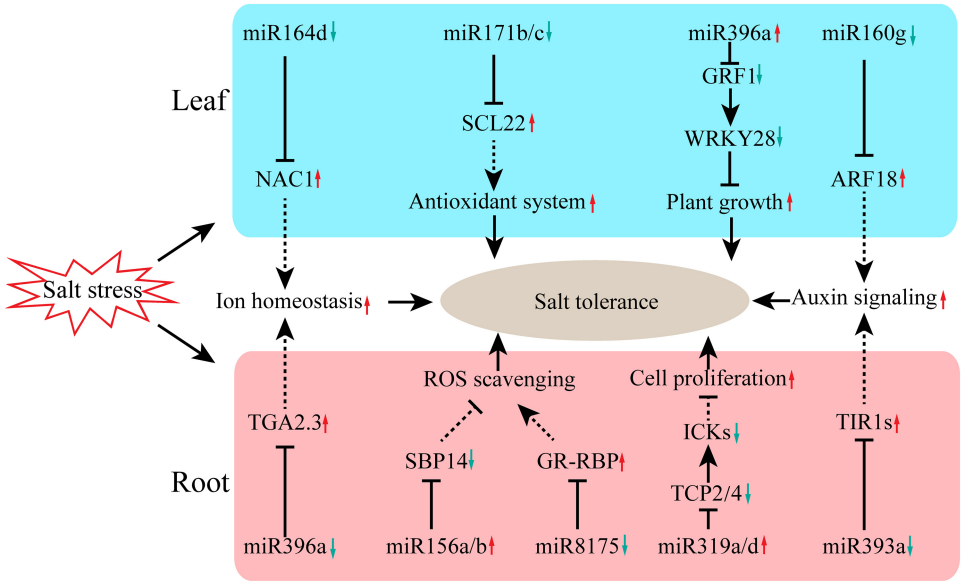


FIGURE 5 | The miRNA-related molecular mechanisms for enhancing salt tolerance of *F. velutina* 'R7'. Under salt stress, multiple miRNA/target modules were involved in R7 response to salt stress. In the leaf, miR164d/NAC1, miR171/SCL22, miR396a/GRF1/WRKY28, and miR160g/ARF18 modules were involved in plant response to salt stress by regulating multiple biological processes, such as antioxidant system and auxin signaling. In the root, miR396a/TGA2.3, miR156/SBP14, miR319/TCPs/ICKs, and miR393a/TIR1 modules enhanced plant tolerance to salt stress by regulating several processes including ROS scavenging, cell proliferation, and ion homeostasis.

we identified several key miRNA-target modules contributing to the high salt tolerance of R7. Specifically, the miRNA-target modules identified in R7 leaf were primarily related with antioxidant system and auxin signaling; while miRNA-target modules in R7 root mainly belonged to ion homeostasis and ROS scavenging.

It is well reported that *miR171/GRAS* module is an important contributor to plant development and biotic and abiotic stress resistance in *Medicago truncatula* (Hirsch et al., 2009). In apple, *miR171i/SCL26.1* module can enhance drought stress tolerance via regulating antioxidant system (Wang et al., 2020). In the present study, we found both miR171b (mes-MIR171b-p3) and miR171c (mes-MIR171b-p3) upregulated a GRAS transcription factor *SCL22* (FRAX13_000246880.2_R0), suggesting that *miR171b/c-SCL22* module is involved in regulating antioxidant system, resulting in the enhanced tolerance to salt stress. The *miR396a-5p/GRF1* module enhances tobacco (*Nicotiana tabacum*) tolerance to salt stress (Chen L.et al., 2015). Under stress, *GRF1* inhibits plant growth by regulating *WRKY28* expression in *Arabidopsis* (Piya et al., 2020). The present study found that miR396a (mtr-miR396a-5p) over-expression down-regulated *GRF1* (FRAX13_000056290.1_R0), further down-regulated *WRKY28*, implying that *miR396a/GRF1/WRKY28* module enhances tolerance to salt stress by promoting plant growth. Thus, *miR171b/c-SCL22* and *miR396a/GRF1/WRKY28* modules are important regulators in enhancing salt stress resistance in *F. velutina*.

Auxin response factors (ARFs) have been confirmed to play crucial roles in plant tolerance to abiotic stress through regulating the auxin signaling pathway (Liu et al., 2018; Song X.et al., 2019; Cui et al., 2020). In poplar, *miR390/TAS3/ARFs* module has been confirmed to be a key regulator of lateral root growth of poplar (*Populus* spp.) plants under salt stress by modulating the auxin pathway (He et al., 2018). In the present study, we found miR160g (ptc-miR160g) up-regulated *ARF18* (FRAX13_000237750.1_R0) expression, suggesting that the *miR160g/ARF18* module contributes to salt tolerance via modulating auxin pathway. Previous study has shown that the overexpressed miR393 regulates rice salt and drought tolerance by inhibiting transport inhibitor response protein (*TIR1*) (Xia et al., 2012). In this study, we found that miR393a (mtr-miR393a.1, mtr-miR393a.2, and mtr-miR393a.3) down-regulated *TIR1* (FRAX13_000333380.1_R0) under salt stress, indicating that *miR393a/TIR1* module enhances plant tolerance to salt stress via mediating auxin signaling pathway. Together, *miR160g/ARF18* and *miR393a/TIR1* modules enhance tolerance to salt stress by regulating auxin signaling pathway in *F. velutina*.

Much evidence shows that NAC transcription factors play vital roles in plant development (Guo et al., 2005; Petricka et al., 2012), cell apoptosis (Lee et al., 2014), and abiotic stress tolerance (Tran et al., 2010). Overexpressed *PeNAC1* in *Arabidopsis* enhances tolerance to salt stress by regulating Na^+/K^+ homeostasis (Wang J. Y. et al., 2013). In this study, we found that miR164d (mtr-miR164d) targeted *NAC1* (FRAX13_000009950.1_R1) and up-regulated NAC1 expression in R7 subjected to salt stress, suggesting that *miR164d/NAC1*

module enhances salt tolerance through regulating Na^+/K^+ homeostasis. The increasing evidence has shown that the TGA (TGACG motif-binding factor) transcription factors, a basic leucine zipper (*bZIP*) gene subfamily, play crucial roles in response to salt stress (Du et al., 2014; Zhang et al., 2014). The overexpression of *GmTGA13* enhances *Arabidopsis* and soybean (*Glycine max*) salt tolerance through regulating ion homeostasis (Ke et al., 2021). In the present study, we found *TGA2.3* (FRAX13_000342850.1_R0) was targeted by miR396a (mtr-miR396a-5p.1), suggesting that *miR396a/TGA2.3* module enhances salt tolerance via mediating ion homeostasis. Thus, *miR164d/NAC1* and *miR396a/TGA2.3* modules contribute to the enhanced salt tolerance in *F. velutina*.

Multiple studies have shown that squamosa-promoter binding protein box (*SBP*-box) can regulate the salt tolerance in many plants, such as *Betula platyphylla* Suk (Ning et al., 2017) and rice (Lan et al., 2019). Silencing *CaSBP12* in pepper (*Capsicum annuum*) enhances tolerance to salt stress and reduces the ROS accumulation (Zhang H. X.et al., 2020). In this study, we found that miR156a/b (gma-miR156a and mtr-miR156b-5p) targeted *SBP14* (FRAX13_000348650.1_R2) and up-regulated *SBP14* expression under salt stress, suggesting that *miR156a/b-SBP14* module increases ROS scavenging, thereafter leading to the enhanced salt tolerance. A previous study has shown that GR-RBPs is a positive regulatory molecule in regulating ROS accumulation to enhance salt tolerance in *Arabidopsis* (Tan et al., 2014). Our study showed that miR8175 (ath-miR8175.1 and ath-miR8175.2) directly bonded to *GR-RBP* (FRAX13_000273800.1_R0) and up-regulated *GR-RBP* expression in R7 subjected to salt stress, indicating that *miR8175/GR-RBP* module enhances tolerance to salt stress. Collectively, *miR156a/b-SBP14* and *miR8175/GR-RBP* modules are key regulators, regulating ROS scavenging, of high salt tolerance in *F. velutina*.

CONCLUSION

In summary, this work revealed the key regulatory network of miRNAs in salt-tolerant clone R7 of *F. velutina* responding to salt stress. The small RNA and degradome sequencing data presented here allowed us to propose potential regulatory roles of miRNAs in the defense response of R7 to salt stress (**Figure 5**). Under salt stress, multiple miRNA/target modules were involved in R7 response to salt stress. In the leaf, *miR164d/NAC1*, *miR171/SCL22*, *miR396a/GRF1/WRKY28*, and *miR160g/ARF18* modules were involved in plant response to salt stress by regulating multiple biological processes, such as antioxidant system and auxin signaling. In the root, *miR396a/TGA2.3*, *miR156/SBP14*, *miR319/TCPs/ICKs*, and *miR393a/TIR1* modules enhanced plant tolerance to salt stress by regulating several processes including ROS scavenging, cell proliferation, and ion homeostasis. The miRNA-target modules identified here pave a novel avenue for improving the salt tolerance of *F. velutina* through miRNA manipulation.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: The small RNA and degradome sequencing data presented in this study can be available at the Sequence Read Archive under accession number PRJNA793056.

AUTHOR CONTRIBUTIONS

KY and DW conceptualized the research program. JL, XM, and LY finished the analysis of this study and wrote the manuscript. QL, HF, CW, and LW conducted the RNA sequencing data analysis. YD and ZC designed the qRT-PCR experiment and finished the operation. RZ, YB, XL, SG, and RC planted the material and finished the physiology analysis. KY, RC, and DW revised the manuscript. All authors discussed the results, commented on the manuscript, and approved the submitted version.

FUNDING

This work was supported by the Improved Variety Program of Shandong Province of China (2019LZGC009).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.847853/full#supplementary-material

Supplementary Figure 1 | Analysis of miRNA sequencing data. **(A)** Number of conserved miRNAs in each family. **(B)** Heatmap showed the correlation coefficient (r^2) between each sample.

Supplementary Figure 2 | The qRT-PCR analysis was performed to verify the reliability of miRNA sequencing data. The heatmap showed the log_2 fold change of 15 selected miRNAs in different comparisons of qRT-PCR results and sequencing data. From blue to red represented the value of log2 fold change from low to high.

Supplementary Table 1 | All primers used in this study.

Supplementary Table 2 | Summary of the small RNA sequences from each sample.

Supplementary Table 3 | Statistics of the small RNAs identified in each sample.

Supplementary Table 4 | The details of all miRNAs identified in this study.

Supplementary Table 5 | The overview of the degradome sequencing data.

Supplementary Table 6 | Identification of target genes of the miRNAs.

Supplementary Table 7 | Identification of the differentially expressed miRNAs.

Supplementary Table 8 | Identification of target genes of the differentially expressed miRNAs.

REFERENCES

Addo-Quaye, C., Eshoo, T. W., Bartel, D. P., and Axtell, M. J. (2008). Endogenous siRNA and miRNA targets identified by sequencing of the *Arabidopsis* degradome. *Curr. Biol.* 18, 758–762. doi: 10.1016/j.cub.2008.04.042

Addo-Quaye, C., Miller, W., and Axtell, M. J. (2009). CleaveLand: a pipeline for using degradome data to find cleaved small RNA targets. *Bioinformatics* 25, 130–131. doi: 10.1093/bioinformatics/btn604

Allen, E., Xie, Z., Gustafson, A. M., and Carrington, J. C. (2005). microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell* 121, 207–221. doi: 10.1016/j.cell.2005.04.004

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/s0022-2836(05)80360-2

Bai, Q., Wang, X., Chen, X., Shi, G., Liu, Z., Guo, C., et al. (2018). Wheat miRNA taemiR408 acts as an essential mediator in plant tolerance to Pi deprivation and salt stress via modulating stress-associated physiological processes. *Front. Plant Sci.* 9:499. doi: 10.3389/fpls.2018.00499

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Butcher, K., Wick, A. F., DeSutter, T., Chatterjee, A., and Harmon, J. (2016). Soil salinity: a threat to global food security. *Agron. J.* 108, 2189–2200. doi: 10.2134/agronj2016.06.0368

Cervera-Seco, L., Marques, M. A. C., Sanz-Carbonell, A., Marquez-Molins, J., Carbonell, A., Darï, S. J., et al. (2019). Identification and characterization of stress-responsive TAS3-derived tasiRNAs in melon. *Plant Cell Physiol.* 60, 2382–2393. doi: 10.1093/pcp/pcz131

Chen, J., and Mueller, V. (2018). Coastal climate change, soil salinity and human migration in Bangladesh. *Nat. Clim. Chang.* 8, 981–985. doi: 10.1038/s41558-018-0313-8

Chen, L., Luan, Y., and Zhai, J. (2015). Sp-miR396a-5p acts as a stress-responsive genes regulator by conferring tolerance to abiotic stresses and susceptibility to phytophthora nicotianae infection in transgenic tobacco. *Plant Cell Rep.* 34, 2013–2025. doi: 10.1007/s00299-015-1847-0

Chen, Z., Hu, L., Han, N., Hu, J., Yang, Y., Xiang, T., et al. (2015). Overexpression of a miR393-resistant form of transport inhibitor response protein 1 (mTIR1) enhances salt tolerance by increased osmoregulation and Na^+ exclusion in *Arabidopsis thaliana*. *Plant Cell Physiol.* 56, 73–83. doi: 10.1093/pcp/pcu149

Cui, J., Li, X., Li, J., Wang, C., Cheng, D., and Dai, C. (2020). Genome-wide sequence identification and expression analysis of ARF family in sugar beet (*Beta vulgaris* L.) under salinity stresses. *PeerJ* 8:e9131. doi: 10.7717/peerj.9131

Denver, J. B., and Ullah, H. (2019). miR393s regulate salt stress response pathway in *Arabidopsis thaliana* through scaffold protein RACK1A mediated ABA signaling pathways. *Plant Signal. Behav.* 14:1600394. doi: 10.1080/15592324.2019.1600394

Djami-Tchatchou, A. T., Sanan-Mishra, N., Ntushelo, K., and Dubery, I. A. (2017). Functional roles of microRNAs in agronomically important plants-potential as targets for crop improvement and protection. *Front. Plant Sci.* 8:378. doi: 10.3389/fpls.2017.00378

Du, X., Du, B., Chen, X., Zhang, S., Zhang, Z., and Qu, S. (2014). Overexpression of the MhTGA2 gene from crab apple (*Malus hupehensis*) confers increased tolerance to salt stress in transgenic apple (*Malus domestica*). *J. Agric. Sci.* 152, 634–641. doi: 10.1017/S0021859613000130

German, M. A., Luo, S., Schroth, G., Meyers, B. C., and Green, P. J. (2009). Construction of parallel analysis of RNA ends (PARE) libraries for the study of cleaved miRNA targets and the RNA degradome. *Nat. Protoc.* 4, 356–362. doi: 10.1038/nprot.2009.8

German, M. A., Pillay, M., Jeong, D. H., Hetawal, A., Luo, S., Janardhanan, P., et al. (2008). Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. *Nat. Biotechnol.* 26, 941–946. doi: 10.1038/nbt1417

Guo, H. S., Xie, Q., Fei, J. F., and Chua, N. H. (2005). MicroRNA directs mRNA cleavage of the transcription factor NAC1 to downregulate auxin signals for arabidopsis lateral root development. *Plant Cell* 17, 1376–1386. doi: 10.1105/tpc.105.030841

He, F., Xu, C., Fu, X., Shen, Y., Guo, L., Leng, M., et al. (2018). The microRNA390/trans-acting short interfering RNA3 module mediates lateral root growth under salt stress via the auxin pathway. *Plant Physiol.* 177, 775–791. doi: 10.1104/pp.17.01559

Hirsch, S., Kim, J., Muñoz, A., Heckmann, A. B., Downie, J. A., and Oldroyd, G. E. (2009). GRAS proteins form a DNA binding complex to induce gene expression during nodulation signaling in *Medicago truncatula*. *Plant Cell* 21, 545–557. doi: 10.1105/tpc.108.064501

Hossain, M. S., and Dietz, K. J. (2016). Tuning of redox regulatory mechanisms, reactive oxygen species and redox homeostasis under salinity stress. *Front. Plant Sci.* 7:548. doi: 10.3389/fpls.2016.00548

Iglesias, M. J., Terrile, M. C., Windels, D., Lombardo, M. C., Bartoli, C. G., Vazquez, F., et al. (2014). MiR393 regulation of auxin signaling and redox-related components during acclimation to salinity in *Arabidopsis*. *PLoS One* 9:e107678. doi: 10.1371/journal.pone.0107678

Ke, D., He, Y., Fan, L., Niu, R., Cheng, L., Wang, L., et al. (2021). The soybean TGA transcription factor GmTGA13 plays important roles in the response to salinity stress. *Plant Biol. (Stuttg)* 24, 313–322. doi: 10.1111/plb.13360

Kechin, A., Boyarskikh, U., Kel, A., and Filipenko, M. (2017). cutPrimers: a new tool for accurate cutting of primers from reads of targeted next generation sequencing. *J. Comput. Biol.* 24, 1138–1143. doi: 10.1089/cmb.2017.0096

Kelly, L. J., Plumb, W. J., Carey, D. W., Mason, M. E., Cooper, E. D., Crowther, W., et al. (2020). Convergent molecular evolution among ash species resistant to the emerald ash borer. *Nat. Ecol. Evol.* 4, 1116–1128. doi: 10.1038/s41559-020-1209-3

Kumar, D., Dutta, S., Singh, D., Prabhu, K. V., Kumar, M., and Mukhopadhyay, K. (2017). Uncovering leaf rust responsive miRNAs in wheat (*Triticum aestivum* L.) using high-throughput sequencing and prediction of their targets through degradome analysis. *Planta* 245, 161–182. doi: 10.1007/s00425-016-2600-9

Kumar, V., Khare, T., Shriram, V., and Wani, S. H. (2018). Plant small RNAs: the essential epigenetic regulators of gene expression for salt-stress responses and tolerance. *Plant Cell Rep.* 37, 61–75. doi: 10.1007/s00299-017-2210-4

Lan, T., Zheng, Y., Su, Z., Yu, S., Song, H., Zheng, X., et al. (2019). OsSPL10, a SBP-box gene, plays a dual role in salt tolerance and trichome formation in rice (*Oryza sativa* L.). *G3 (Bethesda)* 9, 4107–4114. doi: 10.1534/g3.119.400700

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10:R25. doi: 10.1186/gb-2009-10-3-r25

Lee, S., Lee, H. J., Huh, S. U., Paek, K. H., Ha, J. H., and Park, C. M. (2014). The *Arabidopsis* NAC transcription factor NTL4 participates in a positive feedback loop that induces programmed cell death under heat stress conditions. *Plant Sci.* 227, 76–83. doi: 10.1016/j.plantsci.2014.07.003

Litalien, A., and Zeeb, B. (2020). Curing the earth: a review of anthropogenic soil salinization and plant-based strategies for sustainable mitigation. *Sci. Total Environ.* 698:134235. doi: 10.1016/j.scitotenv.2019.134235

Liu, N., Dong, L., Deng, X., Liu, D., Liu, Y., Li, M., et al. (2018). Genome-wide identification, molecular evolution, and expression analysis of auxin response factor (ARF) gene family in *Brachypodium distachyon* L. *BMC Plant Biol.* 18:336. doi: 10.1186/s12870-018-1559-z

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8

Luan, M., Xu, M., Lu, Y., Zhang, L., Fan, Y., and Wang, L. (2015). Expression of zma-miR169 miRNAs and their target ZmNF-YA genes in response to abiotic stress in maize leaves. *Gene* 555, 178–185. doi: 10.1016/j.gene.2014.11.001

Luan, M., Xu, M., Lu, Y., Zhang, Q., Zhang, L., Zhang, C., et al. (2014). Family-wide survey of miR169s and NF-YAs and their expression profiles response to abiotic stress in maize roots. *PLoS One* 9:e91369. doi: 10.1371/journal.pone.0091369

Ma, X., Liu, J. N., Yan, L., Liang, Q., Fang, H., Wang, C., et al. (2022). Comparative transcriptome analysis unravels defense pathways of *Fraxinus velutina* Torr against salt stress. *Front. Plant Sci.* 2022:842726. doi: 10.3389/fpls.2022.842726

Ma, Y., Xue, H., Zhang, F., Jiang, Q., Yang, S., Yue, P., et al. (2021). The miR156/SPL module regulates apple salt stress tolerance by activating *MdWRKY100* expression. *Plant Biotechnol. J.* 19, 311–323. doi: 10.1111/pbi.13464

Mao, P., Tang, Q., Cao, B., Liu, J., Shao, H., Cao, Z., et al. (2017). Eco-physiological adaptability in mixtures of *Robinia pseudoacacia* and *Fraxinus velutina* and

coastal eco-engineering. *Ecol. Eng.* 106, 109–115. doi: 10.1016/j.ecoleng.2017.05.021

Markham, N. R., and Zuker, M. (2008). UNAFold: software for nucleic acid folding and hybridization. *Methods Mol. Biol.* 453, 3–31. doi: 10.1007/978-1-60327-429-6_1

Meyers, B. C., Axtell, M. J., Bartel, B., Bartel, D. P., Baulcombe, D., Bowman, J. L., et al. (2008). Criteria for annotation of plant microRNAs. *Plant Cell* 20, 3186–3190. doi: 10.1105/tpc.108.064311

Ning, K., Chen, S., Huang, H., Jiang, J., Yuan, H., and Li, H. (2017). Molecular characterization and expression analysis of the SPL gene family with BpSPL9 transgenic lines found to confer tolerance to abiotic stress in *Betula platyphylla* Suk. *Plant Cell Tissue Organ. Cult. (PCTOC)* 130, 469–481. doi: 10.1007/s11240-017-1226-3

Ondrasek, G., and Rengel, Z. (2021). Environmental salinization processes: detection, implications & solutions. *Sci. Total Environ.* 754, 142432. doi: 10.1016/j.scitotenv.2020.142432

Petricka, J. J., Winter, C. M., and Benfey, P. N. (2012). Control of *Arabidopsis* root development. *Annu. Rev. Plant Biol.* 63, 563–590. doi: 10.1146/annurev-arplant-042811-105501

Piya, S., Liu, J., Burch-Smith, T., Baum, T. J., and Hewezi, T. (2020). A role for *Arabidopsis* growth-regulating factors 1 and 3 in growth-stress antagonism. *J. Exp. Bot.* 71, 1402–1417. doi: 10.1093/jxb/erz502

Polle, A., and Chen, S. (2015). On the salty side of life: molecular, physiological and anatomical adaptation and acclimation of trees to extreme habitats. *Plant Cell Environ.* 38, 1794–1816. doi: 10.1111/pce.12440

Shamimuzzaman, M., and Vodkin, L. (2012). Identification of soybean seed developmental stage-specific and tissue-specific miRNA targets by degradome sequencing. *BMC Genomics* 13:310. doi: 10.1186/1471-2164-13-310

Song, S., Hao, L., Zhao, P., Xu, Y., Zhong, N., Zhang, H., et al. (2019). Genome-wide identification, expression profiling and evolutionary analysis of auxin response factor gene family in potato (*Solanum tuberosum* Group Phureja). *Sci. Rep.* 9:1755. doi: 10.1038/s41598-018-37923-7

Song, X., Li, Y., Cao, X., and Qi, Y. (2019). MicroRNAs and their regulatory roles in plant-environment interactions. *Annu. Rev. Plant Biol.* 70, 489–525. doi: 10.1146/annurev-arplant-050718-100334

Sun, X., Zhang, Y., Zhu, X., Korir, N. K., Tao, R., Wang, C., et al. (2014). Advances in identification and validation of plant microRNAs and their target genes. *Physiol. Plant* 152, 203–218. doi: 10.1111/ppl.12191

Tan, Y., Qin, Y., Li, Y., Li, M., and Ma, F. (2014). Overexpression of MpGR-RBP1, a glycine-rich RNA-binding protein gene from *Malus prunifolia* (Willd.) Borkh., confers salt stress tolerance and protects against oxidative stress in *Arabidopsis*. *Plant Cell Tissue Organ. Cult. (PCTOC)* 119, 635–646. doi: 10.1007/s11240-014-0563-8

Tran, L. S., Nishiyama, R., Yamaguchi-Shinozaki, K., and Shinozaki, K. (2010). Potential utilization of NAC transcription factors to enhance abiotic stress tolerance in plants by biotechnological approach. *GM Crops* 1, 32–39. doi: 10.4161/gmcr.1.1.10569

Wang, C., Han, J., Korir, N. K., Wang, X., Liu, H., Li, X., et al. (2013). Characterization of target mRNAs for grapevine microRNAs with an integrated strategy of modified RLM-RACE, newly developed PPM-RACE and qPCRs. *J. Plant Physiol.* 170, 943–957. doi: 10.1016/j.jplph.2013.02.005

Wang, J. Y., Wang, J. P., and He, Y. (2013). A *Populus euphratica* NAC protein regulating Na+/K+ homeostasis improves salt tolerance in *Arabidopsis thaliana*. *Gene* 521, 265–273. doi: 10.1016/j.gene.2013.03.068

Wang, M., Guo, W., Li, J., Pan, X., Pan, L., Zhao, J., et al. (2021). The miR528-AO module confers enhanced salt tolerance in rice by modulating the ascorbic acid and abscisic acid metabolism and ROS scavenging. *J. Agric. Food Chem.* 69, 8634–8648. doi: 10.1021/acs.jafc.1c01096

Wang, W., Liu, D., Chen, D., Cheng, Y., Zhang, X., Song, L., et al. (2019). MicroRNA414c affects salt tolerance of cotton by regulating reactive oxygen species metabolism under salinity stress. *RNA Biol.* 16, 362–375. doi: 10.1080/15476286.2019.1574163

Wang, Y., Feng, C., Zhai, Z., Peng, X., Wang, Y., Sun, Y., et al. (2020). The apple microR171i-SCARECROW-LIKE PROTEINS26.1 module enhances drought stress tolerance by integrating ascorbic acid metabolism. *Plant Physiol.* 184, 194–211. doi: 10.1104/pp.20.00476

Xia, K., Wang, R., Ou, X., Fang, Z., Tian, C., Duan, J., et al. (2012). OsTIR1 and OsAFB2 downregulation via OsmiR393 overexpression leads to more tillers, early flowering and less tolerance to salt and drought in rice. *PLoS One* 7:e30039. doi: 10.1371/journal.pone.0030039

Xu, T., Zhang, L., Yang, Z., Wei, Y., and Dong, T. (2021). Identification and functional characterization of plant miRNA under salt stress shed light on salinity resistance improvement through miRNA manipulation in crops. *Front. Plant Sci.* 12:665439. doi: 10.3389/fpls.2021.665439

Yin, Z., Li, C., Han, X., and Shen, F. (2008). Identification of conserved microRNAs and their target genes in tomato (*Lycopersicon esculentum*). *Gene* 414, 60–66. doi: 10.1016/j.gene.2008.02.007

Yu, Y., Wu, G., Yuan, H., Cheng, L., Zhao, D., Huang, W., et al. (2016). Identification and characterization of miRNAs and targets in flax (*Linum usitatissimum*) under saline, alkaline, and saline-alkaline stresses. *BMC Plant Biol.* 16:124. doi: 10.1186/s12870-016-0808-2

Yuan, S., Li, Z., Li, D., Yuan, N., Hu, Q., and Luo, H. (2015). Constitutive expression of rice microRNA528 alters plant development and enhances tolerance to salinity stress and nitrogen starvation in creeping bentgrass. *Plant Physiol.* 169, 576–593. doi: 10.1104/pp.15.00899

Yuan, S., Zhao, J., Li, Z., Hu, Q., Yuan, N., Zhou, M., et al. (2019). MicroRNA396-mediated alteration in plant development and salinity stress response in creeping bentgrass. *Hortic. Res.* 6:48. doi: 10.1038/s41438-019-0130-x

Zhang, B. H., Pan, X. P., Cox, S. B., Cobb, G. P., and Anderson, T. A. (2006). Evidence that miRNAs are different from other RNAs. *Cell Mol. Life Sci.* 63, 246–254. doi: 10.1007/s00018-005-5467-7

Zhang, H. X., Zhu, W. C., Feng, X. H., Jin, J. H., Wei, A. M., and Gong, Z. H. (2020). Transcription factor CaSBP12 negatively regulates salt stress tolerance in pepper (*Capsicum annuum* L.). *Int. J. Mol. Sci.* 21:444. doi: 10.3390/ijms21020444

Zhang, J. Y., Qu, S. C., Qiao, Y. S., Zhang, Z., and Guo, Z. R. (2014). Overexpression of the *Malus hupehensis* MhNPR1 gene increased tolerance to salt and osmotic stress in transgenic tobacco. *Mol. Biol. Rep.* 41, 1553–1561. doi: 10.1007/s11033-013-3001-9

Zhang, Y., Gong, H., Li, D., Zhou, R., Zhao, F., Zhang, X., et al. (2020). Integrated small RNA and degradome sequencing provide insights into salt tolerance in sesame (*Sesamum indicum* L.). *BMC Genomics* 21:494. doi: 10.1186/s12864-020-06913-3

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 08 April 2022
doi: 10.3389/fpls.2022.851531

ZmPP2C26 Alternative Splicing Variants Negatively Regulate Drought Tolerance in Maize

Fengzhong Lu[1†], Wanchen Li[1†], Yalin Peng[1†], Yang Cao[1], Jingtao Qu[1], Fuai Sun[1], Qingqing Yang[1], Yanli Lu[1], Xuehai Zhang[2], Lanjie Zheng[2], Fengling Fu[1] and Haoqiang Yu[1]**

[1] Key Laboratory of Biology and Genetic Improvement of Maize in Southwest Region, Ministry of Agriculture, Maize Research Institute, Sichuan Agricultural University, Chengdu, China, [2] National Key Laboratory of Wheat and Maize Crop Science, Henan Agricultural University, Zhengzhou, China

Serine/threonine protein phosphatase 2C (PP2C) dephosphorylates proteins and plays crucial roles in plant growth, development, and stress response. In this study, we characterized a clade B member of maize PP2C family, i.e., ZmPP2C26, that negatively regulated drought tolerance by dephosphorylating ZmMAPK3 and ZmMAPK7 in maize. The *ZmPP2C26* gene generated *ZmPP2C26L* and *ZmPP2C26S* isoforms through untypical alternative splicing. ZmPP2C26S lost 71 amino acids including an MAPK interaction motif and showed higher phosphatase activity than ZmPP2C26L. ZmPP2C26L directly interacted with, dephosphorylated ZmMAPK3 and ZmMAPK7, and localized in chloroplast and nucleus, but ZmPP2C26S only dephosphorylated ZmMAPK3 and localized in cytosol and nucleus. The expression of *ZmPP2C26L* and *ZmPP2C26* was significantly inhibited by drought stress. Meanwhile, the maize *zmpp2c26* mutant exhibited enhancement of drought tolerance with higher root length, root weight, chlorophyll content, and photosynthetic rate compared with wild type. However, overexpression of *ZmPP2C26L* and *ZmPP2C26S* significantly decreased drought tolerance in *Arabidopsis* and rice with lower root length, chlorophyll content, and photosynthetic rate. Phosphoproteomic analysis revealed that the ZmPP2C26 protein also altered phosphorylation level of proteins involved in photosynthesis. This study provides insights into understanding the mechanism of PP2C in response to abiotic stress.

Keywords: maize, drought stress, protein phosphatase 2C, MAPK, alternative splicing, photosynthesis

OPEN ACCESS

Edited by:
Prasanta Kumar Subudhi, Louisiana State University, United States

Reviewed by:
Mo-Xian Chen, Guizhou University, China
R. Glen Uhrig, University of Alberta, Canada

***Correspondence:**
Fengling Fu
ffl@sicau.edu.cn
Haoqiang Yu
yhq1801@sicau.edu.cn

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

Received: *10 January 2022*
Accepted: *08 March 2022*
Published: *08 April 2022*

Citation:
Lu F, Li W, Peng Y, Cao Y, Qu J, Sun F, Yang Q, Lu Y, Zhang X, Zheng L, Fu F and Yu H (2022) ZmPP2C26 Alternative Splicing Variants Negatively Regulate Drought Tolerance in Maize. Front. Plant Sci. 13:851531. doi: 10.3389/fpls.2022.851531

INTRODUCTION

In plants, numerous proteins will be activated or inactivated *via* dephosphorylation catalyzed by protein phosphatases (PPs) (Cohen, 1989). Based on their substrate specificity, PPs are mainly classified into three families, namely, serine (Ser)/threonine (Thr)-specific phosphoprotein phosphatase (PPP), metal-dependent protein phosphatase (PPM), and protein tyrosine phosphatase (PTP) (Barford et al., 1998). The PPP and PPM families encode Ser/Thr PP, while PTP family includes tyrosine-specific and dual-specificity phosphatase (Barford et al., 1998). The PP2C of PPM family is a kind of the Mg^{2+}- or Mn^{2+}-dependent PPs and specifically dephosphorylates

the phosphorylated Ser/Thr residues of target proteins (Schweighofer et al., 2004; Shi, 2009). The PP2C family is substantially expanded in plants with 80, 90, and 130 members in *Arabidopsis*, rice, and maize, respectively, and divided into eleven clades of A–K (Xue et al., 2008; Singh et al., 2010; Wang et al., 2018). The clade A members of PP2C act as co-receptors of abscisic acid (ABA), interact with the ABA-receptor protein PYR/PYL/PCAR and SNF1-related protein kinase 2s (SnRK2s) to negatively regulate ABA signaling, and play crucial roles in plant growth, development, and stimuli response (Ma et al., 2009; Park et al., 2009; Komatsu et al., 2013). For instance, AHG1 encoding a PP2C interacts with DELAY OF GERMINATION1 (DOG1) and is impaired by DOG1 to negatively regulate ABA response in seed dormancy and germination (Nishimura et al., 2018). The maize *ZmPP2C-A2*, *ZmPP2C-A6*, and *ZmPP2C-A10* negatively regulate drought tolerance by meditating ABA signaling (Xiang et al., 2017; He et al., 2019). Tomato *SlPP2C3* functions as a negative regulator of ABA signaling to negatively regulate drought tolerance, fruit ripening, and glossiness (Liang et al., 2021). In contrast, only few available studies report the function of other clade PP2Cs. A clade G member *AtPP2C49* negatively regulates salt tolerance through inhibition of Na^+ transporter AtHKT1;1 activity (Chu et al., 2021). In *Arabidopsis*, three of six clade B members of PP2C including *AP2C1*, *AP2C3,* and *PP2C5* are well-elucidated for their function in stomata development, immunity, defense, and K^+ deficiency response (Schweighofer et al., 2007; Brock et al., 2010; Umbrasaite et al., 2010; Shubchynskyy et al., 2017; Singh et al., 2018). The tobacco B clade *NtPP2C2b* is found to regulate nicotine biosynthesis (Liu et al., 2021). However, the function of clade B members of PP2C in crops remains unknown.

Alternative splicing (AS) of precursor messenger RNAs (pre-mRNAs) produces more mRNA isoforms from the same pre-mRNA and plays a key role in gene expression and protein diversity (Kelemen et al., 2013). Previous studies show that abundant genes undergo AS to regulate plant growth and stress response, including salt, heat, cold, and drought stress, and photomorphogenesis, flowering, and yield (Farquharson, 2016; Jiang et al., 2017; Qin et al., 2017; Calixto et al., 2018; Gu et al., 2018; Liu et al., 2018; Yu et al., 2018; Li et al., 2019). The AS events likewise can be triggered in different environmental conditions and developmental stages (Liu et al., 2018; Martin et al., 2021). As well known, there are two typical types of intron retention, namely, U2 and U12 type with 5′-GT●●●●●●AG-3′ and 5′-AT●●●●●●AC-3′ splice site, respectively (Sharp and Burge, 1997; Reddy et al., 2013). Intriguingly, two untypical types of AS are recently found in *Arabidopsis*, containing alternative first exon (AFE) and alternative last exon (ALE) (Zhu et al., 2017). The *FT2* gene undergoes AFE-type AS producing *FT2*β and *FT2*α transcripts to control flowering through regulating their amount during reproductive stage (Qin et al., 2017).

Maize is one of the most important crops and is used in food supply, livestock feed, and industries. Its productivity is seriously restricted by drought stress due to its vulnerability to water deficits (Lobell et al., 2014; Sah et al., 2020). Hence, exploring stress-related genes will contribute for facilitating molecular design breeding to improve maize drought tolerance. In our previous studies, a new maize PP2C was found to be clade B member, named as *ZmPP2C26* and inhibited by drought stress (Wang et al., 2018; Lu et al., 2020), indicating its potential role in drought response. In this study, we characterized that *ZmPP2C26* underwent AFE-type AS and generated two isoforms named as *ZmPP2C26L* and *ZmPP2C26S*. Subsequently, their activity, localization, and interacting proteins were analyzed. Their functions in drought tolerance were identified through phenotyping transgenic *Arabidopsis* and rice, as well as *via* maize mutant. Our data clearly demonstrates that *ZmPP2C26* negatively regulates drought tolerance *via* dephosphorylating ZmMAPK3/ZmMAPK7 and impairing photosynthesis in maize.

MATERIALS AND METHODS

Stress Treatment and Expression Analysis

The drought-tolerant maize inbred lines 81565/87-1 and drought-sensitive lines 200B/DAN340 were used for gene expression analysis. The four-leaf stage seedings were treated with 16% (w/v) polyethylene glycol 6000 (PEG-6000). At 0 (blank control), 3, 6, 12, and 24 h of treatment, the seedling of every line was collected and used for total RNA extraction by using RNAiso plus kit (TaKaRa, Japan).

Total RNA samples were reverse-transcribed into cDNA using PrimeScript^{TM} reagent kit (TaKaRa) and used for real-time quantitative PCR (RT-qPCR). The RT-qPCR was performed using ChamQ Universal SYBR qPCR Master Mix (Vazyme, Nanjing) in the CFX96^{TM} Real-Time System (Bio-Rad, Hercules, CA, United States). The *ZmEF1a* gene was used as internal control. The information of all primers used in this study is shown in **Supplementary Table 1**.

Phosphatase Activity Assay

The open reading frame (ORF) of *ZmPP2C26L*/*ZmPP2C26S* without stop codon was amplified and inserted into pET32a vector to generate *His-ZmPP2C26L*/*His-ZmPP2C26S* plasmid, respectively. The reconstructed plasmids were transformed into *Escherichia coli* strain BL21. The *E. coli* strain harboring the above plasmid was induced using 0.1 mM isopropyl-1-thio-b-D-galactopyranoside (IPTG) at 16°C for 16 h to express His-ZmPP2C26L/His-ZmPP2C26S protein, which was purified using by 6 × His-Tagged Protein Purification Kit (CWBIO, China).

Protein phosphatase activity of ZmPP2C26L/ZmPP2C26S protein was detected as previously reported (Han et al., 2018; Yang et al., 2020). Briefly, 2 mg of His-tagged protein was incubated with 1 mL assay buffer (50 mM Tris–HCl, pH 7.5, 1 mM $MgCl_2$, 0.5 mM EDTA, and 0.1 g/L BSA) for 30 min at 37°C. Subsequently, 2 mM p-nitrophenyl phosphate (pNPP) was added into the above mixture to produce p-nitrophenol catalyzed by ZmPP2C26L/ZmPP2C26S protein. Then, the absorbance value at wavelength of 405 nm (OD_{405}) of p-nitrophenol was monitored every 2 min. The relative PP activity was calculated according to the curve of OD_{405}.

Yeast Two-Hybrid, Bimolecular Fluorescence Complementation, and Glutathione-*S*-Transferase Pull-Down

Yeast two-hybrid (Y2H) assay was conducted using the Matchmaker GAL4 Y2H System (Clontech). The ORF of *ZmPP2C26L*/*ZmPP2C26S* was amplified and inserted into prey vector pGADT7 to generate *AD-ZmPP2C26L*/*AD-ZmPP2C26S* plasmid, respectively. The ORF of 13 maize *PYL* genes was inserted into bait vector pGBKT7 for *BD-ZmPYLs* in our previous study (Wang et al., 2018). The ORF of 20 maize *MAPK* genes (e.g., *ZmMAPK1-19* and *ZmSIMK*) was also inserted into pGBKT7 for *BD-ZmMAPKs*, which were kindly provided by Dongtao Ren (China Agriculture University, Chengdu, China). The prey AD plasmid and the bait BD plasmid were cotransformed into *Saccharomyces cerevisiae* strain Y2H gold by using the yeast transformation kit (Coolaber, Beijing, China). The transformants were cultured on synthetic medium plates (SD medium) lacking Trp and Leu (SD/-Trp/-Leu) at 30°C for 2–3 days, then transferred onto SD/-Trp/-Leu/-His/-Ade plates containing 5-bromo-4-chloro-3-indolyl-α-D-galactopyranoside (X-α-gal) for blue color development to detect the interaction between ZmPP2C26L/ZmPP2C26S and ZmPYLs/ZmMAPKs. The interaction was further validated by bimolecular fluorescence complementation (BiFC) and glutathione-*S*-transferase (GST) pull-down assays.

For BiFC assay, the ORF of *ZmPP2C26L/ZmPP2C26S* was cloned into the pSPYNE-*35S-nYFP* vector generating *nYFP-ZmPP2C26L*/*nYFP-ZmPP2C26S*, respectively. The ORF of the *ZmMAPK3*/*ZmMAPK7* gene was inserted into the pSPYNE-*35S-cYFP* vector generating *cYFP-ZmMAPK3*/*cYFP-ZmMAPK7*, respectively. As previously described (Li et al., 2016), the *nYFP-ZmPP2C26L*/*nYFP-ZmPP2C26S* and *cYFP-ZmMAPK3*/*cYFP-ZmMAPK7* constructs were cotransformed into maize protoplast. After 16 h at 28°C, the protoplasts were examined for YFP fluorescence under the confocal laser scanning microscope (ZESS 800, Germany).

For GST pull-down assay, the ORF of *ZmMAPK3*/*ZmMAPK7* was inserted into pGEX-6p-1 vector to create *GST-ZmMAPK3*/*GST-ZmMAPK7*, respectively. The GST-tagged protein was induced by 0.1-mM IPTG and purified using Glutathione-Sepharose Resin kit (CWBIO). A total of 2 μg GST-tagged protein was uploaded to Mag-Beads GST Fusion Protein Purification (Sangon Biotech, Shanghai, China) and incubated at room temperature for 2 h. Then, 2 μg of His-tagged protein was added to protein-beads complex for combination. The protein complex was pulled down by washing five times with the elution buffer containing 50 mM Tris–HCl and 10 mM reduced glutathione (pH 8.0). The proteins were separated by 12.5% SDS-PAGE and transferred onto the PVDF (polyvinylidene fluoride) membrane by wet transfer at 100 V for 80 min. The membrane was blocked in 2.5% (w/v) non-fat milk powder solution (Coolaber, Beijing, China) for 90 min and incubated with primary antibody (anti-GST/anti-His antibody) for 90 min at room temperature and then with the HRP (horseradish peroxidase)-conjugated Goat Anti-Mouse IgG (ABclonal, Wuhan, China) for 60 min at room temperature. Finally, the signal was visualized using the ChemDoc XRS system (Bio-Rad, Hercules, CA, United States).

Subcellular Localization and Co-localization

For subcellular localization, the ORF of *ZmPP2C26L/ZmPP2C26S* without stop codon was amplified and independently inserted into pCAMBIA2300-*35S-eGFP* vector to create *35S-ZmPP2C26L-eGFP/35S-ZmPP2C26S-eGFP* plasmid, respectively. The construct was transformed into *Agrobacterium tumefaciens* strain GV3101 and then used for infiltrating into the leaves of 5-week-old *Nicotiana benthamiana*. The GFP fluorescence was observed using the confocal laser scanning microscope (ZESS 800). For co-localization of ZmPP2C26L/ZmPP2C26S and ZmMAPK3/ZmMAPK7, the ORF of *ZmMAPK3*/*ZmMAPK7* was separately cloned into the pCAMBIA1300-*35S-mCherry* vector to generate *35S-ZmMAPK3-mCherry*/*35S-ZmMAPK7-mCherry*, respectively. The cultures of *Agrobacterium* carrying *35S-ZmPP2C26L-eGFP*/*35S-ZmPP2C26S-eGFP* and *35S-ZmMAPK3-mCherry*/*35S-ZmMAPK7-mCherry* were co-infiltrated into *N. benthamiana* leaves. The GFP and mCherry fluorescence were observed by using the confocal laser scanning microscope (ZESS 800).

Phosphorylation Assay

The phosphorylation assay was performed with minor modification as previously described (Xia et al., 2021). The ORF of *ZmMAPK3*/*ZmMAPK7* was inserted into pCAMBIA1300-*35S-3* × *HA* vector to generate *35S-HA-ZmMAPK3*/*35S-HA-ZmMAPK7*, respectively. The constructs were transformed into maize protoplast. After transformation, the protoplasts were cultured for 16 h at 28°C. Subsequently, the total protein was extracted using total plant protein extraction kit (Coolaber, Beijing, China) and immunoprecipitated with Anti-HA Affinity Beads (Smart Lifesciences, Changzhou, China) in a rotary mixer for 4 h at 4°C. Then, 1 μg HA-tagged protein mixed with 0.25, 0.5, and 1 μg His-ZmPP2C26L/His-ZmPP2C26S protein, respectively, in the presence of 50 mM ATP and 30 μl kinase buffer (20 mM HEPES [*N*-2-hydroxyethylpiperazine-*N*-2-ethane sulfonic acid], 10 mM $MgCl_2$, and 1 mM DTT, pH 7.5), and incubated for 60 min at 30°C. The reaction was stopped by adding SDS-loading buffer and separated by 12.5% phos-tag™ (Wako, Beijing, China) SDS-PAGE and normal SDS-PAGE for immunoblotting using anti-HA and anti-His antibody (ABclonal, Wuhan, China) as previously. The relative density of each band was analyzed using ImageJ software.[1]

Phenotyping of Transgenic *Arabidopsis*, Rice, and Maize Mutant

The T-DNA insertion mutant of *ZmPP2C26* homolog *ap2c1* (SALK_065126) was obtained from the *Arabidopsis* Biological Resource Center (ABRC, Columbus, OH, United States). The *35S-ZmPP2C26L-eGFP/35S-ZmPP2C26S-eGFP* construct

[1]https://imagej.en.softonic.com/

was transformed into *ap2c1* and Col-0 wild type (WT) for complementation and overexpression, respectively, by the floral-dip method (Clough and Bent, 1998). The seeds of homozygous lines were screened by 50 mg/L kanamycin on 1/2 MS plates without separation, planted in the soil, and cultured in greenhouse under optimal condition. The 4-week-old seedlings were kept under water deprivation for 3 weeks, then re-watered with a recovery time of 2 days. The untransformed WT and *ap2c1* line were used as control.

For transgenic rice, embryonic calli were isolated from the japonica rice variety *Nipponbare*, separately transformed by *35S-ZmPP2C26L* and *35S-ZmPP2C26S* using *Agrobacterium*-mediated transformation, screened on 1/2 MS plates containing 50 μg/ml hygromycin, regenerated, and identified by PCR. The seeds of homozygous lines and WT were germinated for 7 days. Then, 30 seedlings were transplanted into the rectangular plastic pots with mud and grown in greenhouse under a photoperiod of 14 h light 30°C/10 h dark at 25°C. Three-week-old seedings were subjected to drought treatment by withholding watering for 2 weeks, then re-watered for 3 days and photographed. The survival rate of every line was calculated. Before treatment, the photosynthetic rate of every line was measured using LI-6400XT portable photosynthesis system (LI-COR, Lincoln, NE, United States). The content of total chlorophyll was detected as previously described (Zhang et al., 2021). Meanwhile, 30 seedlings were transferred into plastic net pots, cultured in 20% PEG solution for 3 days, and measured their root length.

A maize *zmpp2c26* mutant generated by *Mu* transposon insertion within first exon was obtained from ChinaMu.[2] It is verified by PCR and RT-PCR. The seeds of homozygous mutant were grown in soil under 28°C under a photoperiod of 14 h light/10 h dark. The three-leaf stage seedlings were subjected to drought stress by withholding watering for 3 weeks, then re-watered for 5 days and photographed. The survival rate, root length, and root dry weight of every line were measured. Before treatment, the photosynthetic rate and the content of total chlorophyll were detected as earlier. The WT isolated from heterozygous *zmpp2c26* mutant was used as control.

Tandem Mass Tag Based Quantitative Phosphoproteomics

Rice fresh shoots of L1, S1, and Nip, as well as maize *zmpp2c26* mutant with WT were used for phosphoproteomic analysis with three biological replicates. The total protein of every line was extracted using total plant protein extraction kit (Coolaber, Beijing, China) according to the manufacturer's instruction and quantified through BCA assay (Smith et al., 1985). For each sample, 200 μg protein solution was mixed with 5 mM DTT and 10 mM iodotyrosine in the dark at room temperature for 15 min, subsequently precipitated by adding six volumes of acetone at -20°C overnight, then centrifuged at 8,000 *g* for 10 min at 4°C to collect the precipitate, and placed for 3 min at room temperature to volatilize the acetone. After removing supernatant, 100 μl of TEAB (200 mM) was added into tube to re-dissolve the protein. Subsequently, 1/50 of the sample weight of 1 mg/ml trypsin-TPCK was added into solution and digested at 37°C overnight to generate peptides. The peptides were labeled using TMT labeling kit (Thermo Scientific, Waltham, MA, United States) and used for enriching phosphopeptides using the IMAC Phosphopeptide enrichment kit (Thermo Fisher Scientific, United States) according to the manufacturer's instruction. The enriched peptides were used for LC-MS/MS scan on the Q Exactive HF and EASY-nLC 1200 system (Thermo Scientific, Waltham, MA, United States). ProteomeDiscoverer (version 2.4) was used to search raw data against the sample protein database. A global false discovery rate (FDR) was set to 0.01, and protein groups considered for quantification required at least two peptides. The detailed protocol and parameters were set as described in **Supplementary Table 2**. The amino acid sequences of differentially phosphorylated proteins were used to analyze their Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway and functional enrichment.

Statistical Analysis

All experiments were performed with three replicates. The data were showed as mean ± SD (standard deviation) and analyzed by using Student's *t*-test at $^{*}P \leq 0.05$ and $^{**}P \leq 0.01$.

RESULTS

ZmPP2C26 Is Subject to Alternative First Exon-Type as Generating Two Variants

During gene cloning, it is found that two AS variants were amplified using a pair of *ZmPP2C26* primers and defined as *ZmPP2C26L* and *ZmPP2C26S* (**Figure 1A**). Sequence alignment showed that there was 213 bp retention in the first exon of *ZmPP2C26L* compared with *ZmPP2C26S*. The splicing site is 5′-CC●●●●●●GC-3′ and neither U2 nor U12 type, indicating that it is a new AS-type AFE (**Figure 1B** and **Supplementary Figure 1A**). Protein sequence alignment showed that 71 amino acids were encoded by 213 bp nucleotides and possessed a highly conserved MAPK interaction motif (KIM motif, [K/R]-(3-4)-X(1-6)-[L/I]-X-[L/I]) (**Figure 1C** and **Supplementary Figure 1B**; Schweighofer et al., 2007), implying that ZmPP2C26L might interact with some ZmMAPK members.

ZmPP2C26 Splicing Variants Interacts With ZmMAPK3 and ZmMAPK7

To address whether ZmPP2C26L/ZmPP2C26S participated in ABA and MAPK signaling, Y2H assay was performed to determine the interaction of them with 13 ZmPYLs and 20 ZmMAPKs. The results showed that ZmPP2C26L/ZmPP2C26S did not interact with 13 ZmPYLs (**Supplementary Figure 2A**). However, on the quadruple dropout (–Leu/–Trp/–His/–Ade/) SD plates with X-α-Gal, the yeast strains cotransformed by AD-ZmPP2C26L and BD-ZmMAPK3/BD-ZmMAPK7, AD-ZmPP2C26S, and BD-ZmMAPK3, as well as positive control (i.e., AD-T and BD-53) could grow well and be stained blue

[2]http://chinamu.jaas.ac.cn/Default.html

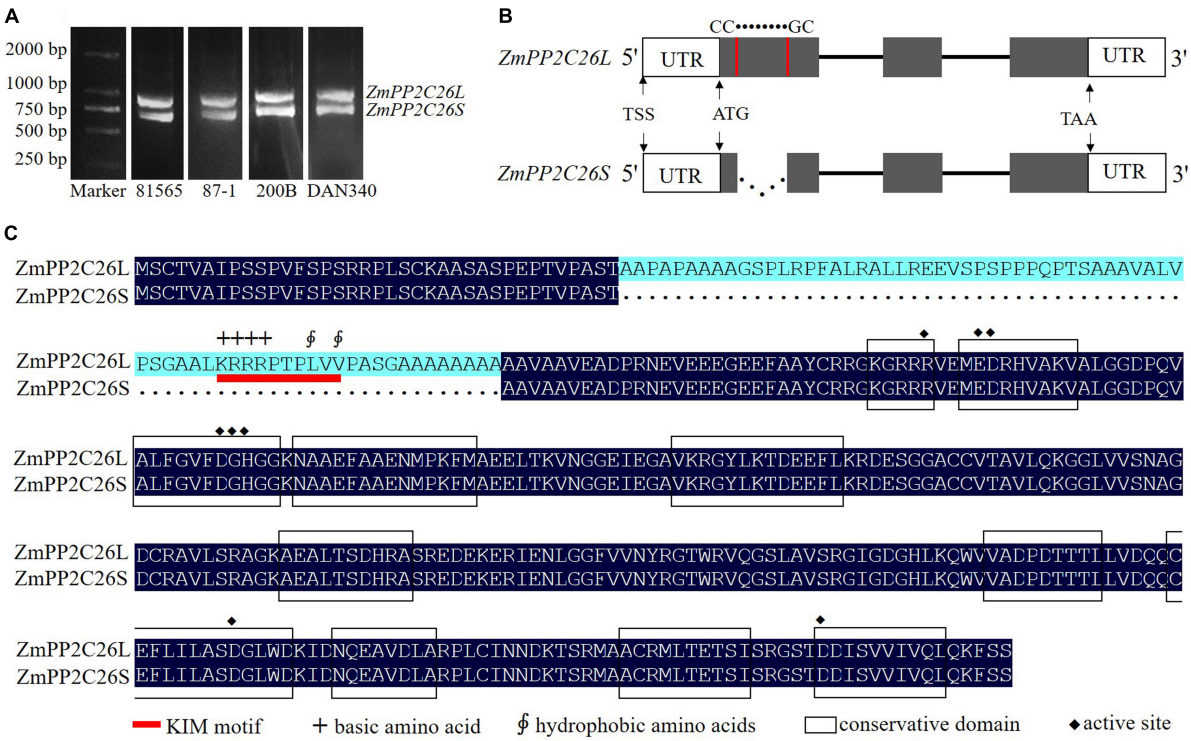


FIGURE 1 | The splicing variants of ZmPP2C26 *via* AFE type. **(A)** The ORF amplification of *ZmPP2C26* gene from maize inbred lines 81565, 87-1, 200B, and DAN340. **(B)** The alternative splicing model of *ZmPP2C26*. The gray boxes indicate exons. Lines correspond to introns; dotted line represents alternative splicing events; UTR, untranslated regions; TSS, the transcriptional start site; The splicing site is 5′-CC●●●●●GC-3′. **(C)** Protein sequence alignment of ZmPP2C26L and ZmPP2C26S. Black boxes indicate conservation domain of protein phosphatase; red line indicates KIM motif; + , basic amino acid; ϕ, hydrophobic amino acids; ◆, active site.

(**Figure 2A** and **Supplementary Figure 2B**). The GST pull-down showed that the His-ZmPP2C26L was pulled down by GST-ZmMAPK3/-ZmMAPK7, and His-ZmPP2C26S was only pulled down by GST-ZmMAPK3 (**Figure 2B**). The BiFC assay further showed that co-expression of nYFP-ZmPP2C26L and cYFP-ZmMAPK3/cYFP-ZmMAPK7 and co-expression of nYFP-ZmPP2C26S and cYFP-ZmMAPK3 in maize protoplasts could produce strong YFP fluorescence signal (**Figure 2C**). These results confirm that ZmPP2C26L physically interact with ZmMAPK3 and ZmMAPK7, but ZmPP2C26S only interacts with ZmMAPK3 *in vitro* and *in vivo*.

ZmPP2C26 Dephosphorylates ZmMAPK3 and ZmMAPK7

To test whether ZmPP2C26L/ZmPP2C26S dephosphorylates ZmMAPK3/ZmMAPK7, the dephosphorylation assay was performed *in vitro*. As shown in **Figure 3**, only phospho-ZmMAPK3/-ZmMAPK7 was detected in the absence of ZmPP2C26L or ZmPP2C26S on the phos-tag^{TM} SDS-PAGE gel. However, de-phosphorylated ZmMAPK3/ZmMAPK7 was detected when HA-ZmMAPK3/HA-ZmMAPK7 was incubated with His-ZmPP2C26L and HA-ZmMAPK3 was incubated with His-ZmPP2C26S. Notably, the phospho-ZmMAPK3/phospho-ZmMAPK7 proteins were decreased with the increase in ZmPP2C26L/ZmPP2C26S concentration. Furthermore, only dephospho-ZmPP2C26L/-ZmPP2C26S was detected using anti-His when HA-ZmMAPK3/HA-ZmMAPK7 was incubated with His-ZmPP2C26L and HA-ZmMAPK3 was incubated with His-ZmPP2C26S. These results show that ZmPP2C26L dephosphorylates ZmMAPK3 and ZmMAPK7, and ZmPP2C26S dephosphorylates ZmMAPK3, whereas ZmPP2C26L and ZmPP2C26S cannot be phosphorylated by ZmMAPK3 and ZmMAPK7. Likewise, the expression of *ZmMAPK3* and *ZmMAPK7* gene was dramatically evaluated elevated by drought stress in maize (**Supplementary Figure 3**).

Subcellular Localization of ZmPP2C26L/ZmPP2C26S

To determine the subcellular localization of ZmPP2C26L/ZmPP2C26S, the ORF of them was fused with *eGFP* under the control of the *35S* promoter and introduced into *N. benthamiana* leaves for transient expression. Confocal laser scanning microscopy showed that the eGFP fluorescence signal was observed in both the cytoplasm and the nucleus from the leaf infiltrated by the empty *eGFP* vector (control) and *ZmPP2C26S-eGFP* vector, but in the chloroplast and nucleus from the leaf infiltrated by the *ZmPP2C26L-eGFP* vector (**Figure 4A**). Co-localization also showed that ZmPP2C26L co-localized with ZmMAPK3/ZmMAPK7 in the nucleus, and

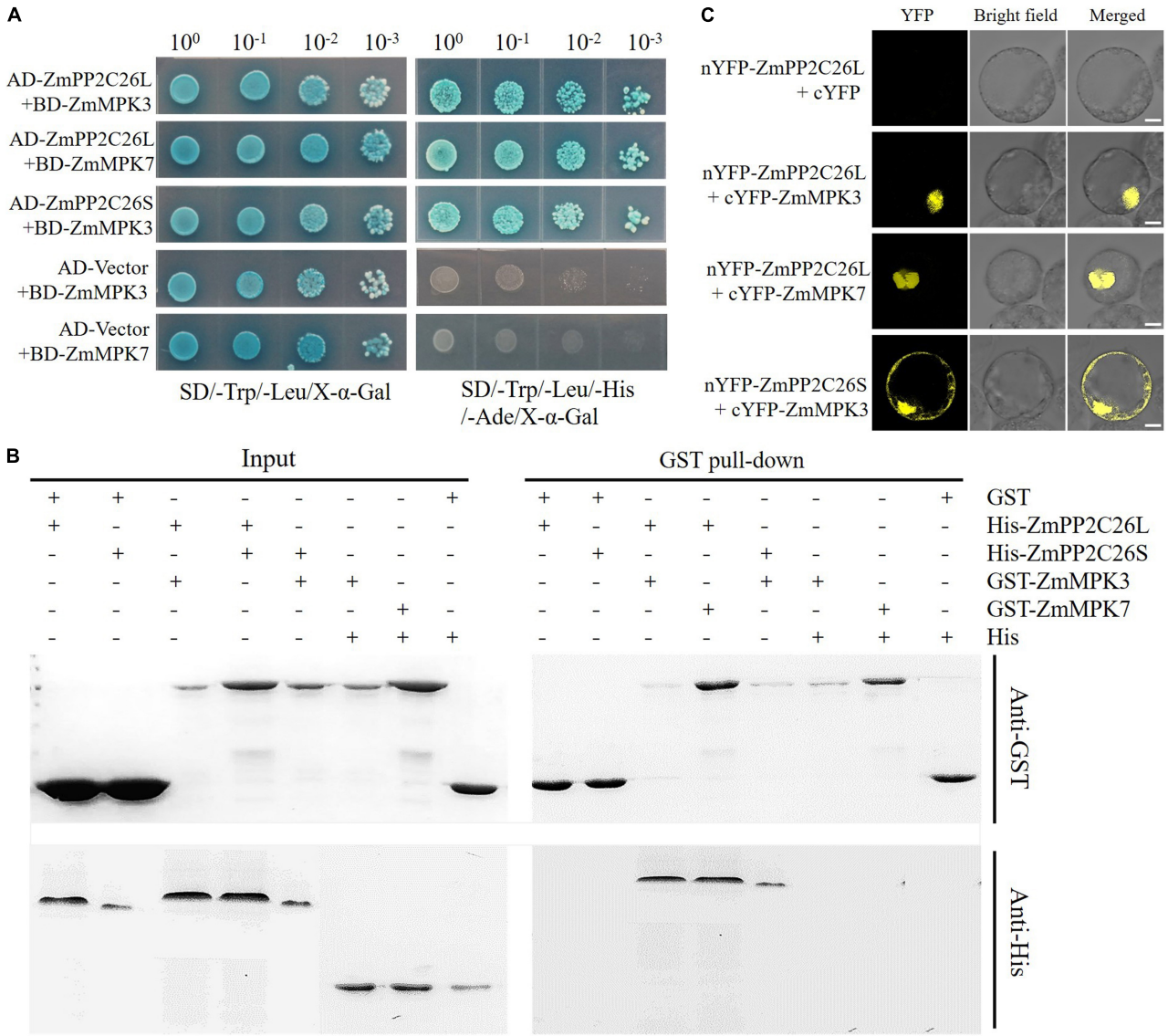


FIGURE 2 | The interaction between ZmPP2C26L/ZmPP2C26S and ZmMAPK3/ZmMAPK7. **(A)** Y2H assay. **(B)** GST pull-down assay. **(C)** BiFC assay. Scale bar = 20 μm.

ZmPP2C26S co-localized with ZmMAPK3 in the cytoplasm and nucleus (**Figure 4B**).

ZmPP2C26S Exhibits Higher Phosphatase Activity Than ZmPP2C26L

Since ZmPP2C26 is a member of PP2C clade B and belongs to PP family (Wang et al., 2018), the phosphatase activity of ZmPP2C26L and ZmPP2C26S was detected using the chromogenic substrate pNPP *in vitro* phosphatase assays. The PP could catalyze pNPP to produce p-nitrophenol with an absorbance value of OD_{405}. After adding the purified His-ZmPP2C26L and His-ZmPP2C26S into the reaction solution, there was a strong OD_{405} values, suggesting that both of them had phosphatase activity. Moreover, the relative phosphatase activity of ZmPP2C26S was significantly higher than ZmPP2C26L (**Supplementary Figure 4**).

ZmPP2C26L and ZmPP2C26S Negatively Regulate Drought Tolerance

The results of RT-qPCR showed that the expression of *ZmPP2C26* and *ZmPP2C26L* was significantly downregulated by drought stress in drought-tolerant lines 81565 and 87-1 and upregulated in drought-sensitive lines 200B and DAN340 (**Supplementary Figure 5**), which was consist with our previous study (Lu et al., 2020). The data imply that *ZmPP2C26* plays a crucial role in regulating drought tolerance.

Hence, *ZmPP2C26L* and *ZmPP2C26S* was complemented and overexpressed in *Arabidopsis*-mutant *ap2c1* (AT2G30020,

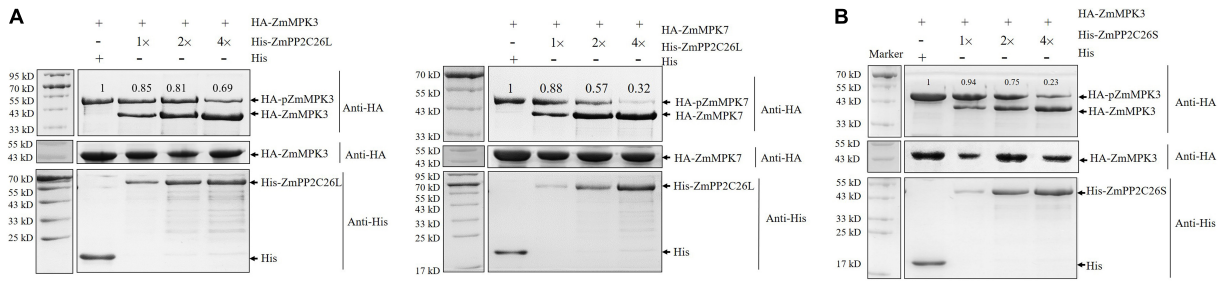


FIGURE 3 | The dephosphorylation assay *in vitro*. **(A)** ZmPP2C26L dephosphorylates ZmMAPK3 and ZmMAPK7. **(B)** ZmPP2C26S dephosphorylates ZmMAPK3. The phospho-ZmMAPK3/-ZmMAPK7 with HA tag was extracted from maize protoplasts, incubated with His-ZmPP2C26L/His-ZmPP2C26S, separated by 12.5% phos-tag SDS-PAGE gel (top panel) or normal SDS-PAGE gel (middle panel), and detected by using an anti-HA or anti-His antibody. Different amounts of His-ZmPP2C26L/His-ZmPP2C26S were used to dephosphorylate ZmMAPK3 and ZmMAPK7. 1×, 2×, and 4× represent 0.25, 0.5, and 1.0 μg purified His-ZmPP2C26L/-ZmPP2C26S, respectively. + and - denote the presence and absence of the protein in each sample, respectively. The relative intensity of the protein bands was measured using ImageJ, and the lane without His-ZmPP2C26L/His-ZmPP2C26S was set to 1.00.

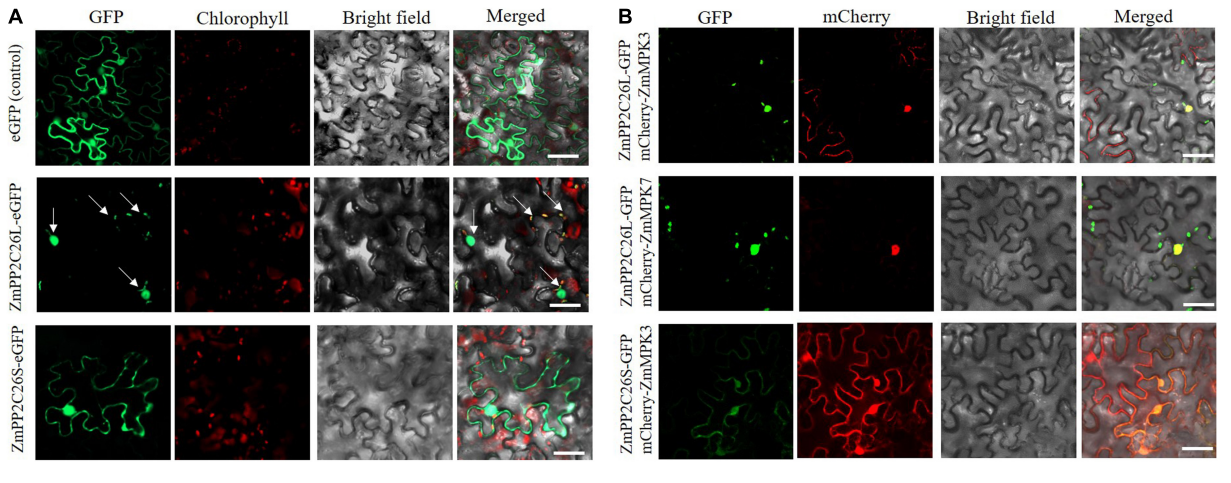


FIGURE 4 | Subcellular localization and co-localization. **(A)** The localization of ZmPP2C26L and ZmPP2C26S. **(B)** Co-localization of ZmPP2C26L and ZmMAPK3/ZmMAPK7 and of ZmPP2C26S and ZmMAPK3. *ZmPP2C26L* and *ZmPP2C26S* were fused with *eGFP*. *ZmMAPK3* and *ZmMAPK7* were separately fused with mCherry. A 4-week-old tobacco (*Nicotiana benthamiana*) seedling was used for infiltrating using *Agro*. Tumefaciens cells harboring above plasmids. Scale bar = 50 μm.

an ortholog of *ZmPP2C26*) and WT. The complementation and overexpression of *ZmPP2C26L*/*ZmPP2C26S* increased the drought sensitivity of transgenic lines compared with *ap2c1* and WT (**Supplementary Figure 6**). Subsequently, they were further overexpressed in rice (*Nipponbare*). Two *ZmPP2C26L*-overexpressing lines (i.e., L1 and L2) and two *ZmPP2C26S*-overexpressing lines (i.e., S1 and S2) were subjected to drought stress by withholding watering. As shown in **Figure 5**, after 2 weeks of drought stress, transgenic lines showed enhancement of drought sensitivity, but WT seedlings were slightly blasted. The survival rates of L1, L2, S1, and S2 lines were 44.5, 54.3, 5.2, and 6.0%, respectively, which were significantly lower than WT (90.8%). Furthermore, the root length of transgenic lines was significantly shorter than WT under drought stress condition. Under optimal condition, the chlorophyll content and photosynthetic rates of L1 and L2 lines were significantly lower than WT. S1 and S2 lines showed no significant difference compared with WT.

Meanwhile, a maize mutant of *zmpp2c26* was used for drought tolerance test. As shown in **Figure 6**, the Mu transposon insertion of *zmpp2c26* resulted in knockout of *ZmPP2C26* identified by RT-PCR. The *zmpp2c26* exhibited drought-tolerant phenotype compared with WT. After drought stress, the survival rate of *zmpp2c26* was 87.5% and that of the WT was only 12.5%. The root length and root dry weight of *zmpp2c26* were also significantly higher than WT. The chlorophyll content and photosynthetic rate of *zmpp2c26* mutant were significantly higher than WT. The above results suggest that *ZmPP2C26* negatively modulates drought tolerance and the *ZmPP2C26S* variant was more sensitive to drought than *ZmPP2C26L*.

ZmPP2C26 Alters Protein Phosphorylation Level

To uncover the global effects of ZmPP2C26 on protein phosphorylation, we performed TMT-based quantitative

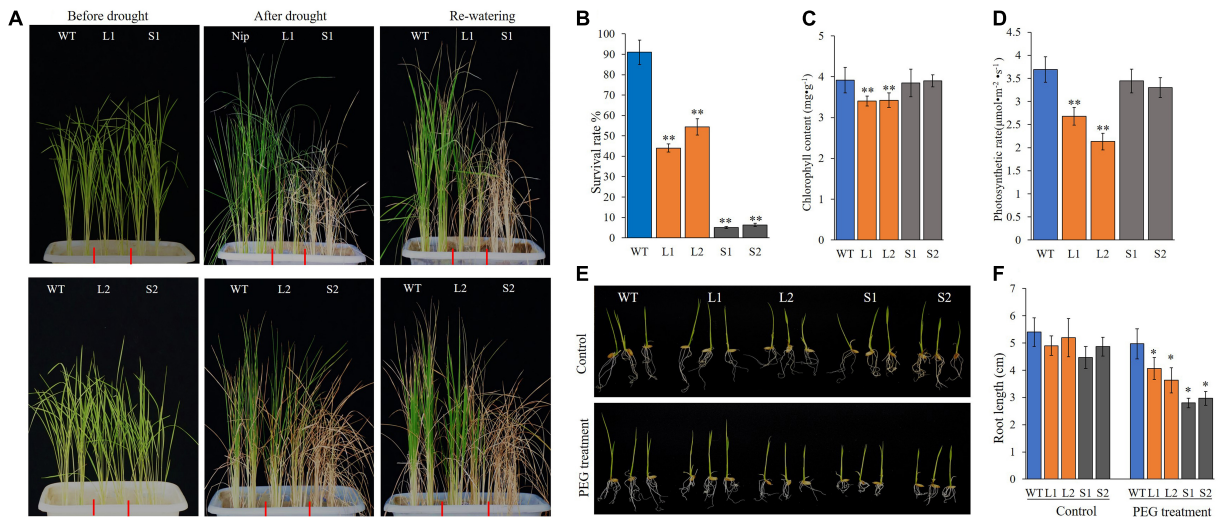


FIGURE 5 | Phenotyping of transgenic rice lines under drought stress. **(A)** Wilting phenotype of every line. Three-week-old seedings were subjected to drought treatment by withholding watering for 2 weeks, then re-watered for 3 days and photographed. **(B)** The survival rate of every line after treatment. **(C,D)** Chlorophyll content and photosynthetic rate of every line before treatment. **(E)** Root phenotype of every line. After germination, seedlings were transferred into plastic net pots and cultured in 20% PEG solution for 3 days. **(F)** Root length. WT, wild type; L1 and L2, *ZmPP2C26L*-overexpressing lines; S1 and S2, *ZmPP2C26S*-overexpressing lines. $^{*}P < 0.05$; $^{**}P < 0.01$.

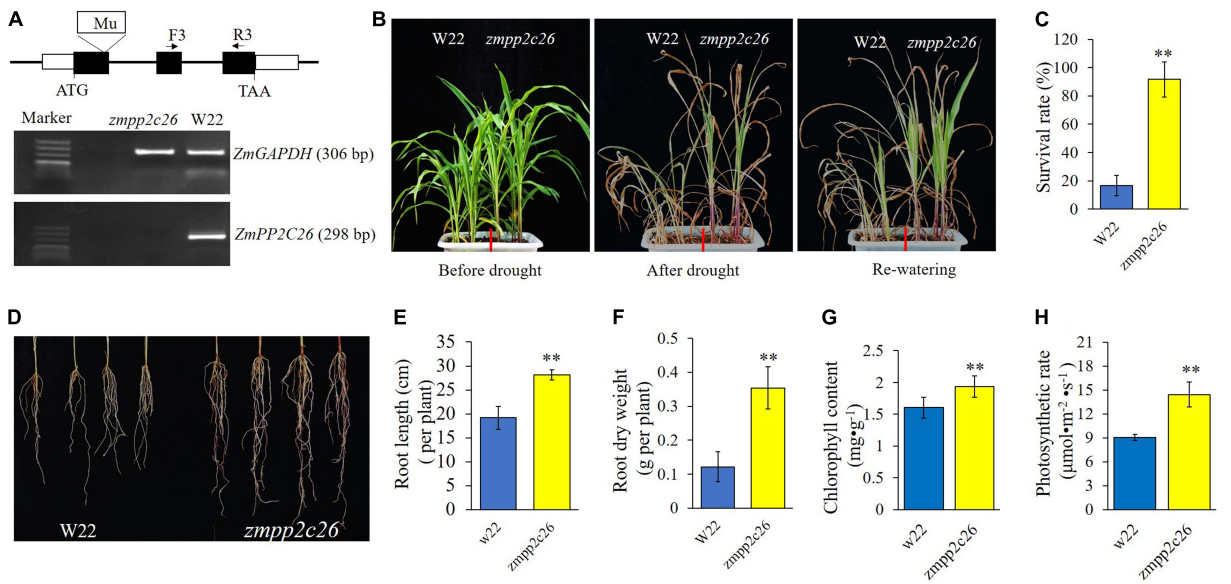


FIGURE 6 | Phenotyping of maize *zmpp2c26* under drought stress. **(A)** Schematic diagram of Mu transposon in *ZmPP2C26* gene and transcriptional detection *ZmPP2C26* by RT-PCR. Black boxes represent exons; F and R, primer position in *ZmPP2C26,* and the sequences of them were listed in **Supplementary Table 1**. A 298 bp fragment of *ZmPP2C26* was amplified from *zmpp2c26* cDNA using F and R primers. A 306 bp fragment of *ZmGAPDH* was amplified and used as reference. W22, the wild type isolated from heterozygous *zmpp2c26* mutant, was used as control. **(B)** Wilting phenotype of every line. The three-leaf stage seedlings were subjected to drought stress *via* withholding watering for 3 weeks, then re-watered for 5 days and photographed. **(C)** The survival rate of every line after treatment. **(D)** Root phenotype of every line. **(E,F)** Root length and root dry weight of every line. **(G,H)** Chlorophyll content and photosynthetic rate of every line before treatment. $^{**}P < 0.01$.

phosphoproteomic analysis in the *ZmPP2C26L*-/*ZmPP2C26S*-overexpressing lines and WT. A total of 3,516 phosphosites and 2,877 phosphopeptides derived from 1,532 phosphoproteins with >1.5-fold change ($P < 0.05$) were found to form transgenic lines compared with WT (**Supplementary Table 3**). After evaluation of raw data, clean data with high reliability were screened and used for KEGG analysis. Compared with WT, 48 and 93 phosphopetides were significantly upregulated and downregulated in the *ZmPP2C26L*-overexpression lines, respectively. In contrast, 55 and 28 phosphopetides

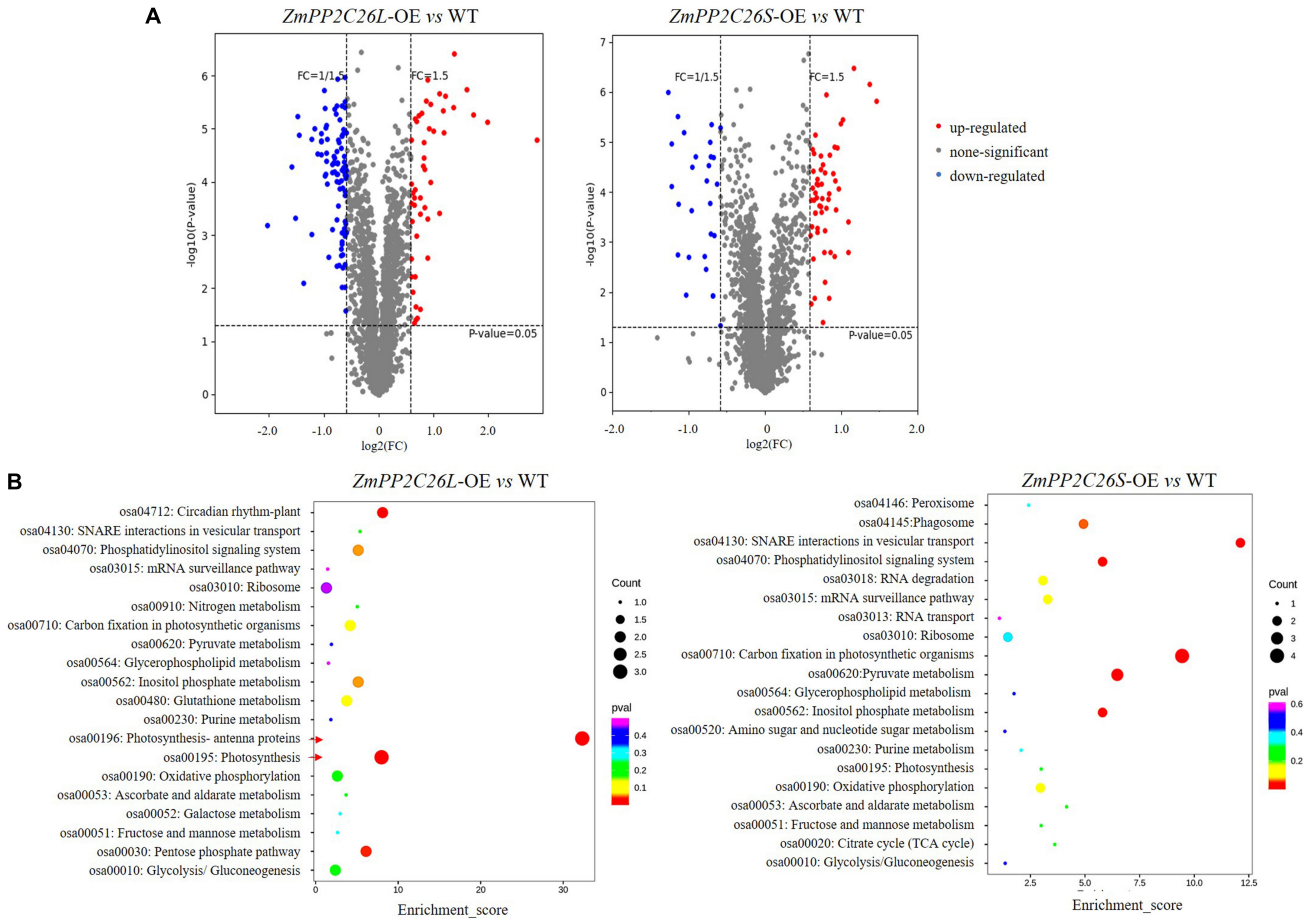


FIGURE 7 | Phosphoproteomics of transgenic rice identifies a set of *ZmPP2C26*-affected phosphoproteins. **(A)** Phosphopeptide abundance for *ZmPP2C26L*- and *ZmPP2C26S*-overexpressing lines compared with WT. **(B)** The GO and KEGG analysis of *ZmPP2C26L*- and *ZmPP2C26S*-affected phosphorylation proteins.

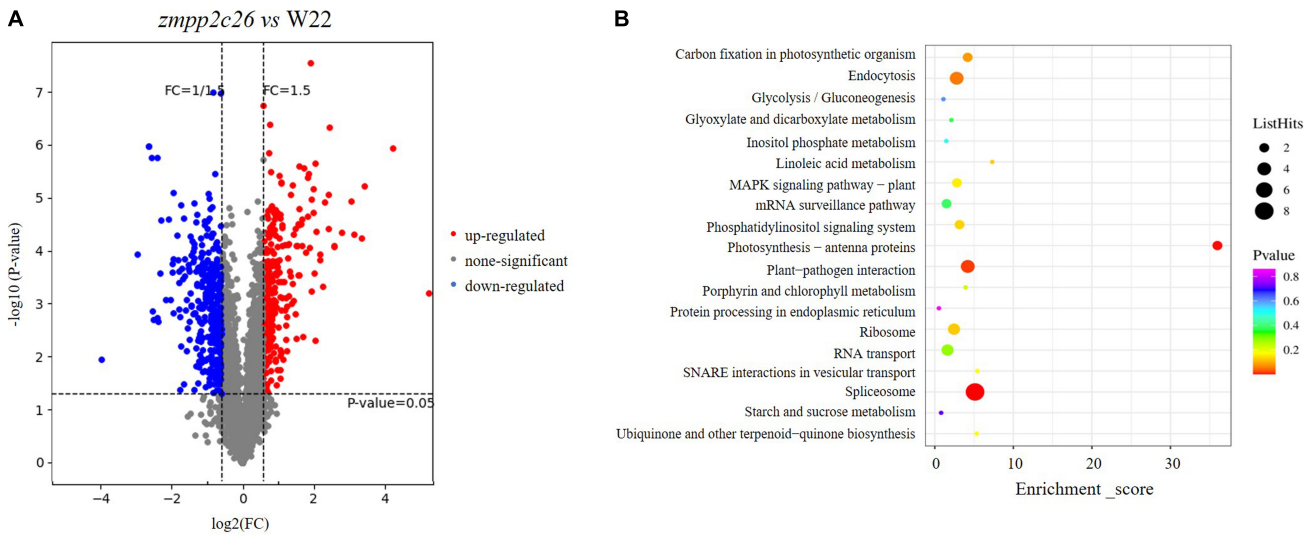


FIGURE 8 | Phosphoproteomics of *zmpp2c26* mutant identifies *ZmPP2C26*-affected phosphoproteins. **(A)** Phosphopeptide abundance for *zmpp2c26* compared with WT. **(B)** The KEGG analysis of *ZmPP2C26*-affected upregulated phosphorylation proteins.

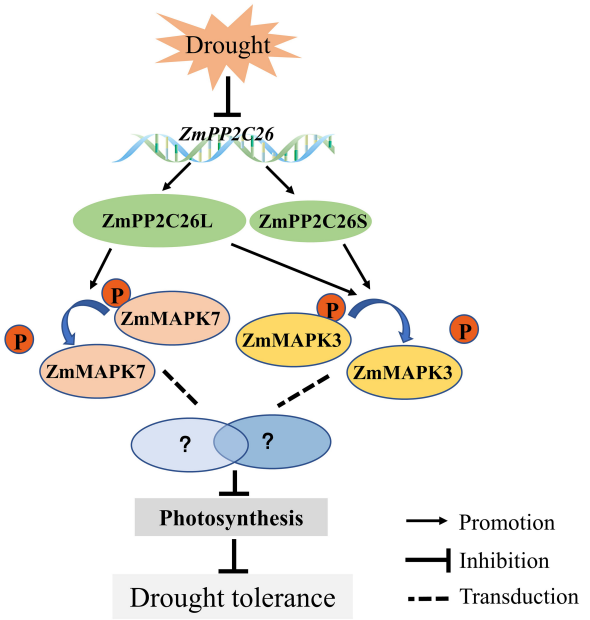


FIGURE 9 | The proposed model depicting ZmPP2C26 alternative splice variants negatively regulate drought tolerance.

were significantly upregulated and downregulated in the *ZmPP2C26S*-overexpression lines, respectively (**Supplementary Table 4**). KEGG enrichment pathway analysis showed that the *ZmPP2C26L*-affected phosphorylation proteins predominately enriched in photosynthesis and *ZmPP2C26S*-affected phosphorylation proteins were associated with carbon fixation in photosynthesis and pyruvate metabolism (**Figure 7**). Among them, 8 upregulated and 12 downregulated phosphoproteins were shared by *ZmPP2C26L-OE* and *ZmPP2C26S-OE* lines compared with WT, and the aquaporin PIP2-7 showed the highly increased phosphorylation level (**Supplementary Table 5**).

Similarly, we analyzed the phosphoproteomic difference of *zmpp2c26* compared with WT. The results showed that a total of 671 unique phosphorylation proteins with >1.5-fold change ($P < 0.05$) were identified in *zmpp2c26* (**Supplementary Table 6**). Also, 324 and 347 phosphoproteins were upregulated and downregulated in *zmpp2c26*, respectively. As ZmPP2C26 is a PP, the phosphorylation level of its targets should be increased in the *zmpp2c26* plants. The KEGG pathway analysis indicated that the upregulated phosphoproteins were enriched in several pathways but mainly involved in photosynthesis (**Figure 8**). Notably, the phosphorylation level of ZmMAPK3 (accession number: B4FN55) was significantly upregulated in *zmpp2c26* mutant (**Supplementary Table 6**), indicating that ZmPP2C26 involved in photosynthesis and MAPK-mediated signaling pathway.

DISCUSSION

Alternative splicing is a critical gene post-transcriptional regulation mechanism for plants in response to surrounding stress, growth, and development (Reddy et al., 2013; Qin et al., 2017). In plants, more than 60% intron-containing genes undergo AS to produce a vast repertoire mRNA isoform (Marquez et al., 2012; Reddy et al., 2013; Iniguez et al., 2017; Jiang et al., 2017). U2 and U12 are two typical intron retention types of AS (Sharp and Burge, 1997; Reddy et al., 2013). In this study, *ZmPP2C26*, a maize PP2C clade B member, is found to undergo AS, producing *ZmPP2C26L* and *ZmPP2C26S* in different maize germplasm lines (**Figure 1**). Its splicing site is 5′-CC●●●●●●GC-3′ (**Figure 1B** and **Supplementary Figure 1A**), which is AFE-type AS and recently also found in *Arabidopsis* (Wang et al., 2015; Qin et al., 2017; Zhu et al., 2017).

The clade B PP2Cs can recognize MAPK phosphatases and interact with MAPK3, MAPK4, or MAPK6 to regulate stomatal aperture, seed germination, and stress response (Brock et al., 2010; Umbrasaite et al., 2010; Sidonskaya et al., 2016; Shubchynskyy et al., 2017). It has been reported that AP2C1 dephosphorylates CBL-interacting protein kinase (CIPK) to regulate K^+ deficiency response (Singh et al., 2018). In this study, it is confirmed that 213 bp is retained in first exon of *ZmPP2C26L* and spliced in *ZmPP2C26S*, and encodes 71 amino acids including a conserved KIM motif (**Figure 1C** and **Supplementary Figure 1B**). However, the ZmPP2C26L physically interacts with ZmMAPK3 and ZmMAPK7 in the nucleus, but ZmPP2C26S only interacts with ZmMAPK3 in both the cytoplasm and the nucleus (**Figures 2**, **4B** and **Supplementary Figure 2B**), which means that the molecular mechanism of these two variants may be different. Considering ZmMAPK3 and ZmMAPK7 can be phosphorylated by ZmMKK10 (Chang et al., 2017), they are transiently expressed in maize protoplast, immunoprecipitated, and used for dephosphorylation catalyzed by ZmPP2C26. It is expectedly found that ZmPP2C26L dephosphorylates ZmMAPK3 and ZmMAPK7, and ZmPP2C26S dephosphorylates ZmMAPK3, while ZmMAPK3 and ZmMAPK7 cannot phosphorylate ZmPP2C26L and ZmPP2C26S (**Figure 3**). The ZmPP2C26L/ZmPP2C26S did not interact with 13 PYLs (**Supplementary Figure 2**), suggesting that ZmPP2C26L and ZmPP2C26S may act on MAPK signaling to regulate downstream signaling in ABA-independent way. Interestingly, ZmPP2C26S lacks KIM motif but interacts with ZmMAPK3, indicating that KIM motif is unnecessary for interaction between ZmMAPKs and other protein. Meanwhile, *ZmMAPK3* and *ZmMPK7* were found to be upregulated by drought stress (**Supplementary Figure 3**) and involved in multiple stresses response in previous reports (Wang J. et al., 2010, Wang Y. L. et al., 2010; Wang et al., 2011; Liu et al., 2013). Their homologs, i.e., AtMPK3 and AtMPK6, regulated salt and cold tolerance and hypoxia signaling in *Arabidopsis* (Li et al., 2017; Yan et al., 2021; Zhou et al., 2021). These studies imply that ZmPP2C26 functions in plants stress response acting on ZmMAPK3/ZmMAPK7.

Since GC content flaking splicing sites is pretty high for primer PP2C26S-F (**Supplementary Figure 1**), fragment of *ZmPP2C26S* was not amplified using primers PP2C26S-F/PP2C26S-R for RT-qPCR, so we detected the relative expression level of *ZmPP2C26* containing the transcript of *ZmPP2C26L* and *ZmPP2C26S* instead of *ZmPP2C26S*. Under drought stress, the expression

of *ZmPP2C26* and *ZmPP2C26L* was significantly downregulated by drought stress in drought-tolerant maize and upregulated in drought-sensitive maize (**Supplementary Figure 5**). The possible cue is due to the inhibition of its promoter by drought (Lu et al., 2020). The results indicate that *ZmPP2C26* may play a negative role in regulating drought tolerance. Hence, they are functional validated *via* ectopic expressing in *Arabidopsis*, rice, as well as phenotyping maize *zmpp2c26* mutant (**Figures 5**, **6** and **Supplementary Figure 6**). But the complementation of *ZmPP2C26L* in *Arabidopsis ap2c1* mutant partially restores drought-sensitive phenotype and overexpression of *ZmPP2C26L* in *Arabidopsis* WT and rice exhibit much lower hypersensitive to drought tolerance compared with *ZmPP2C26S*-overexpression lines (**Figure 5** and **Supplementary Figure 6**). These results may be explained by ZmPP2C26S possesses higher activity compared with ZmPP2C26L (**Supplementary Figure 4**). The mutant of *ZmPP2C26* gene increases chlorophyll and photosynthesis rate (**Figures 6G,H**), which could be explained by affecting the phosphorylation of proteins involved in photosynthesis (**Figure 8**). However, the downstream of ZmMAPK3- and ZmMAPK7-meditated by ZmPP2C26 needs to be explored in further study. In transgenic rice, ZmPP2C26L affects proteins phosphorylation directly enriched in photosynthesis, but ZmPP2C26S regulates proteins associated with carbon fixation in photosynthesis and pyruvate metabolism (**Figure 7**), and the phosphorylation of aquaporin PIP2-7 contributes for drought tolerance influenced by ZmPP2C26L and ZmPP2C26S (Jang et al., 2014; Fan et al., 2015).

CONCLUSION

We characterized a clade B type 2C PP, i.e., ZmPP2C26, that undergoes untypical AS to generate two isoforms. ZmPP2C26 directly dephosphorylates ZmMAPK3 and ZmMAPK7 to negatively regulate drought stress and photosynthesis activity (**Figure 9**). This study provides insights into understanding the mechanism of PP2C in response to abiotic stress.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

FL performed most of the experiments and drafted the manuscript. FL and HY designed the experiments. YP, YC, JQ, and FS carried out the experiments and analyzed the data. LZ and QY assisted in experimental operations. XZ, YL, and FF reviewed the manuscript. WL and HY edited the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was supported by the Sichuan Science and Technology Program (2020YJ0353), the Open Program of National Key Laboratory of Wheat and Maize Crop Science (SKL2021KF09), and the Key Research and Development Project of Chendu (2021-YF05-02024-SN).

ACKNOWLEDGMENTS

We thank Dongtao Ren at China Agricultural University for providing pGBKT7-*ZmMAPK* vectors (BD-*ZmMAPKs*). We also thank the Shanghai Luming Biological Technology Co., Ltd. (Shanghai, China) for providing proteomics services.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.851531/full#supplementary-material

REFERENCES

Barford, D., Das, A. K., and Egloff, M. P. (1998). The structure and mechanism of protein phosphatases: insights into catalysis and regulation. *Annu. Rev. Biophys. Biomol. Struct.* 27, 133–164. doi: 10.1146/annurev.biophys.27.1.133

Brock, A. K., Willmann, R., Kolb, D., Grefen, L., Lajunen, H. M., Bethke, G., et al. (2010). The *Arabidopsis* mitogen-activated protein kinase phosphatase PP2C5 affects seed germination, stomatal aperture, and abscisic acid-inducible gene expression. *Plant Physiol.* 153, 1098–1111. doi: 10.1104/pp.110.156109

Calixto, C. P. G., Guo, W., James, A. B., Tzioutziou, N. A., Entizne, J. C., Panter, P. E., et al. (2018). Rapid and dynamic alternative splicing impacts the *Arabidopsis* cold response transcriptome. *Plant Cell* 30, 1424–1444. doi: 10.1105/tpc.18.00177

Chang, Y., Yang, H., Ren, D., and Li, Y. (2017). Activation of ZmMKK10, a maize mitogen-activated protein kinase kinase, induces ethylene-dependent cell death. *Plant Sci* 264, 129–137. doi: 10.1016/j.plantsci.2017.09.012

Chu, M., Chen, P., Meng, S., Xu, P., and Lan, W. (2021). The *Arabidopsis* phosphatase PP2C49 negatively regulates salt tolerance through inhibition of AtHKT1;1. *J. Integr. Plant Biol.* 63, 528–542. doi: 10.1111/jipb.13008

Clough, S. J., and Bent, A. F. (1998). Floral dip: a simplified method for agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16, 735–743. doi: 10.1046/j.1365-313x.1998.00343.x

Cohen, P. (1989). The structure and regulation of protein phosphatases. *Annu. Rev. Biochem.* 58, 453–508.

Fan, W., Li, J., Jia, J., Wang, F., Cao, C., Hu, J., et al. (2015). Pyrabactin regulates root hydraulic properties in maize seedlings by affecting PIP aquaporins in a phosphorylation-dependent manner. *Plant Physiol. Biochem.* 94, 28–34. doi: 10.1016/j.plaphy.2015.05.005

Farquharson, K. L. (2016). Metabolic Signaling regulates alternative splicing during photomorphogenesis. *Plant Cell* 28:2697. doi: 10.1105/tpc.16.00842

Gu, J., Xia, Z., Luo, Y., Jiang, X., Qian, B., Xie, H., et al. (2018). Spliceosomal protein U1A is involved in alternative splicing and salt stress tolerance in *Arabidopsis thaliana*. *Nucleic Acids Res.* 46, 1777–1792. doi: 10.1093/nar/gkx1229

Han, L., Li, J., Jin, M., and Su, Y. (2018). Functional analysis of a type 2C protein phosphatase gene from *Ammopiptanthus mongolicus*. *Gene* 653, 29–42. doi: 10.1016/j.gene.2018.02.015

He, Z., Wu, J., Sun, X., and Dai, M. (2019). The maize clade A PP2C phosphatases play critical roles in multiple abiotic stress responses. *Int. J. Mol. Sci.* 20:3573. doi: 10.3390/ijms20143573

Iniguez, L. P., Ramírez, M., Barbazuk, W. B., and Hernandez, G. (2017). Identification and analysis of alternative splicing events in *Phaseolus vulgaris* and *Glycine max*. *BMC Genomics* 18:650. doi: 10.1186/s12864-017-4054-2

Jang, H. Y., Rhee, J., Carlson, J. E., and Ahn, S. J. (2014). The *Camelina* aquaporin CsPIP2;1 is regulated by phosphorylation at Ser273, but not at Ser277, of the C-terminus and is involved in salt- and drought-stress responses. *J. Plant Physiol.* 171, 1401–1412. doi: 10.1016/j.jplph.2014.06.009

Jiang, J., Liu, X., Liu, C., Liu, G., Li, S., and Wang, L. (2017). Integrating omics and alternative splicing reveals insights into grape response to high temperature. *Plant Physiol.* 173, 1502–1518. doi: 10.1104/pp.16.01305

Kelemen, O., Convertini, P., Zhang, Z., Wen, Y., Shen, M., Falaleeva, M., et al. (2013). Function of alternative splicing. *Gene* 514, 1–30. doi: 10.1002/9783527678679.dg00350

Komatsu, K., Suzuki, N., Kuwamura, M., Nishikawa, Y., Nakatani, M., Ohtawa, H., et al. (2013). Group A PP2Cs evolved in land plants as key regulators of intrinsic desiccation tolerance. *Nat. Commun.* 4:2219. doi: 10.1038/ncomms3219

Li, H., Ding, Y., Shi, Y., Zhang, X., Zhang, S., Gong, Z., et al. (2017). MPK3- and MPK6-mediated ICE1 phosphorylation negatively regulates ICE1 stability and freezing tolerance in *Arabidopsis*. *Dev. Cell* 43, 630–642. doi: 10.1016/j.devcel.2017.09.025

Li, Y., Chang, Y., Zhao, C., Yang, H., and Ren, D. (2016). Expression of the inactive ZmMEK1 induces salicylic acid accumulation and salicylic acid-dependent leaf senescence. *J. Integr. Plant Biol.* 58, 724–736. doi: 10.1111/jipb.12465

Li, Y., Yang, J., Shang, X., Lv, W., Xia, C., Wang, C., et al. (2019). SKIP regulates environmental fitness and floral transition by forming two distinct complexes in *Arabidopsis*. *New Phytol.* 224, 321–335. doi: 10.1111/nph.15990

Liang, B., Sun, Y., Wang, J., Zheng, Y., Zhang, W., Xu, Y., et al. (2021). Tomato protein phosphatase 2C influences the onset of fruit ripening and fruit glossiness. *J. Exp. Bot.* 72, 2403–2418. doi: 10.1093/jxb/eraa593

Liu, X., Singh, S. K., Patra, B., Liu, Y., Wang, B., Wang, J., et al. (2021). Protein phosphatase NtPP2C2b and MAP kinase NtMPK4 act in concert to modulate nicotine biosynthesis. *J. Exp. Bot.* 72, 1661–1676. doi: 10.1093/jxb/eraa568

Liu, Y., Wang, L., Zhang, D., and Li, D. (2013). Expression analysis of segmentally duplicated ZmMPK3-1 and ZmMPK3-2 genes in maize. *Plant Mol. Biol. Rep.* 31, 457–463. doi: 10.1007/s11105-012-0489-4

Liu, Z., Qin, J., Tian, X., Xu, S., Wang, Y., Li, H., et al. (2018). Global profiling of alternative splicing landscape responsive to drought, heat and their combination in wheat (*Triticum aestivum* L.). *Plant Biotechnol. J.* 16, 714–726. doi: 10.1111/pbi.12822

Lobell, D. B., Roberts, M. J., Schlenker, W., Braun, N., Little, B. B., Rejesus, R. M., et al. (2014). Greater sensitivity to drought accompanies maize yield increase in the U.S. Midwest. *Science* 344, 516–519. doi: 10.1126/science.1251423

Lu, F., Wang, K., Yan, L., Peng, Y., Qu, J., Wu, J., et al. (2020). Isolation and characterization of maize ZmPP2C26 gene promoter in drought-response. *Physiol. Mol. Biol. Plants* 26, 2189–2197. doi: 10.1007/s12298-020-00910-2

Ma, Y., Szostkiewicz, I., Korte, A., Moes, D., Yang, Y., Christmann, A., et al. (2009). Regulators of PP2C phosphatase activity function as abscisic acid sensors. *Science* 324, 1064–1068. doi: 10.1126/science.1172408

Marquez, Y., Brown, J. W., Simpson, C., Barta, A., and Kalyna, M. (2012). Transcriptome survey reveals increased complexity of the alternative splicing landscape in *Arabidopsis*. *Genome Res.* 22, 1184–1195. doi: 10.1101/gr.134106.111

Martin, G., Marquez, Y., Mantica, F., Duque, P., and Irimia, M. (2021). Alternative splicing landscapes in *Arabidopsis thaliana* across tissues and stress conditions highlight major functional differences with animals. *Genome Biol.* 22:35. doi: 10.1186/s13059-020-02258-y

Nishimura, N., Tsuchiya, W., Moresco, J. J., Hayashi, Y., Satoh, K., Kaiwa, N., et al. (2018). Control of seed dormancy and germination by DOG1-AHG1 PP2C phosphatase complex via binding to heme. *Nat. Commun.* 9:2132. doi: 10.1038/s41467-018-04437-9

Park, S. Y., Fung, P., Nishimura, N., Jensen, D. R., Fujii, H., Zhao, Y., et al. (2009). Abscisic acid inhibits type 2C protein phosphatases via the PYR/PYL family of START proteins. *Science* 324, 1068–1071. doi: 10.1126/science.1173041

Qin, Z., Wu, J., Geng, S., Feng, N., Chen, F., Kong, X., et al. (2017). Regulation of FT splicing by an endogenous cue in temperate grasses. *Nat. Commun.* 8:14320. doi: 10.1038/ncomms14320

Reddy, A. S., Marquez, Y., Kalyna, M., and Barta, A. (2013). Complexity of the alternative splicing landscape in plants. *Plant Cell* 25, 3657–3683. doi: 10.1105/tpc.113.117523

Sah, R. P., Chakraborty, M., Prasad, K., Pandit, M., Tudu, V. K., Chakravarty, M. K., et al. (2020). Impact of water deficit stress in maize: phenology and yield components. *Sci. Rep.* 10:2944. doi: 10.1038/s41598-020-59689-7

Schweighofer, A., Hirt, H., and Meskiene, I. (2004). Plant PP2C phosphatases: emerging functions in stress signaling. *Trends Plant Sci.* 9, 236–243. doi: 10.1016/j.tplants.2004.03.007

Schweighofer, A., Kazanaviciute, V., Scheikl, E., Teige, M., Doczi, R., Hirt, H., et al. (2007). The PP2C-type phosphatase AP2C1, which negatively regulates MPK4 and MPK6, modulates innate immunity, jasmonic acid, and ethylene levels in *Arabidopsis*. *Plant Cell* 19, 2213–2224. doi: 10.1105/tpc.106.049585

Sharp, P. A., and Burge, C. B. (1997). Classification of introns: U2-type or U12-type. *Cell* 91, 875–879. doi: 10.1016/s0092-8674(00)80479-1

Shi, Y. (2009). Serine/threonine phosphatases: mechanism through structure. *Cell* 139, 468–484. doi: 10.1016/j.cell.2009.10.006

Shubchynskyy, V., Boniecka, J., Schweighofer, A., Simulis, J., Kvederaviciute, K., Stumpe, M., et al. (2017). Protein phosphatase AP2C1 negatively regulates basal resistance and defense responses to *Pseudomonas syringae*. *J. Exp. Bot.* 68, 1169–1183. doi: 10.1093/jxb/erw485

Sidonskaya, E., Schweighofer, A., Shubchynskyy, V., Kammerhofer, N., Hofmann, J., Wieczorek, K., et al. (2016). Plant resistance against the parasitic nematode *Heterodera schachtii* is mediated by MPK3 and MPK6 kinases, which are controlled by the MAPK phosphatase AP2C1 in *Arabidopsis*. *J. Exp. Bot.* 67, 107–118. doi: 10.1093/jxb/erv440

Singh, A., Giri, J., Kapoor, S., Tyagi, A. K., and Pandey, G. K. (2010). Protein phosphatase complement in rice: genome-wide identification and transcriptional analysis under abiotic stress conditions and reproductive development. *BMC Genomics* 11:435. doi: 10.1186/1471-2164-11-435

Singh, A., Yadav, A. K., Kaur, K., Sanyal, S. K., Jha, S. K., Fernandes, J. L., et al. (2018). A protein phosphatase 2C, AP2C1, interacts with and negatively regulates the function of CIPK9 under potassium-deficient conditions in *Arabidopsis*. *J. Exp. Bot.* 69, 4003–4015. doi: 10.1093/jxb/ery182

Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., et al. (1985). Measurement of protein using bicinchoninic acid. *Anal. Biochem.* 150, 76–85. doi: 10.1016/0003-2697(85)90442-7

Umbrasaite, J., Schweighofer, A., Kazanaviciute, V., Magyar, Z., Ayatollahi, Z., Unterwurzacher, V., et al. (2010). MAPK phosphatase AP2C3 induces ectopic proliferation of epidermal cells leading to stomata development in *Arabidopsis*. *PLoS One* 5:e15357. doi: 10.1371/journal.pone.0015357

Wang, J., Ding, H., Zhang, A., Ma, F., Cao, J., and Jiang, M. (2010). A novel mitogen-activated protein kinase gene in maize (*Zea mays*), ZmMPK3, is involved in response to diverse environmental cues. *J. Integr. Plant Biol.* 52, 442–452. doi: 10.1111/j.1744-7909.2010.00906.x

Wang, J. X., Jiang, M. Y., Ma, F. F., and Ding, H. D. (2011). Expression characterization and function analysis of ZmMPK7 from *Zea mays* seedlings. *J. Nanjing Agric. Univ.* 34, 68–73.

Wang, Y. G., Fu, F. L., Yu, H. Q., Hu, T., Zhang, Y. Y., Tao, Y., et al. (2018). Interaction network of core ABA signaling components in maize. *Plant Mol. Biol.* 96, 245–263. doi: 10.1007/s11103-017-0692-7

Wang, Y. L., Zong, X. J., and Li, D. Q. (2010). Analysis of response to photosynthetic characteristic and antioxidant enzyme activities in transgenic tobacco transformed with zmmpk7 under salt stress. *J. Nuclear Agric. Sci.* 5, 1086–1092.

Wang, Z., Ji, H., Yuan, B., Wang, S., Su, C., Yao, B., et al. (2015). ABA signaling is fine-tuned by antagonistic HAB1 variants. *Nat. Commun.* 6:8138. doi: 10.1038/ncomms9138

Xia, C., Gong, Y., Chong, K., and Xu, Y. (2021). Phosphatase OsPP2C27 directly dephosphorylates OsMAPK3 and OsbHLH002 to negatively regulate cold tolerance in rice. *Plant Cell Environ.* 44, 491–505. doi: 10.1111/pce.13938

Xiang, Y., Sun, X., Gao, S., Qin, F., and Dai, M. (2017). Deletion of an endoplasmic reticulum stress response element in a ZmPP2C-A gene facilitates drought tolerance of maize seedlings. *Mol. Plant* 10, 456–469. doi: 10.1016/j.molp.2016.10.003

Xue, T., Wang, D., Zhang, S., Ehlting, J., Ni, F., Jakab, S., et al. (2008). Genome-wide and expression analysis of protein phosphatase 2C in rice and *Arabidopsis*. *BMC Genomics* 9:550. doi: 10.1186/1471-2164-9-550

Yan, Z., Wang, J., Wang, F., Xie, C., Lv, B., Yu, Z., et al. (2021). MPK3/6-induced degradation of ARR1/10/12 promotes salt tolerance in *Arabidopsis*. *EMBO Rep.* 22:e52457. doi: 10.15252/embr.202152457

Yang, Z., Yang, J., Wang, Y., Wang, F., Mao, W., He, Q., et al. (2020). PROTEIN PHOSPHATASE95 regulates phosphate homeostasis by affecting phosphate transporter trafficking in rice. *Plant Cell* 32, 740–757. doi: 10.1105/tpc.19.00685

Yu, J., Miao, J., Zhang, Z., Xiong, H., Zhu, X., Sun, X., et al. (2018). Alternative splicing of OsLG3b controls grain length and yield in japonica rice. *Plant Biotechnol. J.* 16, 1667–1678. doi: 10.1111/pbi.12903

Zhang, J., Wan, L., Igathinathane, C., Zhang, Z., Guo, Y., Sun, D., et al. (2021). Spatiotemporal heterogeneity of chlorophyll content and fluorescence response within rice (*Oryza sativa* L.) canopies under different nitrogen treatments. *Front. Plant Sci.* 12:645977. doi: 10.3389/fpls.2021.645977

Zhou, Y., Zhou, D. M., Yu, W. W., Shi, L. L., Zhang, Y., Lai, Y. X., et al. (2021). Phosphatidic acid modulates MPK3- and MPK6-mediated hypoxia signaling in *Arabidopsis*. *Plant Cell* 34, 889–909. doi: 10.1093/plcell/koab289

Zhu, F. Y., Chen, M. X., Ye, N. H., Shi, L., Ma, K. L., Yang, J. F., et al. (2017). Proteogenomic analysis reveals alternative splicing and translation as part of the abscisic acid response in *Arabidopsis* seedlings. *Plant J.* 91, 518–533. doi: 10.1111/tpj.13571

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 15 April 2022
doi: 10.3389/fpls.2022.874434

Transcriptomics and Antioxidant Analysis of Two Chinese Chestnut (*Castanea mollissima* BL.) Varieties Provides New Insights Into the Mechanisms of Resistance to Gall Wasp *Dryocosmus kuriphilus* Infestation

Cancan Zhu[1], Wu Wang[1], Yu Chen[1], Yuqiang Zhao[1], Shijie Zhang[1], Fenghou Shi[2], Muhammad Khalil-Ur-Rehman[3] and Niels J. Nieuwenhuizen[4]*

[1] Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing, China, [2] College of Forestry, Nanjing Forestry University, Nanjing, China, [3] Department of Horticultural Sciences, The Islamia University of Bahawalpur, Bahawalpur, Pakistan, [4] The New Zealand Institute for Plant and Food Research Ltd., Auckland, New Zealand

OPEN ACCESS

Edited by:
Mukesh Jain, Jawaharlal Nehru University, India

Reviewed by:
Fang Tang, Chinese Academy of Forestry, China
Melissa Hamner Mageroy, Norwegian Institute of Bioeconomy Research, Norway

***Correspondence:**
Wu Wang
2017204015@njau.edu.cn

Specialty section:
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

Received: *12 February 2022*
Accepted: *21 March 2022*
Published: *15 April 2022*

Citation:
Zhu C, Wang W, Chen Y, Zhao Y, Zhang S, Shi F, Khalil-Ur-Rehman M and Nieuwenhuizen NJ (2022) Transcriptomics and Antioxidant Analysis of Two Chinese Chestnut (Castanea mollissima BL.) Varieties Provides New Insights Into the Mechanisms of Resistance to Gall Wasp Dryocosmus kuriphilus Infestation. Front. Plant Sci. 13:874434. doi: 10.3389/fpls.2022.874434

Chinese chestnut is a popular fruit tree with a high nutritional value of its nuts, which can suffer from infestation by the chestnut gall wasp *Dryocosmus kuriphilus* (GWDK) that results in gall formation and resultant loss of production and profitability. The physiological and molecular mechanisms of GWDK resistance found in certain genotypes currently remains elusive. To gain new insights into this phenomenon, a series of RNA-Seq integrated with metabolomic profiling experiments were executed to investigate the chemical and transcriptional differences in response to GWDK infestation in two contrasting chestnut varieties grown in China (the susceptible "HongLi," HL and the partially resistant "Shuhe_Wuyingli," SW). Three time points were selected for comparison: The initiation stage (A), growth stage (B), and maturation stage (C). Results showed that concentrations of hydrogen peroxide (H_2O_2) and the activities of peroxidase (POD) and superoxide dismutase (SOD) enzyme were elevated in the resistant SW leaves compared with those in HL leaves at all three developmental stages, while catalase (CAT) and polyphenol oxidase (PPO) activities were mostly higher in HL leaves. RNA-Seq transcriptomic analyses of HL and SW leaves revealed that various metabolic pathways involved in GWDK stress responses, such as plant hormone signal transduction, MAPK signaling, and the peroxisome pathway, were enriched in the contrasting samples. Moreover, the weighted gene co-expression network analysis (WGCNA) of differentially expressed genes in the POD pathway combined with transcription factors (TFs) indicated that the expression of TF members of bHLH, WRKY, NAC, and MYB family positively correlated with POD pathway gene expression. The TFs *CmbHLH130 (EVM0032437)*, *CmWRKY31 (EVM0017000)*, *CmNAC50 (EVM0000033),* and *CmPHL12 (EVM0007330)* were identified as putative TFs that participate in the regulation of insect-induced plant enzyme activities in chestnut, which may contribute to GWDK resistance in SW. Expression levels of 8 random differentially expressed

genes (DEGs) were furthermore selected to perform quantitative reverse transcription PCR (qRT-PCR) to validate the accuracy of the RNA-Seq-derived expression patterns. This study guides the functional analyses of further candidate genes and mechanisms important for GWDK resistance in chestnuts in the future as well as can help in identifying the master transcriptional regulators and important enzyme steps that support major insect defense pathways in chestnut.

Keywords: Chinese chestnut, *Castanea mollissima*, gall wasp, *Dryocosmus kuriphilus*, RNA-seq, POD, WGCNA, MAPK

INTRODUCTION

Castanea mollisima Blume, also known as Chinese chestnut, belongs to the Fagaceae family and is an economically valuable tree in China. More than 300 wild and cultivated varieties have been reported (Nishio et al., 2021). Most of Chinese chestnut species were reported as highly resistant to abiotic stresses while being susceptible to herbivory infestations (Acquadro et al., 2020); specialized metabolites and antioxidants have been reported to modulate interactions of plants with their abiotic environment (Chapman et al., 2019). The Asian chestnut gall wasp, *Dryocosmus kuriphilus* Yasumatu (Hymenoptera: Cynipidae), is one of the main pests of the genus *Castanea* and can form a large number of galls on plant leaves or buds. *D. kuriphilus* can also harm the chestnut, cause the leaves to be deformed, and may result in the death of whole branches. As a result from the infestation, the fruiting branches and the vegetative branches cannot grow normally, which eventually leads to a progressive loss of the photosynthetic biomass (Zhu et al., 2019). This pest is also considered as a significant problem in other di- and monocot species such as in grape (Granett et al., 2001), wheat (Smiley et al., 2004), blueberry (Sampson et al., 2002), and rice (Way et al., 1991). *D. kuriphilus* infestation in chestnut is difficult to prevent and control because most of its life cycle occurs inside the galls.

Chestnut galls induced by *D. kuriphilus* are among the most fascinating structures found in fruit crops, as they are the product of sophisticated plant-insect interactions (Ronquist and Liljeblad, 2001; Stone and Schönrogge, 2003). Insect-induced galls are an extreme example of an extended phenotype (Martinson et al., 2021), in which the galls are produced and maintained by plant genes and enzymes, while the initiation, development, and morphology of the galls are rather controlled by the insect through manipulation of the cells of the plant (Raman, 2011; Korgaonkar et al., 2021). To date, the genes/pathways of plants that respond to *D. kuriphilus* infestations and that affect the galls productions need to be further clarified.

External factors that affect the chestnut galls productions include the phenology of the host, the phenology of the parasitoids, and environmental conditions such as temperature. Apart from these factors (Geng et al., 2015), the genetics of the plant host may also play an important role and affects the *D. kuriphilus* feeding. In previous studies, a chestnut variety "Shuhe_Wuyingli," SW, was identified, which is partially resistant to gall infestations (Geng et al., 2015; Zhu et al., 2019). Morphological and cytochemical differences between susceptible and resistant varieties have been surveyed to clarify the nature of the resistance on the chestnut trees to the gall wasps (Matsui and Torikata, 1970). In this study, the GWDK-susceptible cultivar of "HongLi," HL, and GWDK-partially resistant "SW" chestnut were used to further explore the molecular mechanism of GWDK-resistance through physiological and temporal RNA-Seq analysis.

Plants have been interacting with herbivores for millions of years and evolved sophisticated strategies to protect themselves against attacks by biotic stressors. Previous studies have demonstrated that herbivory feeding commonly induces jasmonic acid (JA)-dependent defense pathways in tea (Jing et al., 2021), soybean (Selig et al., 2016), and tomato (Bosch et al., 2014), where jasmonate can contribute to protecting plants against herbivores. The induction of JA in plants that experience insect damage can lead to increased production of secondary defense metabolites such as terpene compounds or increase the polyphenol oxidase (PPO) activity to maintaining considerable resistance against insect pests through a variety of possible mechanisms. Terpenes are important compounds involved in the interactions between plants and insects as well between plants and pathogens, such as in tea leaves, where the homoterpene *(E)*-4,8-dimethyl-1,3,7-non-atriene (DMNT) induced the accumulation of JA and thus promoted resistance in neighboring uninfested plants to herbivorous insects. In this process, activation of the terpene biosynthesis pathway genes in response to the herbivory infestations is highly beneficial (Jing et al., 2021).

Another significant pest and stress defense mechanisms that has been reported is the antioxidant defense system. When plants are infested by insects, induction of antioxidant enzymes is part of the early responses to the herbivory signals. This includes antioxidant enzymes and low-molecular antioxidants, such as superoxide dismutase (SOD), peroxidase (POD), catalase (CAT), phenylalanine ammonia lyase (PAL), and hydrogen peroxide (H_2O_2) (War et al., 2012). SOD can convert superoxide radicals (O^{-2}) into H_2O_2 (Liochev and Fridovich, 2007); POD and CAT play major roles in reducing/scavenging H_2O_2 to water using different substrates as electron donors. The rapid detoxification of both O^{-2} and H_2O_2 are essential processes to prevent excessive oxidative damage (Shi et al., 2014). Several studies have demonstrated that the activity of antioxidant enzymes was positive correlated with plant tolerance to biotic stresses, including pests and fungi (Chojak-Koźniewska et al., 2018).

The Chinese chestnut genome sequence was recently released (Wang et al., 2020), and some studies on the transcriptional regulation of chestnut secondary metabolites and QTL mapping of chestnut disease resistance (Barakat et al., 2012) have been completed. However, a comprehensive transcriptome analyses to identify gene expression and expressional variations of genes of the Chinese chestnut in response to GWDK infestation using comprehensive transcriptome resources are still lacking. Deep sequencing by RNA-Seq-based approaches can enable researchers to generate an unprecedented global view of the transcriptome changes and to identify the signaling pathways responsible for plant defense to various biotic stresses. Recently, several studies have used RNA-Seq assays to quantify changes in the transcriptome upon herbivory infections, and a small subset of novel defense-response candidate genes were also identified (Huang et al., 2020; Trujillo-Moya et al., 2020).

To determine the gene expression profile changes in chestnuts leaves upon gall wasp infestation and to identify developmental stage-specific genes and pathways involved in this interaction, in this study, temporal RNA-Seq was combined with antioxidant enzyme metabolomic profiling. This approach included an in-depth characterization of the genes and pathways altered in susceptible vs. partially resistant chestnut-GWKD interactions. This dataset provides new insights into the chestnut-GWKD interactions and identifies candidate genes/enzymes in plant responses to GWDK infestation that could contribute to the molecular-aided breeding of gall wasp–resistant chestnuts and other woody plants.

MATERIALS AND METHODS

Materials and Treatments

Four-year-old Chinese chestnut (*Castanea mollissima* Blume) plants of two varieties, namely, "HL" and "SW," were planted in Chinese chestnut germplasm resources unit, Nanjing, Jiangsu Province, China, under same growth conditions (i.e., temperature, humidity, and light). The variety "HL" was susceptible to GWDK infestation, while "SW" was partially resistant to GWDK infestation as manifested by the appearance of less and smaller galls on the resistant plants following the methodology by Acquadro et al. (2020). Buds with galls were harvested from 12 plants at 3 time points starting from budburst from April 7 until April 26 [Initiation stage (A) at April 7, growth (B) at April 15, and maturation (C) at April 26] in 2019 (**Figure 1**). Each sample contained three biological replicates. Bud samples of chestnut were harvested from 12 plants (four plants per replicate) and used for the enzyme activity test, RNA extraction, and RNA-sequencing. The fresh buds were sampled and frozen in liquid nitrogen and stored at −80°C until further use.

Determination of Hydrogen Peroxide Content

The H_2O_2 content (n mol g^{-1} fresh weight) from chestnut buds was assessed using the xylenol orange method, as described by Loreto and Velikova (2001). The absorbance was measured using spectrophotometer at 560 nm, and the experiment was repeat at least 3 times.

Antioxidant Enzyme Activities Assay

All the antioxidant enzymes described below were extracted by homogenizing 0.3 g of frozen buds in phosphate buffer solution (pH 7.4 and 4% polyvinylpolypyrrolidone) and then centrifugated at 16,000 × *g* for 30 min at 4°C; the supernatants were collected and used for antioxidant enzyme activities assays. The activities of SOD, CAT, PPO, and POD were assayed according to the methodologies proposed by Beers and Sizer (1952), Giannopolitis and Ries (1977), Wissemann and Lee (1980) and Amako et al. (1994), respectively, with some modifications. The activities of CAT, POD, PPO, and SOD were determined from the measurement of absorbance at 240, 420, 410, and 560 nm, respectively. The results of aforementioned enzymes were expressed as U mg^{-1} FW; three biological replicates were utilized to make each of the measurements.

RNA Preparation, Transcriptome Sequencing, and Functional Annotation

A total of 100 mg frozen bud samples for 2 species of 3 development stages were ground in a mortar under liquid nitrogen, and the resultant powder was used for RNA extraction with the TRIzol® extraction reagent (Invitrogen, United States.). Nano Drop and Agilent 2100 bioanalyzer (Agilent Technologies, United States) were used to assess the quality of the RNA sample, and high-quality RNA samples with an A260/A280 ratio of 1.8–2.1 were used as template material for RNA-Seq and quantitative reverse transcription PCR (qRT-PCR) analyses. In total, 18 libraries were prepared (i.e., 3 biological replicates per sample) and sequenced using the SE 50 using Illumina Hiseq2500 platform (BGI, Shenzhen, China); the raw reads were filtered by removal of adapter and low-quality sequences using SOAPnuke v1.5.2. Data are available under NCBI SRA accession PRJNA791965. The clean reads were mapped to Chinese chestnut genome[1] using HISAT2 (v2.0.4); fragments per kilobase per million reads (FPKM) values were applied to calculate the gene expression levels. The DESeq (V1.4.5) package was used to identify the differentially expressed genes (DEGs); false discovery rate (FDR) value $\leq$ 0.05 and $|\log2\text{Fold Change}| \geq 1$ were set for significant gene expression differences between two samples. To gain insights into the changes in phenotype, the enrichment analysis of the DEGs was performed with gene ontology (GO)[2] and Kyoto Encyclopedia of Genes and Genomes (KEGG[3]) annotation. Annotations of some DEGs also referred to the genome of hardwood[4]. The significant levels of the terms and pathways were corrected using a rigorous threshold (*q*-Value $\leq$ 0.05) *via* Bonferroni correction.

[1]https://ngdc.cncb.ac.cn/search/?dbId=gwhandq=GWHANWH00000000
[2]http://www.geneontology.org/
[3]https://www.kegg.jp/
[4]https://www.hardwoodgenomics.org/

FIGURE 1 | Physicochemical changes during "HongLi (HL)" and "Shuhe_Wuyingli (SW)" bud development. The average number of galls per tree were analyzed. Mean ± standard error (n = 12). Scale bar = 1 cm.

Candidate Differentially Expressed Genes Validation by Quantitative Reverse Transcription-PCR

Expression pattern verification of 8 randomly selected DEGs from RNA-Seq was performed using qRT-PCR, as described previously (Zhu et al., 2019). Complementary DNA (cDNA) was transcribed using the Takara Prime Script TM-RT PCR reagent Kit (Takara, Japan) according to the manufacturer's instructions. Specific primers (**Supplementary Table 1**) for qRT-PCR were designed using the Primer Premier 5 software. *Actin* gene was selected as internal reference in this study. qRT-PCR was conducted using the ABI7500 RT-PCR system according to the manufacturer's instructions (Applied Biosystems, Foster City, CA, United States). All reactions and non-template controls were performed in triplicate. Relative transcription levels were calculated using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). Data are expressed as mean ± standard error (SE) (n = 3).

Correlation Network Construction and Visualization

Candidate genes associated with the GWDK infestation were identified based on the correlation of transcript levels with antioxidant enzymes activities during berry development among the 2 varieties. Weighted gene coexpression network analysis (WGCNA) was conducted to get the gene coexpression network based on total 29,886 DEGs, Visualization of the gene coexpression network was performed using Cytoscape (version 3.8.2) (Vasaikar et al., 2018). Correlation network was performed using the OmicStudio tools[5].

[5]https://www.omicstudio.cn/tool

Statistical Analysis

Figures were made using GraphPad prism 8.0 (San Diego, CA, United States.). All the analyzed data were expressed as means ± standard error (SE). SPSS Version 17.0 (SPSS Inc. Chicago, IL, U.S.A.) and Excel software were used to perform statistical analyses. A one-way ANOVA with a Duncan's t-test was used to evaluate the significant differences ($p \leq 0.05$). TBtools (v 0.665) software were applied for heatmaps.

RESULTS

Basic Physicochemical Parameters of Bud Development and Galls Formation

The basic physiological indicators of the different development stages of gall formation by *D. kuriphilus* are shown in **Figure 1**. For "HL," the galls are smaller at the initiation stage (A), while during the bud development, the galls become larger and more visible (B). At the maturation stage (C), galls in the buds showed a darker red appearance, while in the buds of "SW," galls were more difficult to spot during all the development stages, and the average number of galls in the "HL" was significantly higher than "SW" (7.5 times higher).

Hydrogen Peroxide Content and Antioxidant Enzyme Activities of "HongLi" and "Shuhe_Wuyingli" Buds

Hydrogen peroxide content is an important indicator of oxidative damage. Therefore, the H_2O_2 content was investigated during the 3 development stages of "HL" and "SW" buds. At the initiation stage A, the H_2O_2 content in buds of the 2 experiment species was similar, while in the growth and maturation stage (B and C), increasing H_2O_2 was observed compared with that in the

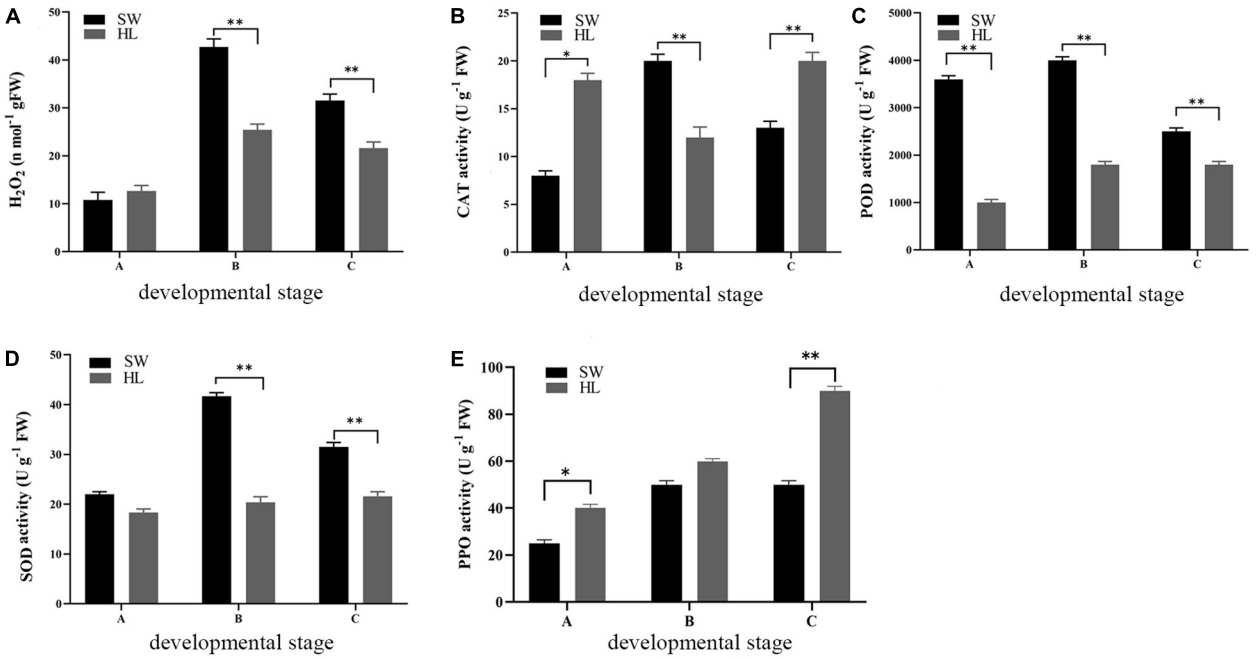


FIGURE 2 | Hydrogen peroxide (H_2O_2) content **(E)** and antioxidant enzyme activities of "HL" and "SW" buds during infestation by gall wasp *Dryocosmus kuriphilus* (GWDK) gall wasps [**(A)** polyphenol oxidase (PPO); **(B)** catalase (CAT); **(C)** peroxidase (POD); and **(D)** superoxide dismutase (SOD)]. Asterisks indicate a significant difference (*$P < 0.05$, **$P < 0.01$) between 'HL' and 'SW'.

initiation stage. The maximum content of H_2O_2 (42.7 n mol g^{-1} FW) was observed in the "SW" buds at the growth stage (B), and the content of H_2O_2 was significant higher in "SW" buds at the growth and maturation stages compared with "HL" (**Figure 2A**).

The antioxidant enzymes system, which includes CAT and POD, PPO, and SOD, play significant roles in preventing or alleviating the herbivory damage of plants resulting from reactive oxygen. These enzyme activities were therefore measured in "HL" and "SW" buds to evaluate the relative antioxidant levels. As shown in **Figures 2B–E**, at the initiation stage, PPO and CAT activity was higher in "HL" buds, while in contrast, higher POD and SOD activities were observed in "SW" buds. At the growth and maturation stages (B and C), the activity of CAT, POD, and SOD was generally higher in "SW" buds compared with that in "HL." The only exception was the CAT activity at stage C, which was higher in HL. In contrast to CAT, POD, and SOD, a higher PPO activity was observed in "HL" buds during all three stages, which was statistically significant for stages A and C. Overall, in the majority of cases, the four enzyme activities were lowest at the initially time (i.e., stage A) and increased rapidly toward stage B and remained stable or increase slightly for rest of the examination period. The higher values of H_2O_2 content and antioxidant enzyme activities were generally observed in the resistant "SW" buds, as compared with "HL" buds, apart from the PPO activity. The values of these parameters almost showed an increasing to decreasing trend during the development of both "HL" and "SW" buds.

TABLE 1 | Statistical results of the RNA-Seq data.

Classification	Maximum	Minimum	Average
Number of clean bases (Gb)	1.21	1.19	1.2
Number of clean reads (M)	24.1	23.87	23.98
Number of reads aligned (M)	18.07	16.99	17.51
Percentage of reads aligned (%)	74.96	71.01	73.01
Number of unique mapping reads (M)	17.24	15.26	16.58
Percentage of unique mapping reads (%)	71.97	64.95	68.82
Q30 (%)	92.12	91.12	91.61

Q30 indicates the percentage of data/bases with a phred quality score > 30. Phred quality score > 30 means > 99.9% base call accuracy.

Libraries Construction and Sequencing

The samples of the two chestnut varieties at three developmental stages (i.e., initiation, growth, and maturation) were subjected to RNA-Seq transcriptomics analysis. High-throughput RNA-Seq generated between 1.19 and 1.21 Gb of clean base reads from each library (**Table 1**). After stringent filtering of low-quality sequences, on average of 23.87 M clean reads were obtained with quality scores > Q30, which represented > 91.12% of the reads that were subsequently mapped to the reference genome (see text footnote 1). The aligned ratios of the reads ranged from 71.01% to 74.96% among the eighteen libraries, and 64.95–71.97% of the reads were uniquely mapped (**Table 1**). These data represented a high sequencing depth and quality sufficient for further transcriptomics analysis.

The expression levels of all the sequenced samples were then calculated *via* the FPKM method and subjected to principle

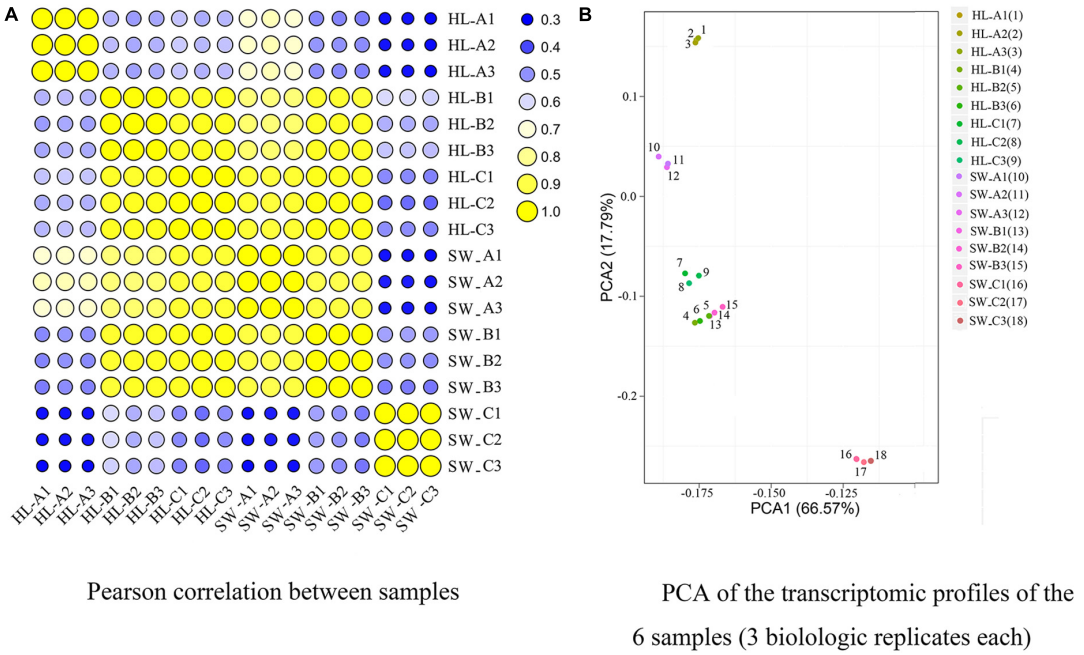


FIGURE 3 | Pearson's correlation matrix **(A)** and principle component analysis (PCA) **(B)** of differentially expressed genes (DEGs) were performed on three biological replicates of each sample sets to evaluate correlations and variance between samples. "HL" and "SW," "HongLi" and "Shuhe_Wuyingli" cultivars, respectively; A/B/C: bud infestation stage infestation/growth/maturation.

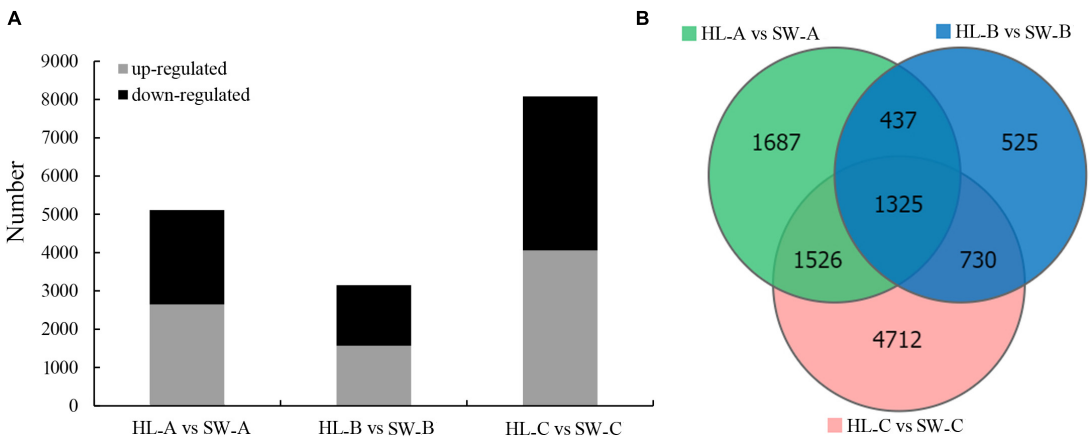


FIGURE 4 | Differential gene expression in chestnut buds in the presence of the GWDK in resistant/susceptible cultivars. **(A)** The number of upregulated and downregulated DEGs (log_2 (fold change) ≥ 1 and false discovery rate (FDR) < 0.05). **(B)** Quantity statistics Venn diagrams of differentially expressed genes among the three sampling groups (log_2 (fold change) ≥ 1 and FDR < 0.05).

component analysis (PCA) and Pearson correlation analysis (**Figure 3A**). All samples showed a degree of correlation (the within-group correlation was at least $R^2 > 0.962$). PCA of the expression profiles of the 18 libraries [i.e., 6 samples and 3 biological replicates (**Figure 3B**)] revealed that the samples from the three different stages and 2 cultivars could be clearly separated, confirming that between-stage variation was relatively high and that the three stages showed distinct global expression patterns, which were suitable for further analysis. PCA data (PCA1 variance explained 66.57% and PCA2 17.79%) represent 84.36% of the variance across the two dimensions. Overall samples HL and SW stages B clustered most closely together and both clustered closely with HL stage C, while the other combinations were more distant.

Identification of Differentially Expressed Genes

After the estimating of the gene expression levels through the FPKM method, 12,793 protein-coding DEGs were identified

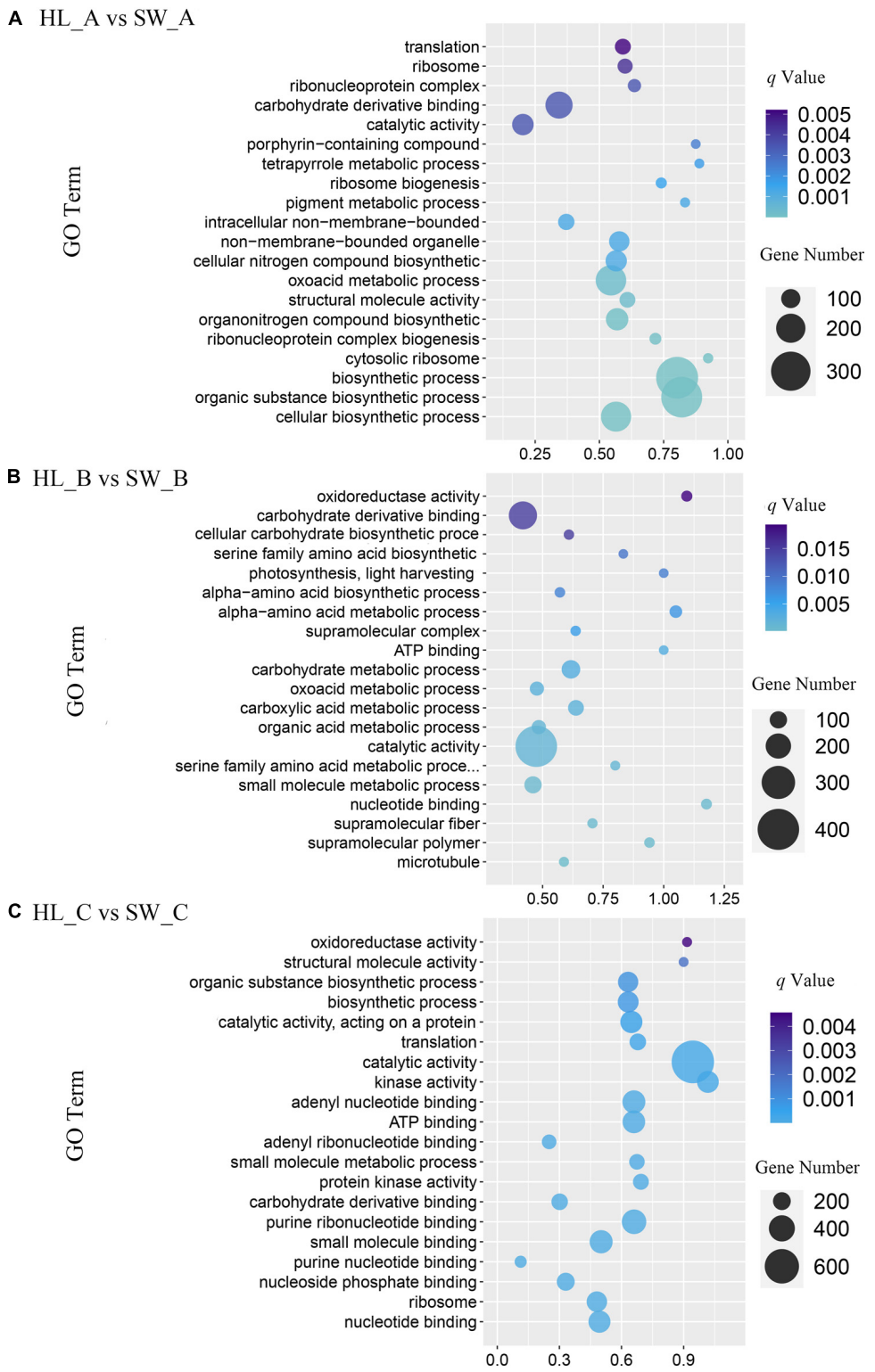


FIGURE 5 | Gene ontology (GO) pathway analysis of differentially expressed genes. Advanced bubble chart shows enrichment of differentially expressed genes in certain pathways. **(A)** HL_A vs. SW_A; **(B)** HL_B vs. SW_B; and **(C)** HL_C vs. SW_C. Count represents the DEGs number, and gene ratios are calculated as the ratio of DEGs in each GO pathway compared with total number of DEGs.

TABLE 2 | Significantly enriched pathways of differentially expressed genes (DEGs) among HL and SW samples through KEGG enrichment analyses at bud infestation stages A, B, and C.

Pathway ID	Pathway	Number of annotated	Up-regulated	Down-regulated	*q*-Value (<0.05)
		log2 (fold-change) ≥ 1			
HL_A vs SW_A					
ko04075	Plant hormone signal transduction	120	79	41	8.65E-06
ko04016	MAPK signaling pathway	105	73	32	1.67E-02
ko03010	Ribosome	37	17	20	3.25E-03
ko03008	Ribosome biogenesis in eukaryotes	25	12	13	2.09E-05
ko04145	Phagosome	10	8	2	4.96E-02
ko00196	Photosynthesis - antenna proteins	9	9	0	4.92E-06
ko03050	Proteasome	8	1	7	2.45E-02
ko00590	Arachidonic acid metabolism	4	3	1	3.49E-02
ko00860	Porphyrin and chlorophyll metabolism	3	3	0	4.48E-06
ko00750	Vitamin B6 metabolism	3	1	2	4.56E-04
HL_B vs SW_B					
ko04016	MAPK signaling pathway	55	28	27	0.005288
ko04075	Plant hormone signal transduction	36	22	14	0.019426
ko00511	Other glycan degradation	11	9	2	0.017392
ko03008	Ribosome biogenesis in eukaryotes	10	5	5	0.049904
ko04146	Peroxisome	6	4	2	0.049904
ko01212	Fatty acid metabolism	5	4	1	0.025405
ko01200	Carbon metabolism	5	1	4	0.017667
ko00051	Fructose and mannose metabolism	5	0	5	0.00222
ko00943	Isoflavonoid biosynthesis	4	2	2	0.005288
ko00860	Porphyrin and chlorophyll metabolism	3	2	1	0.000547
HL_C vs SW_C					
ko04075	Plant hormone signal transduction	187	65	122	3.93E-08
ko04016	MAPK signaling pathway	179	104	75	2.70E-08
ko00511	Other glycan degradation	30	8	22	1.91E-02
ko03050	Proteasome	18	4	14	4.62E-06
ko00941	Flavonoid biosynthesis	17	11	6	5.48E-03
ko04146	Peroxisome	16	8	8	1.28E-05
ko00564	Glycerophospholipid metabolism	16	6	10	4.19E-04
ko00592	alpha-Linolenic acid metabolism	14	11	3	4.19E-04
ko00562	Inositol phosphate metabolism	14	5	9	7.40E-03
ko01212	Fatty acid metabolism	13	5	8	8.94E-03

(log2 (fold change) $\geq$ 1 and FDR $<$ 0.05) and further analyzed as the critical genes that associate with GWDK infestation (**Figure 4A**). Among them, 4,975, 3,017, and 8,293 DEGs were differentially expressed ($\log_2$ (fold change) $\geq$ 1 and FDR $<$ 0.05) at stages A, B, and C, respectively, when comparing the two cultivars (**Figure 4B**).

Gene Ontology and Kyoto Encyclopedia of Genes and Genomes Functional Enrichment Analysis of Differentially Expressed Genes

To further investigate the functions of the detected 12,793 DEGs [$\log_2$ (fold change) $\geq$ 1], GO-based enrichment was performed with a threshold value of p-Value $<$ 0.05 to evaluate significantly enriched GO pathways. A total of 6,674 DEGs (52.16%) were annotated as "biological process," which contained the majority of GO terms, followed by 4,820 DEGs (37.15%) annotated as "cellular component" and 5,162 DEGs (39.79%) that were annotated as "molecular function." Through GO assignments, the assigned DEGs were divided into 41 functional groups. The major subcategories along with the analysis of all the transcripts among the three different stages are shown in **Figure 5A**. Transcripts associated with small molecule binding (GO: 0044283) accounted for the highest number of enriched GO terms among stage A comparisons, whereas catalytic activity (GO: 0003824) include the highest number of DEGs among both stages B and C classification (**Figures 5B,C**).

Various DEGs are coordinately expressed to perform their different biological functions; therefore, the KEGG pathway-based analysis of DEGs was performed to identify the candidate pathways and genes associated with GWDK infestation. Based

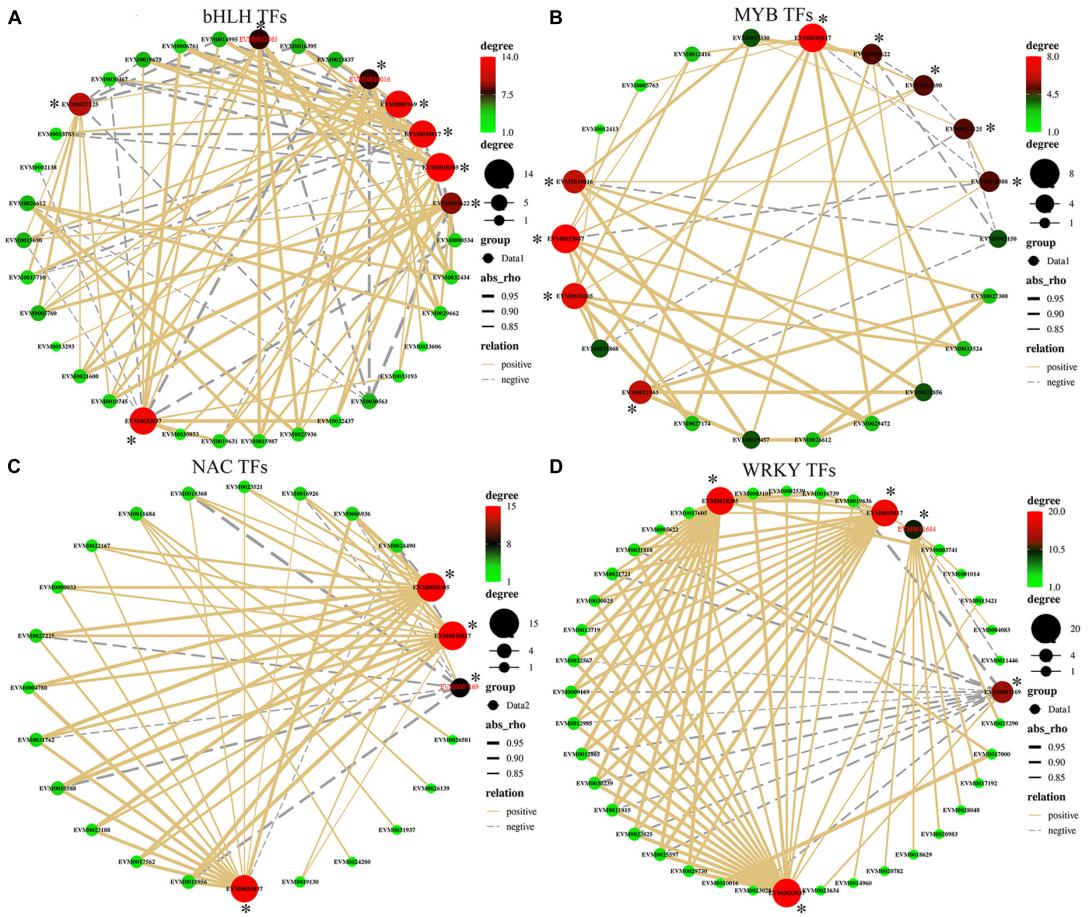


FIGURE 6 | (A–D) Correlation of gene expression between the four main DEG transcription factor (TF) families (i.e., bHLH, WRKY, NAC, and MYB) and peroxidase activity-related genes. Line color represent positive (solid yellow) and negative correlation (dotted gray); line thickness represents the strength of the correlation; dot size/color represents the number of correlated objects. * represents the antioxidant genes.

on the KEGG analysis, 12,793 DEGs were allocated to 136 pathways with the most enriched pathways as shown in **Table 2** (q-Value < 0.05, top 10 numbers of DEGs). In the comparisons of stages A, B, and C, plant hormone signal transduction (ko04075) and MAPK signaling pathway (ko04016) were detected as the top 2 significant KEGG pathways (q-Value < 0.0053) at all three stages and included the highest number of DEGs. Furthermore, in the comparison at stages B and C, the peroxisome pathway (ko04146) was also significantly enriched (q-Value of 0.049 and 1.28E-05, respectively). GO and KEGG annotations for all the DEGs are shown in **Supplementary Table 2**.

Transcripts Associated With the Plant Hormone Signal Transduction and MAPK Signaling Pathway

In this study, 157 and 178 transcripts ($\log_2$ (fold change) ≥ 1 and q-Value < 0.05) linked to plant hormone signal transduction (**Supplementary Table 3**) and MAPK signaling (**Supplementary Table 4**) were identified by comparing the three development stages of two Chinese chestnut species, respectively. In the hormone signal transduction pathway, the ABA, auxin, brassinosteroid, cytokinin, ethylene, gibberellin, and jasmonic acid (JA) signaling pathways were further analyzed (**Supplementary Table 3**). Among them, the JA signaling pathways genes were commonly reported and are often associated with plant herbivory defense. In this pathway, 19 DEGs were identified, which were annotated as coronatine-insensitive, TIFY or MYC, and bHLH transcription factors. Across the three stage comparisons, 5 TIFY protein (i.e., *EVM0015448, EVM0017479, EVM0018121, EVM0013408,* and *EVM0017717*) were all higher expressed in SW compared with HL (**Supplementary Table 3**). In the MAPK signaling pathway, the majority of DEGs were annotated as leucine-rich repeat (LRR) receptor-like kinase/receptor like (85 out of 178) and members of the transcription factor family WRKY (17 out of 178) and bHLH [9 out of 178 (**Supplementary Table 4**)]. Among the 178 DEGs of the MAPK signaling pathway, *EVM0018807* and *EVM0025267* were detected as the top two highest most highly expressed genes at the maturity stages A and B, respectively, which were annotated as acanthoscurrin-1-like and DNA damage-repair/toleration protein DRT100-like, respectively.

EVM0025267 showed consistently higher expression in the HL variety at all three stages, while *EVM0018807* was most downregulated at stage B in the HL (33-fold) variety vs. most downregulated at stage C in SW (350-fold).

Transcription Factor Analyses

Transcription factors can play a major role in regulating gene expression of target genes by binding to their promoter regulatory elements and thus affecting transcriptional activity of the targets. A total of 244 DEGs encoding TFs belonging to 20 major transcription factor families were identified across the three stages (**Supplementary Table 5**). Members of the bHLH (42/244), WRKY (40/244), NAC (30/244), and MYB families (21/244) represented the largest number of DEG transcription factors (**Supplementary Table 5**). To identify the potential functions of TFs that are associated with antioxidant enzyme-related genes, a correlation analysis of gene expressions between these four TF families and POD-related genes (POD, **Supplementary Table 6**) was performed (**Figure 6** and **Supplementary Table 7**). *CmbHLH130 (EVM0032437)* was positively correlated with the expression of POD pathway genes *EVM0033037* (r = 0.98), *EVM0018305* (r = 0.95), and *EVM0010817* (r = 0.92), while *CmIRL3* (*EVM0030563*) was negatively correlated with the expression of *EVM0027125* (r = –0.86), *EVM0026612* (r = –0.82), *EVM0021463* (–0.91), *EVM0010016* (–0.93), and *EVM0005622* (–0.95). Expression of the POD gene *EVM0033037* positive correlated with 20 members of WRKY TFs family (**Figure 6D** and **Supplementary Table 7**) (20 out of 40), while the expression of POD genes *EVM0018305*, *EVM0011684,* and *EVM0010817* correlated with 19, 9, and 18 WRKY TFs, respectively, highlighting that WRKY TF members mostly positively correlated with these 4 POD-related genes. In contrast, only POD gene *EVM0007169* has shown a negative correlation with the expression of a large suite of WRKY family genes (12 out of 40). This trend was also apparent in the correlation network of NAC with POD, where *EVM0018305* and *EVM0010817* (both 13 out of 30) were positive correlated with NAC expression (**Figure 6C** and **Supplementary Table 7**). For the MYB family, the majority of members showed a positive correlation with POD-related gene expression (**Figure 6B** and **Supplementary Table 8**) with about 70% positive correlations. For example, *EVM0018305* only positively correlated with seven MYBs (7/21), while in contrast, *EVM0007169* only negatively correlated (5/21).

Integrative Analysis of MAPK Signaling Pathway Genes Expression and Antioxidant Enzyme Activities

Based on the transcriptomics data and antioxidant enzyme activities data, we constructed a set of correlation networks between the five main antioxidant enzyme activities (i.e., PPO, CAT, POD, SOD, and H_2O_2) and the related MAPK signaling pathway gene expressions of 6 tested samples (stages A, B, and C of "HL" and "SW"). As shown in **Figure 7** and **Supplementary Table 8**, most of MAPK signaling pathway members showed a positive correlation with PPO and CAT activities, while the majority was negatively associated with POD, SOD, and H_2O_2 activity/content. POD was significantly correlated with the MAPK-related genes basic leucine zipper 61 (*EVM0029813*, r = 0.84, p-Value < 0.05), receptor-like protein 12 (*EVM0032794*, r = 0.89, p-Value < 0.05), and the transcription factor MUTE (*EVM0020177*, r = 0.87, p-Value < 0.05). In contrast, PPO activity was positively correlated to several gene members of the MAPK signaling pathway.

Correlations of Antioxidant Enzyme Activities and Expression of Antioxidant Enzyme Genes

To identify the antioxidant enzyme genes that may correlated to antioxidant enzyme activities, we constructed a set of correlation networks among the five main antioxidant enzyme activities (i.e., PPO, CAT, POD, SOD, and H_2O_2) to evaluate the expression of antioxidant enzyme-related genes. The results revealed that PPO activity was positively correlated with *EVM0026323* (r = 0.85, p-Value < 0.05) and *EVM0006016* (r = 0.72, p-Value < 0.05); CAT activity was positively correlated with *EVM002624729* (r = 0.81, p-Value < 0.05); POD activity was positively correlated with *EVM0026612* (r = 0.89, p-Value < 0.05), *EVM0005622* (r = 0.84, p-Value < 0.05), and negatively correlated with *EVM0011684* (r = –0.71, p-Value < 0.05). SOD activity was positively correlated with *EVM0032156* (r = 0.9, p-Value < 0.05), *EVM0033049* (r = 0.79, p-Value < 0.05), while *EVM0005687* expression was negatively correlated (r = –0.79, p-Value < 0.05) with buds H_2O_2 contents (**Supplementary Table 9**). All the aforementioned antioxidant enzyme activity-related genes could be important candidates for further Chinese chestnuts "antioxidant enzyme" study.

Effect of Gall Wasp *Dryocosmus kuriphilus* Infestation on Expression of Terpene Synthase Enzymes

Terpenes are a class of secondary metabolite volatile compounds derived from the mevalonate and MEP substrate pathways that are often associated with direct and indirect defenses to pests and pathogens. In this study, GWDK feeding exhibits a significant effect on the elicited expression of terpene synthase (TPS) genes (**Supplementary Table 10**). For example, the expression of TPS gene *EVM0000042* was annotated as (*E,E*)-alpha-farnesene synthase and presented an 8-fold higher expression in HL than SW at the initiation stage A (**Supplementary Table 10**), while two (*3S,6E*)-nerolidol synthase genes *EVM0024276* and *EVM0000019* were induced 3- to 12-fold in both cultivars after infestation (**Supplementary Table 10**). Interestingly, (α)-farnesene and (*E*)-nerolidol have both been reported to be associated with herbivory defenses in plants (Jing et al., 2021).

Validation of Differentially Expressed Genes by Quantitative Reverse Transcription-PCR

To validate the expression patterns for the DEGs generated by RNA-Seq, the expression levels of 8 randomly selected genes were further analyzed by qRT-PCR across the 6 sequence samples (i.e.,

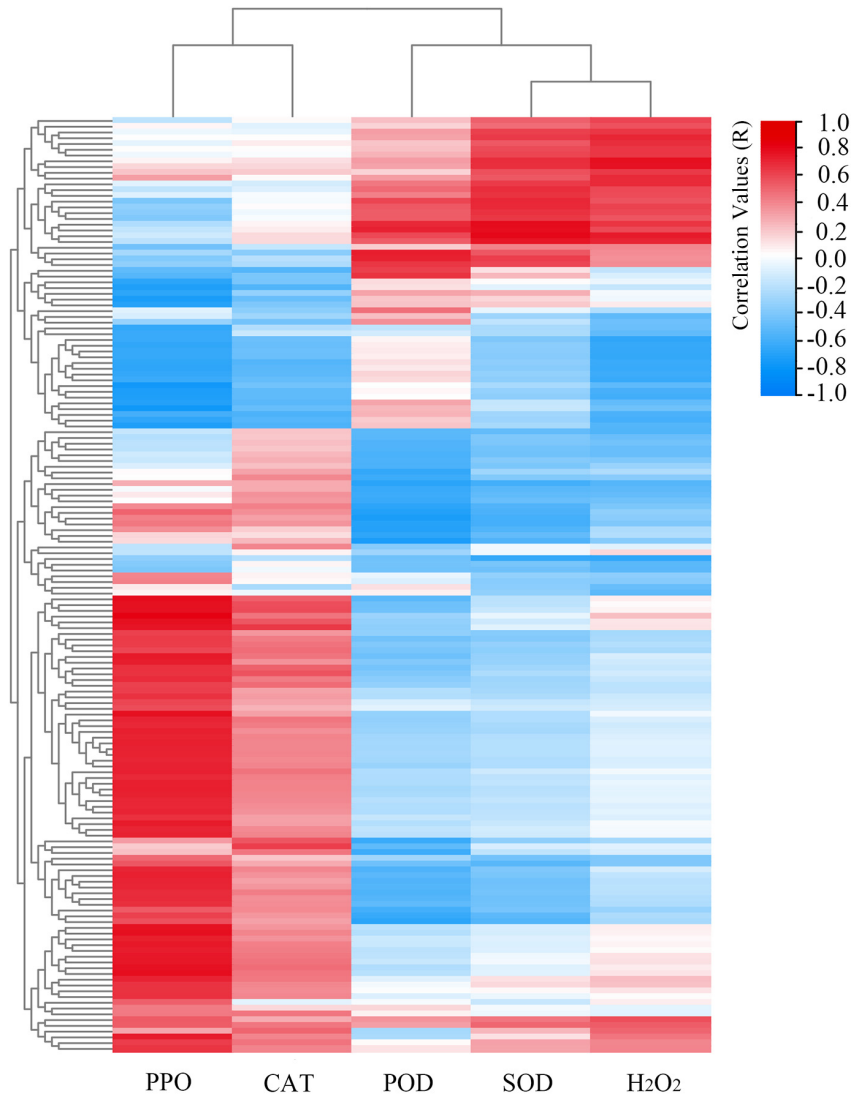


FIGURE 7 | Representative candidate genes of MAPK signaling pathway likely linked to the H_2O_2 content and the main antioxidant enzyme activities (i.e., PPO, CAT, POD, and SOD). Scale bar represents the correlation values.

3 biological repeats per samples**).** The qRT-PCR data showed very similar patterns and trends of expression compared with the RNA-Seq transcriptomics data (**Figure 8** lines vs. bar graphs). These results validate that the RNA-Seq data can be considered reliable and reproducible by an independent method (i.e., qRT-PCR) using a random selection of genes ($n = 8$).

DISCUSSION

The oriental gall wasp *D. kuriphilus* (GWDK) is considered the most serious pest to the Chinese chestnut industry, and infestation results in gall formation on vegetative parts such as buds, leaves, and petioles. It also contributes to large losses in fruit yield due to weakening of the plant health and indirectly by plants becoming more susceptible to other pests and diseases. The molecular and physiological aspects of GWDK infestation in European chestnuts (*Castanea sativa*, Fagaceae) have been previously examined in several studies (Sartor et al., 2012, 2015; Acquadro et al., 2020). The interaction of a specialist herbivore with a plant host can be exploited to dissect the architecture of plant defense signaling networks that can induce plant resistance or tolerance (Walling, 2008; Kaloshian and Walling, 2016). However, the molecular mechanisms and targets involved in GWDK infestations of Chinese chestnuts remain largely unknown. In this study, the comparative transcriptomic analysis of susceptible "HL" and the partially resistant "SW" Chinese chestnut (*C. mollisima*) buds at different stages during GWDK infestation were examined. A large number of DEGs that differed among the budburst stages of the 2 species were observed. Additionally, through KEGG enrichment analysis, a large portion of the DEGs were involved in the likely recognition

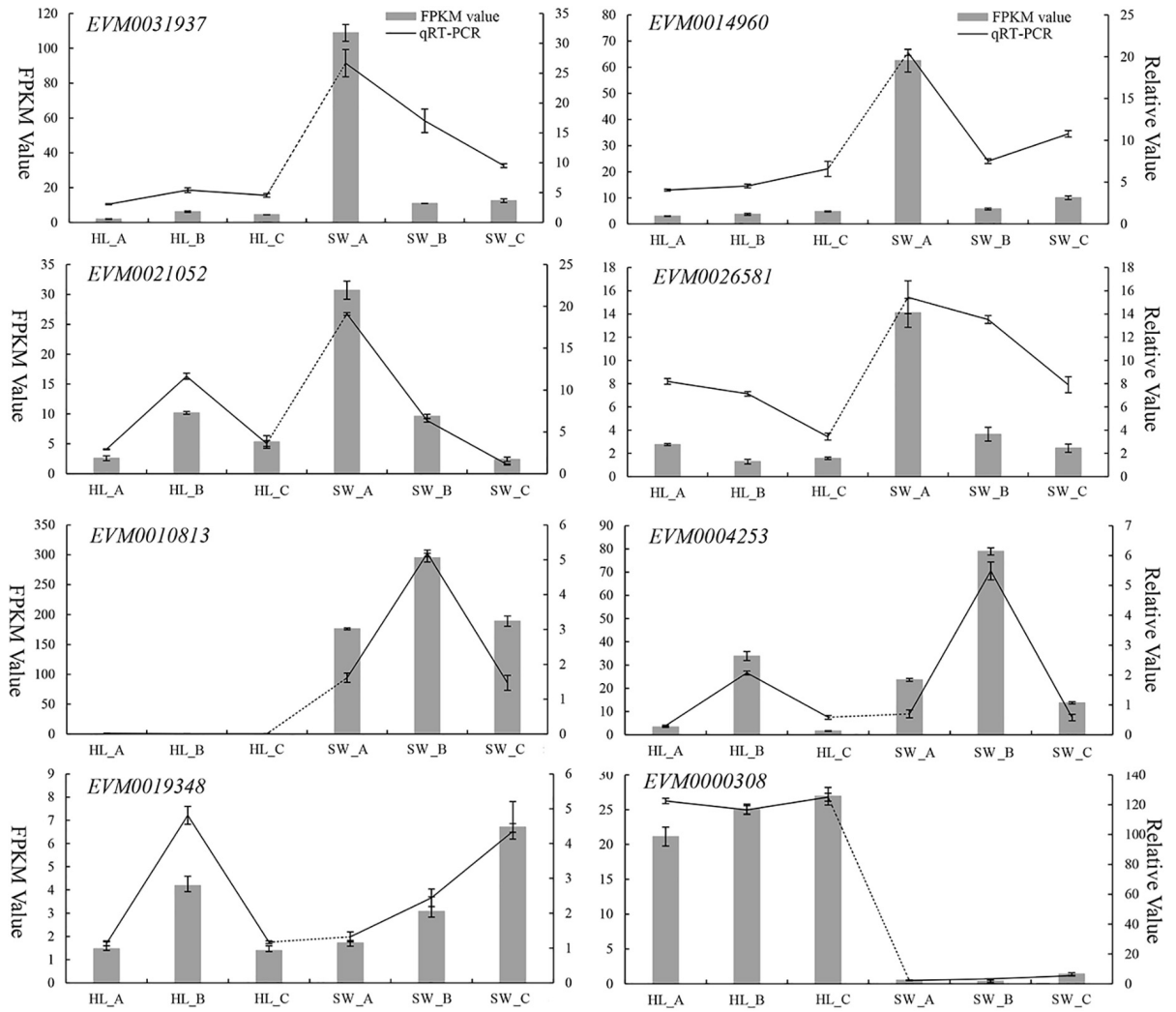


FIGURE 8 | Quantitative reverse-transcriptase (qRT)-PCR validation of DEGs expression in developing "HL" and "SW" chestnut buds vs. by RNA-Seq. dark line: qRT-PCR expression, gray bars: transcriptomics RNA-Seq analysis [fragments per kilobase per million reads (FPKM)].

of specific elicitors and herbivore damage, and substantial variations were identified in four of the main pathways that were enriched, including plant hormone signal transduction, MAPK signaling, ribosome biogenesis in eukaryotes, and the peroxisome pathway.

Oxidative stress signaling is proposed to be an important process induced by herbivory infestation (Caverzan et al., 2016), where H_2O_2 and ROS are common elicitors of defense signaling that are involved in the elicitation of plant responses to herbivory attack. Oxidative enzymes have been reported as a signaling factor to increase the expression of genes related to biotic stresses (Maffei et al., 2007). This study highlights that POD, H_2O_2, and SOD were higher in resistant cultivar leaves compared with susceptible HL at most of the development stages of infestation, while CAT and PPO activities were mostly higher in susceptible (i.e., HL) leaves (**Figure 2**). Through the GO and KEGG enrichment analysis, we demonstrated that the "catalytic activity" was a significant GO-enriched term in stages B and C, which relates to enzyme catalysis (**Figure 5**), while in the KEGG enrichment analysis, "MAPK signaling pathway" and "plant hormone signal transduction" featured prominently in all three stages (**Table 2**). Modifying the levels of SOD, POD, and CAT gene expression and enzyme activities affects the H_2O_2 balance through production and scavenging and may directly (insecticidal) or indirectly (defense signaling) influence galls formation and growth. Our results were consistent with the studies carried out on tea (Ning et al., 1996) and wheat (Han et al., 2009) plants that stronger activities of glutamine dehydrogenase and PAL were observed in the shoots, which were resistant to pink mite feeding. In the wheat study, resistance against cereal aphid was associated with higher constitutive PAL and PPO activity, while aphid infestation further enhanced the levels of PAL and PPO activities in both resistant and susceptible cultivars. In contrast, aphid infestation induced POD activity in

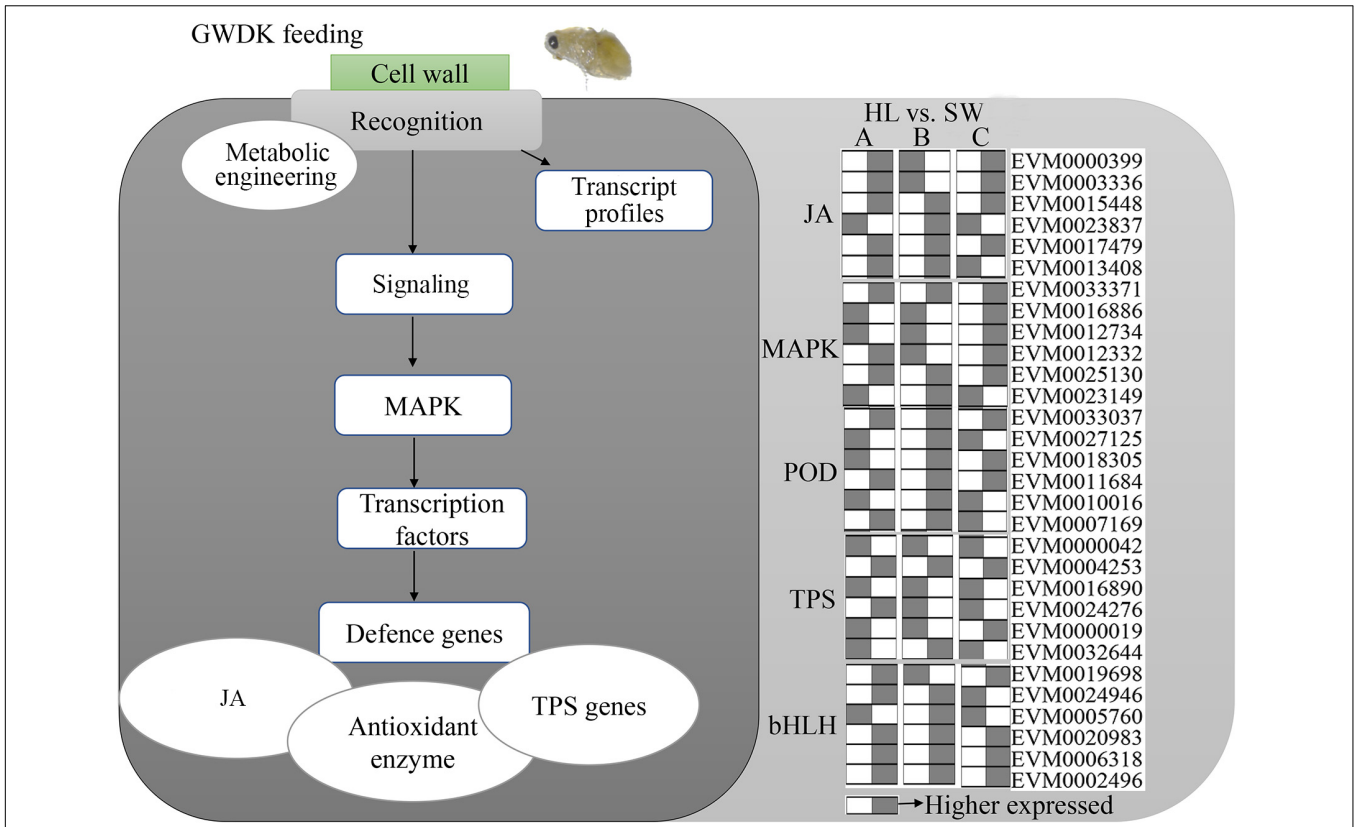


FIGURE 9 | Genetic model of Chinese chestnut buds responses upon GWDK infestation. The heatmap includes the top 6 highest expressed genes in jasmonic acid (JA), MAPK, POD, terpene synthase (TPS), and bHLH gene classes; gray color represents higher expressed of a certain gene compared with white (lower).

both tested cultivars, especially in susceptible ones. Therefore, further investigation into the complex relationship between several defense pathways in relation to GWDK infestation and the role of the specialized antioxidant enzyme systems is required.

Constitutive defense mostly directly acts on the invading agent, while induced responses such as ROS and phytoalexins defense signals, in addition to direct defense, can benefit the plant by employment "on demand." For example, herbivore-induced release of volatile compounds has received much attentions, including terpenoids, aromatic compounds, and green leaf volatiles, and transcriptional analysis of these responses has been widely studied (He et al., 2020a; Pingault et al., 2021). Considering that a number of TPS genes were significantly upregulated after GWDK infestation in both HL and SW plants at the later infestations stages B and C compared with the initially budburst stage A, we hypothesized that these TPS genes are good candidate genes for further study that may contribute to the production and emission of several terpene compounds either involved in inhibition of herbivory or by attracting predatory insects (Clancy et al., 2020). Our study provides some indirect evidence to support a network between herbivory feeding and terpene production and release, but more studies are required to link this to the terpenoid secondary metabolites that are induced.

In the process of herbivory feedings, transcription factors often play pivotal roles in regulating transcriptional responses by the plant. The functional category of "transcription factors" was significantly enriched in the transcript expression profiles of many biotic defense-related transcriptomics comparative studies, and the WRKY TF class is widely reported in the context of biotic and abiotic stress responses (Bai et al., 2018). Induction of the expression of WRKY TF genes in response to herbivory was observed in *N. attenuata* (Skibbe et al., 2008), rice (Akagi et al., 2014), and other plants (Kundu and Vadassery, 2021), which affected the biosynthesis of defensive secondary metabolites or by reprogramming the expression of associated pathway genes. Apart from WRKY TFs, other classes of TFs associated with herbivory infestations have also been widely reported, for example, in rice, leaf, and phloem feeding by brown planthopper (BPH) induced the transcript levels of genes encoding transcription factors of the AP2/ERF, MYB, bZIP, and bHLH families (Li et al., 2021). *OsMYB30* may play a key role in modulating the resistance of rice to BPH by regulating the biosynthesis of salicylic acid and expressions of genes in the PAL pathway (He et al., 2020b). In Chinese chestnuts, the genetic mechanisms of TF action involved in GWDK feedings have not been previously experimentally elucidated. We identified 244 TF members belonging to 20 major TFs families that were differentially expressed in the susceptible vs. resistant cultivar during one or more of the stages of bud infestation by GWDK. Among them, four transcription factors family,

namely, bHLH, WRKY, NAC, and MYB, were widely represented (**Supplementary Table 5**). Candidate TFs identified in this study can contribute in future to a more detailed understanding of how the transcriptional regulatory networks may contribute to a successful herbivory resistance response.

When attacked by herbivories, plants can respond in terms of signal transduction *via* phytohormonal pathways, induce gene expression changes, which may lead to the biosynthesis of secondary metabolites and defense proteins (Erb and Reymond, 2019). Jasmonic acid (JA) has been long recognized as a pivotal mediator that regulates a myriad of plant developmental and biotic stress responses especially against herbivores, and pathogens as well as abiotic stress responses such as wounding and UV damage (Sun et al., 2011; Alhaithloul and Soliman, 2021). Chewing herbivores such as caterpillars and piercing-sucking insects such as white flies can significantly induce the jasmonate pathway in *Arabidopsis thaliana* (Zhang et al., 2013, 2020). Methyl jasmonate treatment can also induce the defense responses in bilberry (Benevenuto et al., 2019), Norway spruce (Mageroy et al., 2020), and Jacobaea vulgaris (Wei et al., 2021). In our study, KEGG enrichment identified "plant hormone signal transduction" and "MAPK signaling" (**Supplementary Table 3**) among the top enriched pathways in all three stages. Additionally, five proteins annotated as TIFY protein and JA-related transcription factors MYC2 were detected and showed higher expression in resistant SW compared with the susceptible HL variety.

Understanding of insect-plant interactions is of interest not only from an ecological and evolutionary perspective but also for the development of novel crop protection strategies such as through resistance breeding. The data presented in this study provide genetic information for the discovery of new regulatory steps (TFs) and can identify novel biosynthetic genes such as terpene synthases that may, for example, produce repellent, antifeeding, or insecticidal compounds or act as volatile signals that attract natural enemies of herbivores. WRKY TFs and related genes may influence the levels of target modifying/sequestration enzymes (e.g., PAL and PPO) involved in the GWDK resistance. Our correlation analysis of TFs and oxidative enzyme levels (**Figure 6** and **Supplementary Table 7**) revealed that aforementioned enzymes and TFs could potentially contribute to the resistance process as part of a network of induced defense responses.

To further categorize and visualize the various functions and components of the GWDK-responsive genes in buds, all DEGs were assembled into a model for GWDK responses (**Figure 9**). A large proportion of the DEGs belong to the signals derived from JA and MAPK signaling and involves enzymatic metabolism such as POD and TPS. Most of these pathways have been reported to play major roles in the stress responses in plants.

CONCLUSION

In summary, we investigated the defense response of GWDK-susceptible and GWDK-partial resistant Chinese chestnuts through *de novo* transcriptome sequencing and oxidative enzyme metabolism analysis. It shows that the activities of CAT and POD, PPO, and SOD were closely associated with the Chinese chestnuts GWDK infestation, where the CAT and PPO activities were significant higher in HL buds compared with SW. RNA-Seq transcriptomic analyses of HL and SW leaves revealed that various metabolic pathways involved in GWDK biotic stress/defense responses, such as plant hormone signal transduction, MAPK signaling, ribosome biogenesis in eukaryotes, and peroxisome pathway, were enriched. Moreover, changes of gene expression levels of terpenoid biosynthetic, plant hormone signaling transduction, and WRKY TFs were observed. Further study on the herbivory feeding and related DEGs can provide us more detailed insights into GWDK resistance mechanisms, and this knowledge can be exploited and used in agriculture to inform better management practices and in plant breeding to improve the genetics and ultimately the quality and quantity of Chinese chestnuts production.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://www.ncbi.nlm.nih.gov/sra, PRJNA791965.

AUTHOR CONTRIBUTIONS

CZ, WW, and YC participated in the design of the study and revised the manuscript. WW, SZ, and YZ contributed to performing the experiments, data analyses, and manuscript writing. CZ conducted antioxidant enzyme metabolite data analyses. MK-U-R and NN contributed to the writing and reviewing of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China under Grant No. (31500514), the Independent Scientific Research Project of Institute of Botany, Jiangsu Province, and Chinese Academy of Sciences under Grant Nos. (JSPKLB202030, JSPKLB202043, and JSPKLB202024) and Special Fund of the Chestnut Germplasm Resources Repositories in Jiangsu Province, China.

ACKNOWLEDGMENTS

We acknowledge NN for the linguistic editing during the preparation of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.874434/full#supplementary-material

REFERENCES

Acquadro, A., Torello Marinoni, D., Sartor, C., Dini, F., Macchio, M., and Botta, R. (2020). Transcriptome characterization and expression profiling in chestnut cultivars resistant or susceptible to the gall wasp *Dryocosmus kuriphilus*. *Mol. Genet. Genomics* 295, 107–120. doi: 10.1007/s00438-019-01607-2

Akagi, A., Fukushima, S., Okada, K., Jiang, C. J., Yoshida, R., and Nakayama, A. (2014). WRKY45-dependent priming of diterpenoid phytoalexin biosynthesis in rice and the role of cytokinin in triggering the reaction. *Plant Mol. Biol.* 86, 171–183. doi: 10.1007/s11103-014-0221-x

Alhaithloul, H. A. S., and Soliman, M. H. (2021). "Methyl jasmonate and brassinosteroids: emerging plant growth regulators in plant abiotic stress tolerance and environmental changes," in *Plant Growth Regulators*, eds T. Aftab and K. R. Hakeem (Cham: Springer Nature Switzerland AG), 173–195. doi: 10.1007/978-3-030-61153-8_8

Amako, K., Chen, G. X., and Asada, K. (1994). Separate assays specific for ascorbate peroxidase and guaiacol peroxidase and for the chloroplastic and cytosolic isozymes of ascorbate peroxidase in plants. *Plant Cell Physiol.* 35, 497–504.

Bai, Y., Sunarti, S., Kissoudis, C., Visser, R. G., and van der Linden, C. (2018). The role of tomato WRKY genes in plant responses to combined abiotic and biotic stresses. *Front. Plant Sci.* 9:801. doi: 10.3389/fpls.2018.00801

Barakat, A., Staton, M., Cheng, C. H., Park, J., Yassin, N. M. B., Ficklin, S., et al. (2012). Chestnut resistance to the blight disease: insights from transcriptome analysis. *BMC Plant Biol.* 12:38. doi: 10.1186/1471-2229-12-38

Beers, R. F., and Sizer, I. W. (1952). A spectrophotometric method for measuring the breakdown of hydrogen peroxide by catalase. *J. Biol. Chem.* 195, 133–140. doi: 10.1016/s0021-9258(19)50881-x

Benevenuto, R. F., Seldal, T., Hegland, S. J., Rodriguez-Saona, C., Kawash, J., and Polashock, J. (2019). Transcriptional profiling of methyl jasmonate-induced defense responses in bilberry (*Vaccinium myrtillus* L.). *BMC Plant Biol.* 19:70. doi: 10.1186/s12870-019-1650-0

Bosch, M., Berger, S., Schaller, A., and Stintzi, A. (2014). Jasmonate-dependent induction of polyphenol oxidase activity in tomato foliage is important for defense against *Spodoptera exigua* but not against *Manduca sexta*. *BMC Plant Biol.* 14:257. doi: 10.1186/s12870-014-0257-8

Caverzan, A., Casassola, A., and Brammer, S. P. (2016). "Reactive oxygen species and antioxidant enzymes involved in plant tolerance to stress," in *Abiotic and Biotic Stress in Plants: Recent Advances and Future Perspectives*, Vol. 17, eds A. Shanker and C. Shanker (London: Intechopen), 463–480.

Chapman, J. M., Muhlemann, J. K., Gayomba, S. R., and Muday, G. K. (2019). RBOH-dependent ROS synthesis and ROS scavenging by plant specialized metabolites to modulate plant development and stress responses. *Chem. Res. Toxicol.* 32, 370–396. doi: 10.1021/acs.chemrestox.9b00028

Chojak-Koźniewska, J., Kuźniak, E., and Zimny, J. (2018). The effects of combined abiotic and pathogen stress in plants: insights from salinity and *Pseudomonas syringae* pv *lachrymans* interaction in cucumber. *Front. Plant Sci.* 9:1691. doi: 10.3389/fpls.2018.01691

Clancy, M. V., Haberer, G., Jud, W., Niederbacher, B., Niederbacher, S., Senft, M., Schnitzler, J. P., et al. (2020). Under fire-simultaneous volatilome and transcriptome analysis unravels fine-scale responses of tansy chemotypes to dual herbivore attack. *BMC Plant Biol.* 20:551. doi: 10.1186/s12870-020-02745-1

Erb, M., and Reymond, P. (2019). Molecular interactions between plants and insect herbivores. *Annu. Rev. Plant Biol.* 70, 527–557. doi: 10.1146/annurev-arplant-050718-095910

Geng, G., Zhu, C., and Zhou, J. (2015). Resistance of *Castanea mollissima* Shuhe-WYL strain to *Dryocosmus kuriphilus* and its molecular mechanism. *Genet. Mol. Res.* 14, 11456–11461. doi: 10.4238/2015.September.25.12

Giannopolitis, C. N., and Ries, S. K. (1977). Superoxide dismutases: I. Occurrence in higher plants. *Plant Physiol.* 59, 309–314. doi: 10.1104/pp.59.2.309

Granett, J., Walker, M. A., Kocsis, L., and Omer, A. D. (2001). Biology and management of grape phylloxera. *Annu. Rev. Entomol.* 46, 387–412. doi: 10.1146/annurev.ento.46.1.387

Han, Y., Wang, Y., Bi, J. L., Yang, X. Q., Huang, Y., Zhao, X., et al. (2009). Constitutive and induced activities of defense-related enzymes in aphid-resistant and aphid-susceptible cultivars of wheat. *J. Chem. Ecol.* 35, 176–182. doi: 10.1007/s10886-009-9589-5

He, J., Bouwmeester, H. J., Dicke, M., and Kappers, I. F. (2020a). Transcriptional and metabolite analysis reveal a shift in direct and indirect defences in response to spider-mite infestation in cucumber (*Cucumis sativus*). *Plant Mol. Biol.* 103, 489–505. doi: 10.1007/s11103-020-01005-y

He, J., Liu, Y., Yuan, D., Duan, M., Liu, Y., Shen, Z., et al. (2020b). An R2R3 MYB transcription factor confers brown planthopper resistance by regulating the phenylalanine ammonia-lyase pathway in rice. *Proc. Natl. Acad. Sci. U.S.A.* 117, 271–277. doi: 10.1073/pnas.1902771116

Huang, J., Lu, X., Wu, H., Xie, Y., Peng, Q., Gu, L., et al. (2020). Phytophthora effectors modulate genome-wide alternative splicing of host mRNAs to reprogram plant immunity. *Mol. Plant* 13, 1470–1484. doi: 10.1016/j.molp.2020.07.007

Jing, T., Du, W., Gao, T., Wu, Y., Zhang, N., Zhao, M., et al. (2021). Herbivore-induced DMNT catalyzed by CYP82D47 plays an important role in the induction of JA-dependent herbivore resistance of neighboring tea plants. *Plant Cell Environ.* 44, 1178–1191. doi: 10.1111/pce.13861

Kaloshian, I., and Walling, L. L. (2016). Hemipteran and dipteran pests: effectors and plant host immune regulators. *J. Integr. Plant Biol.* 58, 350–361. doi: 10.1111/jipb.12438

Korgaonkar, A., Han, C., Lemire, A. L., Siwanowicz, I., Bennouna, D., Kopec, R. E., et al. (2021). A novel family of secreted insect proteins linked to plant gall development. *Curr. Biol.* 31, 1836–1849.e12.

Kundu, P., and Vadassery, J. (2021). Role of WRKY transcription factors in plant defense against lepidopteran insect herbivores: an overview. *J. Plant Biochem. Biotechnol.* 30, 698–707. doi: 10.1007/s13562-021-00730-9

Li, Y., Cheah, B. H., Fang, Y. F., Kuang, Y. H., Lin, S. C., Liao, C. T., et al. (2021). Transcriptomics identifies key defense mechanisms in rice resistant to both leaf-feeding and phloem feeding herbivores. *BMC Plant Biol.* 21:306. doi: 10.1186/s12870-021-03068-5

Liochev, S. I., and Fridovich, I. (2007). The effects of superoxide dismutase on H_2O_2 formation. *Free Radic. Biol. Med.* 42, 1465–1469.

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2- ΔΔCT method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Loreto, F., and Velikova, V. (2001). Isoprene produced by leaves protects the photosynthetic apparatus against ozone damage, quenches ozone products, and reduces lipid peroxidation of cellular membranes. *Plant Physiol.* 127, 1781–1787. doi: 10.1104/pp.010497

Maffei, M. E., Mithöfer, A., and Boland, W. (2007). Insects feeding on plants: rapid signals and responses preceding the induction of phytochemical release. *Phytochemistry* 68, 2946–2959. doi: 10.1016/j.phytochem.2007.07.016

Mageroy, M. H., Wilkinson, S. W., Tengs, T., Cross, H., Almvik, M., Pétriacq, P., et al. (2020). Molecular underpinnings of methyl jasmonate-induced resistance in Norway spruce. *Plant Cell Environ.* 43, 1827–1843. doi: 10.1111/pce.13774

Martinson, E. O., Werren, J. H., and Egan, S. P. (2021). Tissue-specific gene expression shows a cynipid wasp repurposes oak host gene networks to create a complex and novel parasite-specific organ. *Mol. Ecol.* [Online ahead of print]. doi: 10.1111/mec.16159

Matsui, S., and Torikata, H. (1970). Studies on the resistance of chestnut trees to chestnut gall wasp. II. The resistance of trees to gall wasp and the histochemistry of the chestnut buds. *Engei Gakkai Zasshi* 39, 44–54. doi: 10.2503/jjshs.39.44

Ning, X., Chen, X., Chen, H., and Chen, Z. (1996). Morphological and biochemical parameters of tea varieties resistant to pink mite (*Acaphylla theae* Watt). *J. Tea Sci.* 16, 125–130.

Nishio, S., Takada, N., Terakami, S., Takeuchi, Y., Kimura, M. K., Isoda, K., et al. (2021). Genetic structure analysis of cultivated and wild chestnut populations reveals gene flow from cultivars to natural stands. *Sci. Rep.* 11:240. doi: 10.1038/s41598-020-80696-1

Pingault, L., Varsani, S., Palmer, N., Ray, S., Williams, W. P., Luthe, D. S., Louis, J., et al. (2021). Transcriptomic and volatile signatures associated with maize defense against corn leaf aphid. *BMC Plant Biol.* 21:138. doi: 10.1186/s12870-021-02910-0

Raman, A. (2011). Morphogenesis of insect-induced plant galls: facts and questions. *Flora Morphol. Distrib. Funct. Ecol. Plants* 206, 517–533. doi: 10.1016/j.flora.2010.08.004

Ronquist, F., and Liljeblad, J. (2001). Evolution of the gall wasp-host plant association. *Evolution* 55, 2503–2522. doi: 10.1111/j.0014-3820.2001.tb00765.x

Sampson, B. J., Stringer, S. J., and Spiers, J. M. (2002). Integrated pest management for *Dasineura oxycoccana* (Diptera: Cecidomyiidae) in blueberry. *Environ. Entomol.* 31, 339–347. doi: 10.1603/0046-225x-31.2.339

Sartor, C., Dini, F., Marinoni, D. T., Mellano, M. G., Beccaro, G. L., Alma, A., et al. (2015). Impact of the Asian wasp *Dryocosmus kuriphilus* (Yasumatsu) on cultivated chestnut: yield loss and cultivar susceptibility. *Sci. Hortic.* 197, 454–460. doi: 10.1016/j.scienta.2015.10.004

Sartor, C., Marinoni, D. T., Quacchia, A., and Botta, R. (2012). Quick detection of *Dryocosmus kuriphilus* Yasumatsu (Hymenoptera: Cynipidae) in chestnut dormant buds by nested PCR. *Bull. Entomol. Res.* 102, 367–371. doi: 10.1017/S0007485311000812

Selig, P., Keough, S., Nalam, V. J., and Nachappa, P. (2016). Jasmonate-dependent plant defenses mediate soybean thrips and soybean aphid performance on soybean. *Arthropod Plant Interact.* 10, 273–282. doi: 10.1007/s11829-016-9437-9

Shi, Y., Zhang, Y., Yao, H., Wu, J., Sun, H., and Gong, H. (2014). Silicon improves seed germination and alleviates oxidative stress of bud seedlings in tomato under water deficit stress. *Plant Physiol. Biochem.* 78, 27–36. doi: 10.1016/j.plaphy.2014.02.009

Skibbe, M., Qu, N., Galis, I., and Baldwin, I. T. (2008). Induced plant defenses in the natural environment: *Nicotiana attenuata* WRKY3 and WRKY6 coordinate responses to herbivory. *Plant Cell* 20, 1984–2000. doi: 10.1105/tpc.108.058594

Smiley, R. W., Gourlie, J. A., Whittaker, R. G., Easley, S. A., and Kidwell, K. K. (2004). Economic impact of Hessian fly (Diptera: Cecidomyiidae) on spring wheat in Oregon and additive yield losses with Fusarium crown rot and lesion nematode. *J. Econ. Entomol.* 97, 397–408. doi: 10.1093/jee/97.2.397

Stone, G. N., and Schönrogge, K. (2003). The adaptive significance of insect gall morphology. *Trends Ecol. Evol.* 18, 512–522. doi: 10.1016/s0169-5347(03)00247-7

Sun, J. Q., Jiang, H. L., and Li, C. Y. (2011). Systemin/jasmonate-mediated systemic defense signaling in tomato. *Mol. Plant* 4, 607–615. doi: 10.1093/mp/ssr008

Trujillo-Moya, C., Ganthaler, A., Stöggl, W., Kranner, I., Schüler, S., Ertl, R., et al. (2020). RNA-Seq and secondary metabolite analyses reveal a putative defence-transcriptome in Norway spruce (*Picea abies*) against needle bladder rust (*Chrysomyxa rhododendri*) infection. *BMC Genomics* 21:336. doi: 10.1186/s12864-020-6587-z

Vasaikar, S. V., Straub, P., Wang, J., and Zhang, B. (2018). LinkedOmics: analyzing multi-omics data within and across 32 cancer types. *Nucleic Acids Res.* 46, D956–D963. doi: 10.1093/nar/gkx1090

Walling, L. L. (2008). Avoiding effective defenses: strategies employed by phloem-feeding insects. *Plant Physiol.* 146, 859–866. doi: 10.1104/pp.107.113142

Wang, J., Tian, S., Sun, X., Cheng, X., Duan, N., Tao, J., et al. (2020). Construction of pseudomolecules for the Chinese chestnut (*Castanea mollissima*) genome. *G3* 10, 3565–3574. doi: 10.1534/g3.120.401532

War, A. R., Paulraj, M. G., Ahmad, T., Buhroo, A. A., Hussain, B., Ignacimuthu, S., et al. (2012). Mechanisms of plant defense against insect herbivores. *Plant Signal. Behav.* 7, 1306–1320. doi: 10.4161/psb.21663

Way, M. O., Grigarick, A. A., Litsinger, J. A., Palis, F., and Pingali, P. (1991). "Economic thresholds and injury levels for insect pests of rice," in *Rice Insects: Management Strategies*, eds E. A. Heinrichs and T. A. Miller (New York, NY: Springer), 67–105. doi: 10.1007/978-1-4612-3124-0_2

Wei, X., Vrieling, K., Kim, H. K., Mulder, P. P., and Klinkhamer, P. G. (2021). Application of methyl jasmonate and salicylic acid lead to contrasting effects on the plant's metabolome and herbivory. *Plant Sci.* 303:110784. doi: 10.1016/j.plantsci.2020.110784

Wissemann, K. W., and Lee, C. (1980). Polyphenoloxidase activity during grape maturation and wine production. *Am. J. Enol. Vitic.* 31, 206–211.

Zhang, J., Zhang, X., Ye, M., Li, X. W., Lin, S. B., and Sun, X. L. (2020). The jasmonic acid pathway positively regulates the polyphenol oxidase-based defense against tea geometrid caterpillars in the tea plant (*Camellia sinensis*). *J. Chem. Ecol.* 46, 308–316. doi: 10.1007/s10886-020-01158-6

Zhang, P. J., Broekgaarden, C., Zheng, S. J., Snoeren, T. A., van Loon, J. J., Gols, R., et al. (2013). Jasmonate and ethylene signaling mediate whitefly-induced interference with indirect plant defense in *Arabidopsis thaliana*. *New Phytol.* 197, 1291-1299. doi: 10.1111/nph.12106

Zhu, C., Shi, F., Chen, Y., Wang, M., Zhao, Y., and Geng, G. (2019). Transcriptome analysis of Chinese chestnut (*Castanea mollissima* Blume) in response to *Dryocosmus kuriphilus* Yasumatsu infestation. *Int. J. Mol. Sci.* 20:855. doi: 10.3390/ijms20040855

Conflict of Interest: NN was employed by the New Zealand Institute for Plant and Food Research Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 18 April 2022
doi: 10.3389/fpls.2022.843725

Transcriptomic Analysis of Cadmium Stressed *Tamarix hispida* Revealed Novel Transcripts and the Importance of Abscisic Acid Network

*Pei-Long Wang[1,2], Xiao-Jin Lei[1], Yuan-Yuan Wang[1], Bai-chao Liu[1], Dan-ni Wang[1], Zhong-Yuan Liu[1] and Cai-Qiu Gao[1]**

[1] State Key Laboratory of Tree Genetics and Breeding, Northeast Forestry University, Harbin, China, [2] Zhejiang Institute of Subtropical Crops, Zhejiang Academy of Agricultural Sciences, Wenzhou, China

OPEN ACCESS

Edited by:
Prasanta Kumar Subudhi, Louisiana State University, United States

Reviewed by:
Parviz Heidari, Shahrood University of Technology, Iran
Wenjing Yao, Nanjing Forestry University, China

***Correspondence:**
Cai-Qiu Gao
gaocaiqiu@nefu.edu.cn

Specialty section:
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

Received: *26 December 2021*
Accepted: *18 February 2022*
Published: *18 April 2022*

Citation:
Wang P-L, Lei X-J, Wang Y-Y, Liu B-c, Wang D-n, Liu Z-Y and Gao C-Q (2022) Transcriptomic Analysis of Cadmium Stressed Tamarix hispida Revealed Novel Transcripts and the Importance of Abscisic Acid Network. Front. Plant Sci. 13:843725. doi: 10.3389/fpls.2022.843725

Cadmium (Cd) pollution is widely detected in soil and has been recognized as a major environmental problem. *Tamarix hispida* is a woody halophyte, which can form natural forest on the desert and soil with 0.5 to 1% salt content, making it an ideal plant for the research on response to abiotic stresses. However, no systematic study has investigated the molecular mechanism of Cd tolerance in *T. hispida*. In the study, RNA-seq technique was applied to analyze the transcriptomic changes in *T. hispida* treated with 150 μmol L^{-1} $CdCl_2$ for 24, 48, and 72 h compared with control. In total, 72,764 unigenes exhibited similar sequences in the Non-redundant nucleic acid database (NR database), while 36.3% of all these unigenes may be new transcripts. In addition, 6,778, 8,282, and 8,601 DEGs were detected at 24, 48, and 72 h, respectively. Functional annotation analysis indicated that many genes may be involved in Cd stress response, including ion bonding, signal transduction, stress sensing, hormone responses and ROS metabolism. A *ThUGT* gene from the abscisic acid (ABA) signaling pathway can enhance Cd resistance ability of *T. hispida* by regulating the production of ROS under Cd stress and inhibit absorption of Cd. The new transcriptome resources and data that we present in this study for *T. hispida* may facilitate investigation of molecular mechanisms governing Cd resistance.

Keywords: ***Tamarix hispida*** **Willd, cadmium stress, transcriptomic analysis, differentially expressed genes, H_2O_2, ABA**

Abbreviations: ABA, abscisic acid; APX, ascorbate peroxidase; Cd, cadmium; cDNA, complementary DNA; DAB, 3,3-diaminobenzidine; DEGs, differentially expressed genes; ETH, Ethylene; FDR, false discovery rate; FPKM, expected number of fragments per kilobase of transcript sequence per million base pairs sequenced; GA, gibberellin; GO, Gene Ontology; GPX, glutathione peroxidase; GR, glutathione reductase; H_2O_2, hydrogen peroxide; IAA, indole-3-acetic acid; JA, jasmonic acid; KEGG, the Kyoto Encyclopedia of Genes and Genomes; N bases, unknown bases; NBT, Nitro blue tetrazolium; NR, Non-redundant nucleic acid database; POD, peroxidase; qRT-PCR, quantitative real-time polymerase chain reaction; RNA, ribonucleic acid; ROS, reactive oxygen species; SA, Salicylic acid; SOD, superoxide dismutase; STEM, Short Time-series Expression Miner; Swiss-Prot, Swiss-port protein sequence database; *T. hispida*, *Tamarix hispida* Willd; UGTs, uridine diphosphate glucosyltransferases.

INTRODUCTION

Heavy metal pollution in soil has become a worldwide problem. It not only inhibits crop growth and reduces yield and quality but also poses a considerable threat to human health (Wu et al., 2007). The heavy metal cadmium is a biologically non-essential element. Cd is highly toxic and migratory with a biological half-life of 10 to 30 years (Suwazono et al., 2009). Therefore, Cd can easily enter the human body through the food chain and accumulate in the body, causing injury to the kidneys, lungs, liver, testicles, brain, bones, and blood system (Hamada et al., 1991).

High concentration of Cd is toxic to plants, but some plants grow under high Cd stress without exhibiting toxic effects. Approximately 10–33% of *Arabidopsis halleri* subsp. *gemmifera* can accumulate more than 100 mg kg^{-1} of Cd in contaminated soil (Bert et al., 2002). Yang et al. (2004) showed no reduction in shoot and root dry matter yields when *S. alfredii* were grown at Cd levels of 200 μmol L^{-1} in nutrient solution. Under natural conditions, the aboveground part of *Noccaea caerulescens* can accumulate up to 164 mg kg^{-1} Cd (Baker et al., 1994). *T. praecox* is a hyperaccumulator plant of zinc, cadmium and lead, and the aboveground part can accumulate up to 5,030 mg kg^{-1} of Cd (Vogel-Mikus et al., 2005). Under Cd pollution level of 25 mg kg^{-1}, the Cd content in stems and leaves of *S. nigrum* exceeded 100 mg kg^{-1}, and it was greater in shoots than in roots (Wei et al., 2005). These studies suggested that these super enriched plants with good cadmium tolerance can provide the theoretical basis for the study of plant remediation of Cd-contaminated soil.

At present, it is generally believed that accumulation of Cd in plants is primarily reflected in two aspects. On the one hand, at the cellular level, Cd primarily accumulates in the vacuoles and apoplasts of plants. On the other hand, at the organ level, this process is manifested in the epidermal cells, subepithelial cells and epidermal hairs of plants. According to Küpper et al. (2000), mustard mesophyll cells are important sites for Cd accumulation. In addition, Salt and Wagner (1993) reported deposit of large amount of Cd in the leaf epidermis and epidermis hairs in mustard. Recently, researchers found that the vacuolar membrane of rapeseed and *A. thaliana* play an important role in regulating the ion channel protein activity of NO_3^- and Cd, and the vacuolar compartmentalization and cell wall fixation of Cd may be the main physiological reasons for the difference in Cd toxicity resistance between Cd-resistant cultivar Z11 and Cd-sensitive cultivar W10 of rapeseed, which provides a means of synergistically improving the NUE and Cd toxicity of rapeseed (Zhang et al., 2019b). In *A. thaliana*, the defensive protein AtPDF2.5 may chelate cytoplasmic Cd and mediate its efflux, promote Cd accumulation in apoplasts, and regulate plant detoxification and accumulation of Cd (Luo et al., 2019). Therefore, accumulation of Cd in cell walls, vacuoles, epidermal cells or epidermal hair is likely to be one of the ways in which plants achieve detoxification.

In recent years, with of availability of transcriptional data, a growing body of knowledge regarding the genetic basis underlying Cd stress physiological processes has greatly increased our understanding of the molecular mechanism of Cd transcription and toxicity in some Cd hyperaccumulating plants, such as *A. halleri* (Herbette et al., 2006), *Brassica juncea* (Farinati et al., 2010), *S. alfredii* (Gao et al., 2013) and *Noccaea caerulescens* (Halimaa et al., 2014; Milner et al., 2014). At the same time, the molecular mechanisms of Cd stress on some cultivated plants such as cabbage (*Brassica oleracea subsp. capitata f. alba*) (Bączek-Kwinta et al., 2019), pea (*Pisum sativum* L.) (Rodriguez-Serrano et al., 2009), barley (*Hordeum vulgare* L.) (Cao et al., 2014), rice (*Oryza sativa* L.) (Oono et al., 2014), tobacco (*Nicotiana tabacum* L.) (Martin et al., 2012), ramie (*Boehmeria nivea* L.) (Liu et al., 2015) and pakchoi (*Brassica chinensis* L.) (Zhou et al., 2016) were studied.

Tamarix hispida is a woody halophyte that grows in arid and semiarid regions. In a previous study, the transcriptome of *T. hispida* treated with $NaHCO_3$ was constructed and analyzed to detect the response of *T. hispida* to alkaline treatment (Wang et al., 2013). Some transcription factors, including *ThNAC7*, *ThCRF1*, *ThZFP1*, and *ThbHLH1*, involved in the process of reducing ROS to confer salt or osmotic tolerance in transgenic plant have been cloned (Zang et al., 2015; Ji et al., 2016; Qin et al., 2017; He et al., 2019). There were studies in which multiple *T. hispida* genes enhanced tolerance to Cd. For example, the transfer of the metallothionein gene *ThMT3,* increased resistance to Cd in transgenic tobacco and yeast (Yang et al., 2011; Zhou et al., 2014). Overexpression of vacuolar membrane H^+-ATPase c subunit gene *ThVHAc1* improved Cd tolerance of *Saccharomyces cerevisiae*, *Arabidopsis*, and *T. hispida*. *ThWRKY7,* a possible upstream gene of *ThVHAc1*, exhibited similar expression patterns as *ThVHAc1* under $CdCl_2$ treatment and improved Cd tolerance in *T. hispida* (Gao et al., 2011; Yang et al., 2016).

Investigating transcriptomic response of Cd-stressed leaves would be particularly useful for furthering the genetic improvement of *T. hispida* to Cd stress. To elucidate the initial perception mechanism in response to Cd stimuli in *T. hispida* leaves, we examined gene expression changes at different time points and identified Cd-specific regulatory networks. This study helps to elucidate the mechanism of Cd tolerance in *T. hispida* and provides a useful reference for further exploration in woody plants and used for remediation of heavy metals (Cd) from contamination soils.

MATERIALS AND METHODS

Plant Materials and Cadmium Treatments

Seeds of *T. hispida* sourced from The Turpan Desert Botanical Garden (Xinjiang, 293 China) were germinated in plastic pots containing a mixture of turf peat and sand (1:1 v/v) under constant photoperiod conditions (14/10 h light/dark) with a light intensity of 1,500∼2,000 lx at temperature (24 ± 1°C). After culturing for 3 months in a greenhouse, at least 800 healthy seedlings of similar size (9 cm in height) were selected for Cd treatment. Based on preliminary test results (Gao et al., 2011), 150 μmol L^{-1} $CdCl_2$ was used to irrigate the seedlings. At the same time, samples irrigated with fresh water were treated

as control. After 24, 48 or 72 h treatment (each treatment contained three separate repeats with at least 200 seedlings), the leaves were washed with clean water and frozen in liquid nitrogen immediately, after which they were stored at −80°C for subsequent experiments. Each sample contains three replicates.

Determination of Cadmium Concentration and H_2O_2-Related Physiological Indices

To detect the Cd concentration of the samples, leaves from samples (each sample contains at least 20 seedlings) containing three replicates for each control, 24, 48 and 72 h were dried at 72°C to a constant weight and then digested with HNO_3. Subsequently, the Cd ion content was determined using ICP-OES 5110 VDV (Agilent Instruments Inc., CA, United States) (Gdbrijel et al., 2009; Han et al., 2016). The H_2O_2 content was detected by a hydrogen peroxide assay kit (Nanjing Jiancheng Bioengineering Institute), and detailed operating procedures were carried out according to the manufacturer's instructions. At the same time, hydrogen peroxide (H_2O_2) content in the leaves after Cd stress was detected by 3,3-diaminobenzidine (DAB) staining. Briefly, the above mentioned samples were placed in PBS (pH 7.0) solution containing 1 mg mL^{-1} DAB and treated in the dark for 12 h at 37°C. After exposure for 1 h, samples were decolorized with ethanol, and finally, the seedlings were photographed (Daudi and O'Brien, 2012).

Ribonucleic Acid Extraction, Sequencing and *de novo* Assembly

Total RNA was extracted from the leaf tissues of each sample by CTAB method (Jiang and Zhang, 2003). The degree of degradation of RNA samples was verified by RNase-free agarose gel electrophoresis. The RNA concentration was detected by qubit, and the RNA integrity was accurately detected with an Agilent 2100 Bioanalyzer. Equal quantities of high-quality RNA from the samples was used for the subsequent RNA sequencing.

Complementary DNA libraries were constructed for each of the samples and sequenced on the Illumina HiSeq 2000 platform (Illumina Inc., CA, United States). The original sequenced reads or raw reads containing reads with adapters or low quality were filtered to obtain clean reads. Specifically, reads with adapters were removed, unknown bases (N bases) over 10% and/or low-quality reads (the number of bases with a mass value of Qphred $\leq$ 20 accounts for more than 50% of the total reads) were removed from each data. Then, the clean reads of the twelve samples with high quality were spliced to construct unique sequences as the reference sequences using the Trinity package (Grabherr et al., 2011). The quality of transcripts were estimated by the value of FPKM (expected number of fragments per kilobase of transcript sequence per million base pairs sequenced).

Normalization of Gene Expression Levels and Identification of Differentially Expressed Genes

The clean reads of each sample were remapped to reference sequences using RSEM software (Li et al., 2009). RSEM counts the results of the bowtie comparison and further obtains the number of read counts for each sample that was aligned to each gene and performed FPKM conversion to analyze gene expression levels. For genes with more than one alternative transcript, the longest transcript was selected to calculate the FPKM.

To infer transcriptional changes over time under Cd stress conditions, differentially expressed genes (DEGs) after 24, 48, and 72 h of Cd treatment were identified by comparing the expression levels with control. The false discovery rate (FDR) was calculated to adjust the threshold of *p*-value to correct for multiple testing (Rajkumar et al., 2015). Transcripts with a minimal four-fold difference in expression ($|\log_2 \text{Ratio}| \geq 2$) and an FDR ≤ 0.001 were considered differentially expressed between two time points (Audic and Claverie, 1997). For convenience, DEGs with higher expression levels at 24, 48, and 72 h than control were denoted as "upregulated," whereas those with the opposite were denoted as "downregulated." At the same time, Venn diagrams of these differentially expressed genes were made to distinguish the differences between them using online software[1].

Short Timeseries Expression Miner (STEM) version 1.3.8 was used to analyze expression pattern (Ernst and Bar-Joseph, 2006). To further explore the temporal expression patterns, K-means clustering was applied to the identified DEGs. The DEGs belonging to the same cluster have expression patterns similar to each other. For each genotype, the clustering profiles of DEGs with *p*-values < 0.05 were considered to be significantly different from the reference group.

Validation of Differentially Expressed Genes With Quantitative Real-Time Polymerase Chain Reaction

Eight genes were randomly selected for quantitative real-time RT-PCR (qRT-PCR) to determine the expression patterns revealed by RNA sequencing. RNA was extracted from *T. hispida* leaves of Cd treated samples (control, 24, 48, and 72 h). The PrimeScript^{TM} RT reagent Kit (Takara) was used for first-strand cDNA synthesis. Primers for qRT-PCR analysis are listed in **Supplementary Table 1**. β-*actin* (FJ618517) was used as an internal control (Vandesompele et al., 2002). qRT-PCR was performed using a real-time PCR instrument (qTOWER 2.0) (analytik jena, Jena, Germany). The reaction mixture (20 μl) consisted of 10 μl of TransStar^{R} Top Green qPCR SuperMix (TRANS), 2 μl of cDNA template (equivalent to 500 ng of total RNA), 0.5 μmol L^{-1} of forward and reverse primers. The reaction procedure was as follows: one cycle at 95°C for 3 min, followed by 45 cycles of 95°C for 30 s, 58°C for 15 s, 72°C for 30 s. Three independent replicates were performed to ensure the reproducibility of results. Expression levels of these genes were determined according to the $2^{-\Delta\Delta(Ct)}$ (Livak and Schmittgen, 2001). Then, qRT-PCR results and sequencing results were analyzed together to verify the accuracy of the sequencing results.

[1]http://bioinfogp.cnb.csic.es/tools/venny/index2.0.2.html

Sequence Annotation, Functional Classification, and Biological Pathway Analysis

The unigenes were analyzed for functional annotation and functional classification. After splicing, the unigene sequences were compared with the protein database by Blast, and the annotation included the NCBI Non-redundant nucleic acid database (NR) and the Swiss-port protein sequence database (Swiss-Prot) with a threshold of *e*-value < 0.000 01. The GO annotation information was obtained by Blast2GO analysis based on the NR annotation information (Conesa et al., 2005) and classification of all unigenes was performed by WEGO (Ye et al., 2006). At the same time, the Kyoto Encyclopedia of Genes and Genomes (KEGG) was performed to further characterize the metabolic pathways and biological functions of DEGs in the transcriptome.

Cloning of *ThUGT* Gene in Abscisic Acid Signaling Pathway and It's Cadmium Resistance Function Analysis

The abscisic acid (ABA) signaling pathway gene *ThUGT* was successfully cloned using F: ATGGCTTCAGAATCCCATGAT and R: TTAGTTAATCCGGCCAC CTTT as primer, then an overexpression vector (*ThUGT*-pROKII) was constructed using F: GCTCTAGAATGGCTTCAGAATCCCATGAT and R: CGGGGTACCTTAG TTAATCCGGCCACCTTT following Wang et al. (2020). Afterward, the overexpression vector strain was transiently transformed into *T. hispida* according to the method of Ji et al. (2014) and treated with 100 μmol L^{-1} $CdCl_2$ for 24 h, and the pROKII empty vector transiently transformed seedlings were used as control. Each treatment contained three separate repeats with at least 45 seedlings. Then the expression levels of *ThUGT* in transient overexpression *T. hispida* and control seedlings were analyzed by qRT-PCR. At the same time, the Cd content, ABA content (SenBeiJia Biological Technology Co., Ltd., Nanjing, China) and H_2O_2 content (Nanjing Jiancheng Bioengineering Institute) of the samples were determined. DAB, NBT and Evans Blue staining were also performed on each sample (Zhang et al., 2011). Each experiment was repeated three times.

RESULTS

Changes of Cadmium and H_2O_2 Concentrations in *Tamarix hispida* Subjected to Cadmium Stress

In this study, the Cd and H_2O_2 contents were detected after *T. hispida* was treated with 150 μmol L^{-1} $CdCl_2$. The results showed that Cd concentration showed an upward trend after Cd stress compared with control. Especially at 72 h,

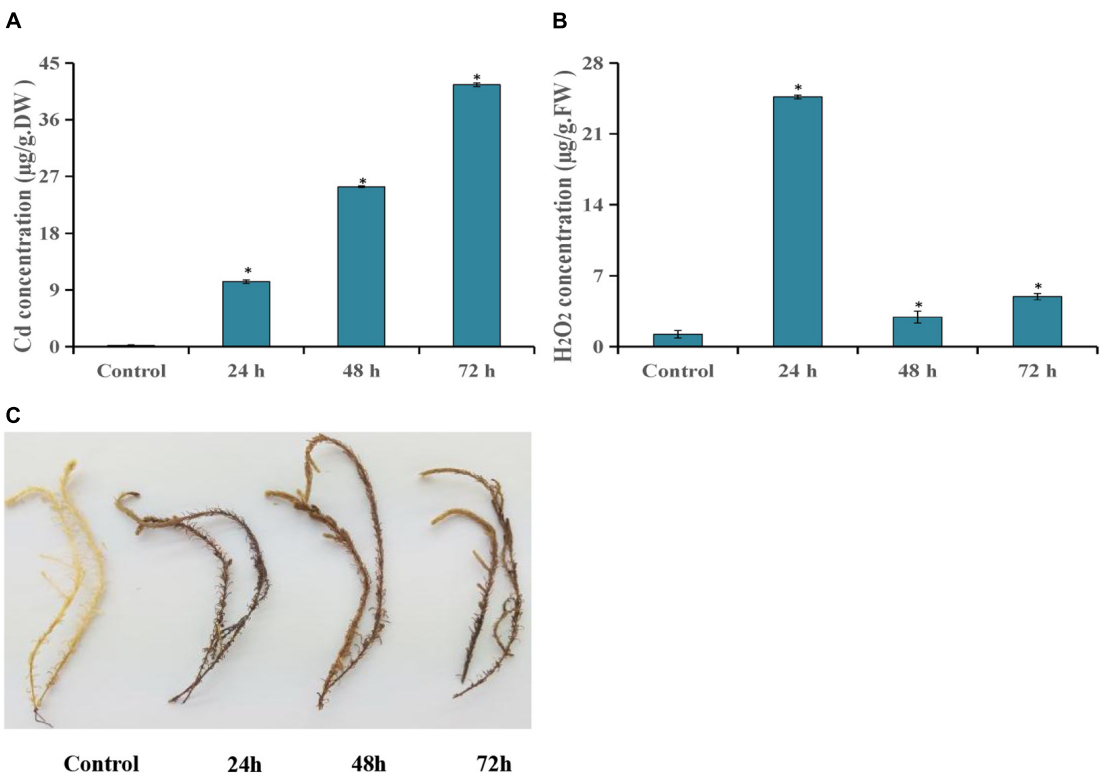


FIGURE 1 | Cd stress analysis of *T. hispida*. The leaves Cd **(A)** and H_2O_2 **(B)** concentrations in *T. hispida* at control or under 150 μmol L^{-1} $CdCl_2$ treated for 24, 48, or 72 h. **(C)** DAB staining of *T. hispida* leaves under control or after cadmium stress. * ($P < 0.05$) indicate signification difference compared with control.

the concentration reached a peak with an absolute value of 41.6 mg kg^{-1}, which was 227 times that of control (**Figure 1A**). The H_2O_2 concentration was also significantly increased after Cd stress and reached the highest level at 24 h (20.2 times that of control) (**Figure 1B**). At the same time, the dense DAB staining of the leaves of *T. hispida* was observed at 24 h (**Figure 1C**).

Pairwise Comparisons of Transcriptome Between Control and Cadmium Stressed Leaves

Using a cutoff of four-fold difference in gene expression, 3,505, 3,983, and 4,443 upregulated genes and 3,273, 4,299, and 4,158 downregulated genes were identified at 24, 48, and 72 h, respectively, compared with those in the control (**Figure 2A**). Among these genes, 2,673, 3,058, and 3,077 DEGs at 24, 48, and 72 h had no similar sequences in the NR database, respectively (**Supplementary Table 2**). Interestingly, the expression of many DEGs changed significantly only at a certain time point. For example, there were 5,012 DEGs at 72 h and there were 1,069 DEGs common to all three treatment time points compared with the control (**Figure 2B**).

To validate the expression data obtained from RNA sequencing, eight genes were randomly selected from the identified DEGs to perform qRT-PCR analysis. The results showed a strong correlation between the RNA sequencing and qRT-PCR data (**Figure 3**), which supports the reliability of the expression results generated by RNA sequencing.

Through GO and KEGG pathway enrichment analysis, the function of the DEGs were characterized. GO annotation suggested that biological processes and molecular functions related to ROS functions and biosynthetic and metabolic processes were enriched among the DEGs at different time points (**Figures 4A,B**). In biological processes, biosynthetic, and metabolic processes, DEGs were enriched at all three time points (**Figure 4A**). In addition, the molecular function term "phenylalanine ammonia-lyase activity" was enriched at 24 and 72 h. The term "methylenetetrahydrofolate reductase NAD(P)H activity" occurred at all three time points (**Figure 4B**). These results indicated that these genes or proteins participate in hormone and ROS metabolism play crucial roles in the *T. hispida* response to Cd stress.

The Kyoto Encyclopedia of Genes and Genomes pathway enrichment analysis results showed that 90, 91, and 105 pathways were categorized from the pairwise comparisons between 24 h vs. control, 48 h vs. control, and 72 h vs. control, respectively. Based on these results, the DEGs involved in the biosynthetic pathways of seven hormones and one pathway for "plant hormone signal transduction" were enriched (**Figure 4C**). At the same time, we observed that the number of genes involved in six hormone synthetic pathways were the highest at 72 h, whereas brassinosteroid biosynthesis pathways peaked at 24 h. Interestingly, six hormone synthetic pathways were included in the top 30 pathways from 24 h vs. control (**Table 1**). In the top 20 pathways in 24 h vs. control, five hormone synthesis pathways were found. These results demonstrated that the expression of genes involved in hormone synthesis may play an important role in *T. hispida* response to Cd stress.

Analysis of Temporal Expression Pattern of Genes

Hierarchical clustering produced six groups with similar expression trends to those of K-means clustering. Specifically, 1,069 qualifying genes were categorized into six groups (referred to as G1, G2, G3, G4, G5, and G6), comprising 240, 110, 384, 142, 164, and 29 genes, respectively (**Figure 5**), of which the G2, G3, G4, G5, and G6 groups mainly showed an upregulated trend under Cd treatment.

The expression of G1 group showed a downregulation trend, and reached its lowest value at 48 h, then slightly increased at 72 h. The G2 group mainly reached a peak at 24 h, then showed a downward trend and slightly increased at 72 h. The G3 group showed an upregulation trend at the beginning

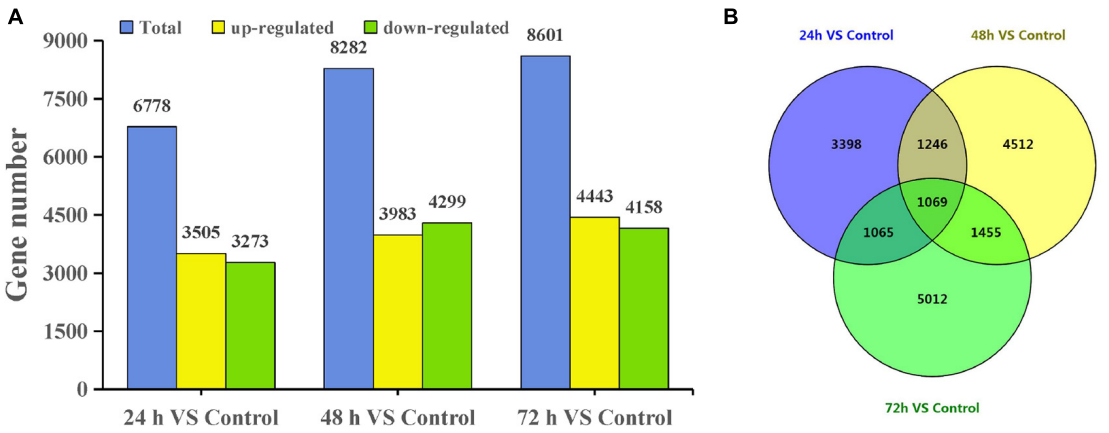


FIGURE 2 | DEGs of *T. hispida* under cadmium treatments. **(A)** Gene number analysis of DEGs between the cadmium treated transcriptomes compared with the control. **(B)** Venn diagrams of these DEGs.

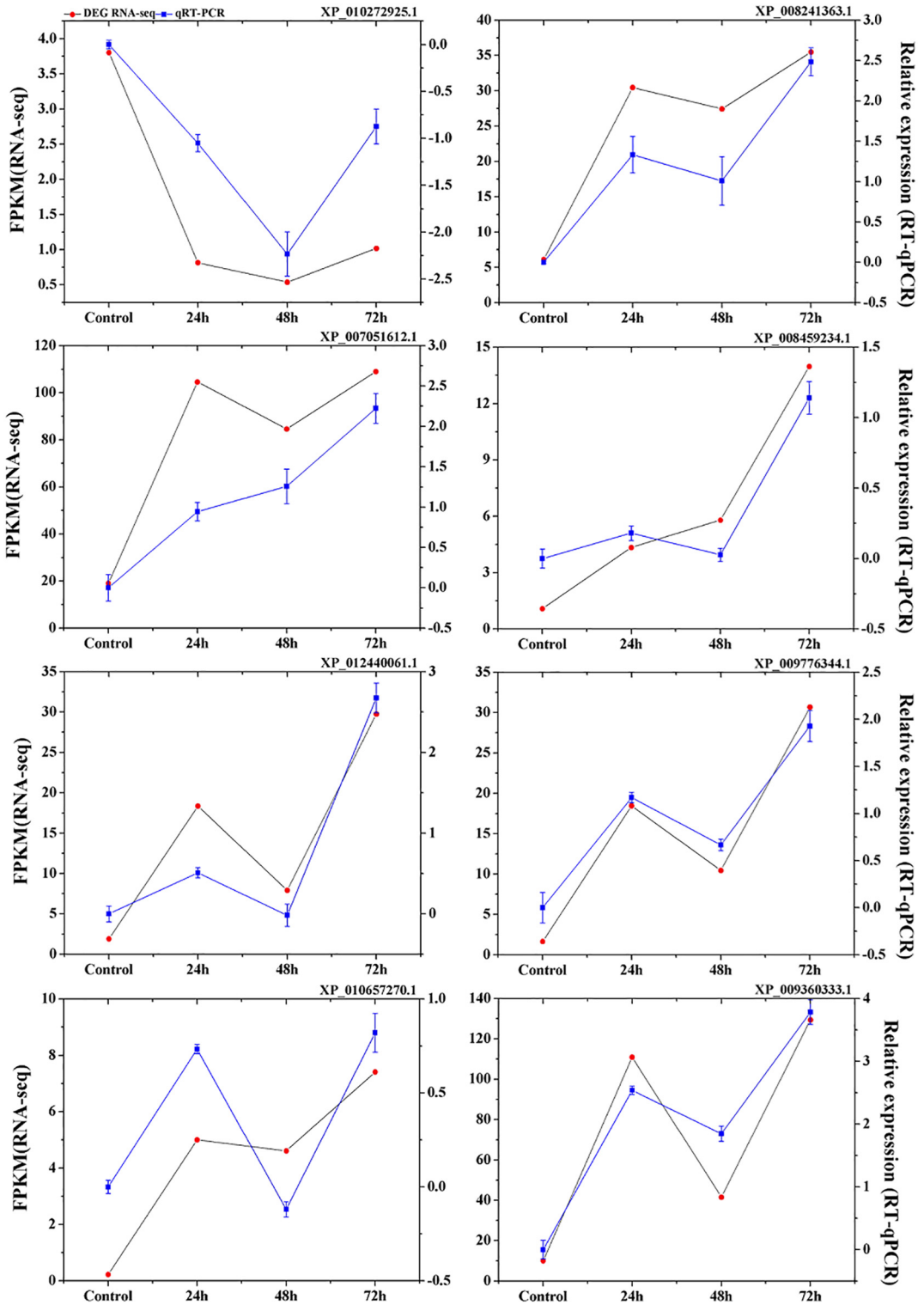


FIGURE 3 | Verification of eight selected DEGs by qRT-PCR. Comparison of RNA-seq data (Red point) with qRT-PCR data (Blue point). The normalized expression level (FPKM; expected number of Fragments Per Kilobase of transcript sequence per Millions base pairs sequenced) of RNA-seq is indicated on the *y*-axis to the left. The relative qRT-PCR expression level of selected DEGs is shown on the *y*-axis to the right. β-*actin* was used as the internal control. Three biological replicates were used.

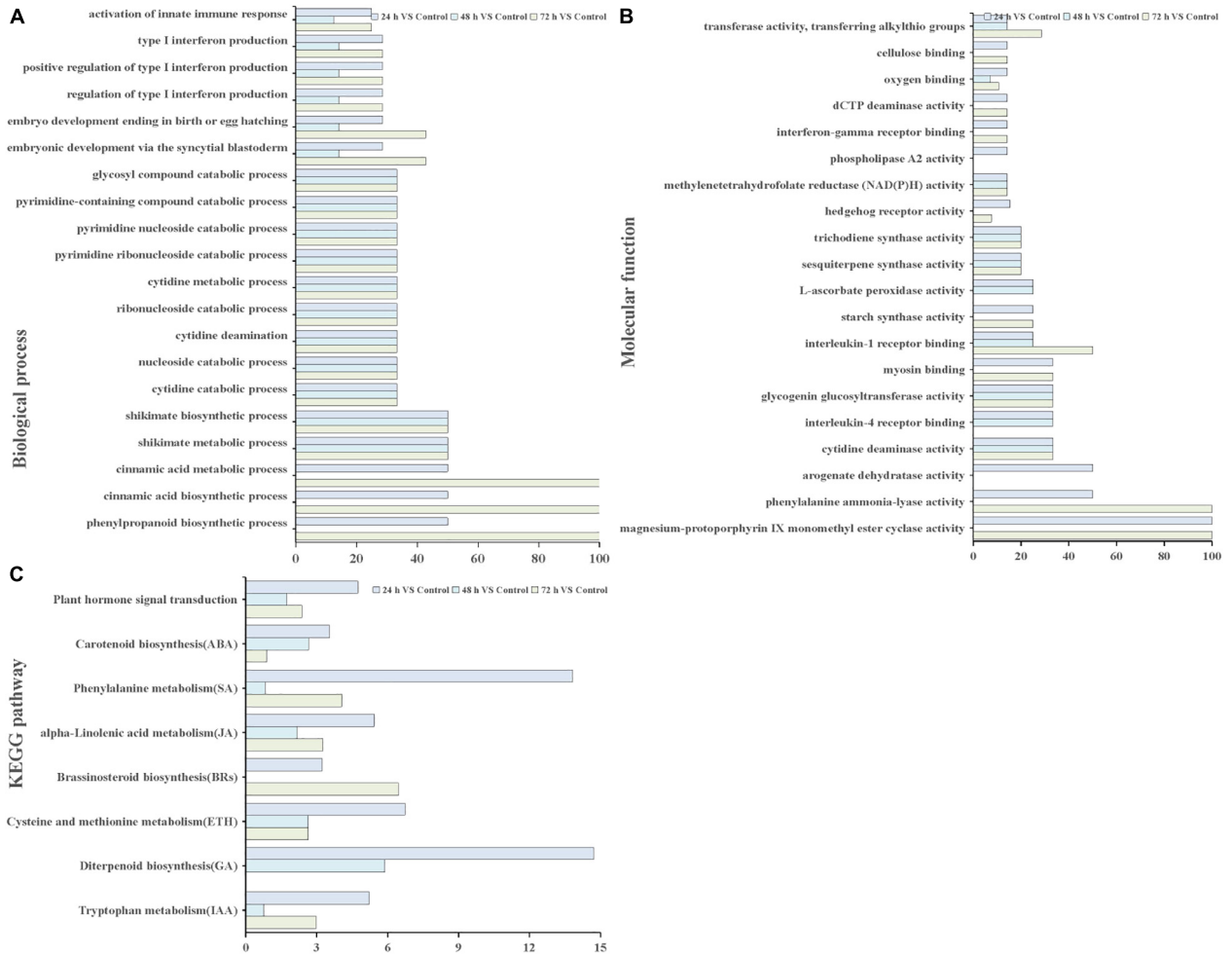


FIGURE 4 | GO terms and KEGG pathways involved in ROS and hormones production analysis. GO analysis of biological process terms **(A)**, molecular function terms **(B,C)** KEGG pathways involved in hormones between the cadmium treated transcriptomes compared with the control. The *x*-axis in **(A,B)** indicates the percentage of DEG snumbers vs. background gene numbers in each GO term. The *x*-axis in **(C)** indicates the percentage of DEGs numbers vs. background gene numbers in each KEGG pathway.

(24 h) followed by stable expression in the subsequent stage. The G4 group showed a continuous increasing trend at 24 h and 48 h. Interestingly, the G5 group showed an increasing expression trend during the whole stage and peaked at 72 h. The G6 group showed an upregulation trend at 24 and 48 h and a slight decrease at 72 h. From these results, we can assume that the genes in G3 and G5 group can be rapidly induced by Cd stress, and the genes in G2, G4, and G6 group showed a time-dependent trend in the process. In contrast, the genes in G1 group showed inhibition trend in response to cadmium stress.

In the GO functional annotation analysis, the number of genes in each group based on their biological pathway, molecular function and cellular component were counted. The results showed that all selected DEGs were involved in 282 biological pathways. Among them, the number of DEGs involved in protein binding, ATP binding, and DNA binding were 88, 52, and 45, respectively (**Table 2**).

In total, the K-means/hierarchical clustering and GO function annotation results indicated that the genes related to hormones were significantly affected during Cd stress in *T. hispida* leaves. Hence, the DEGs involved in the metabolism of hormones and their signaling pathways were further explored systematically.

Differentially Expressed Genes Involved in Hormone Biosynthetic Pathways

To further explore the genes involved in hormone biochemical pathways following Cd stress treatment, the DEGs involved in the seven hormone biosynthesis or metabolism KEGG pathways (ABA, ETH, IAA, SA, GA, BRs, and JA) were analyzed (**Table 1**). The number of genes in the ABA, SA, ETH, IAA, GA, and JA biosynthetic or metabolism pathways primarily increased, while BR biosynthesis showed decreasing trends (**Figure 4C**).

TABLE 1 | Top 30 KEGG pathways based on the percentage of DEGs in 24 h vs. control.

Term	Total	24 h vs. control		48 h vs. control		72 h vs. control	
		DEGs	%	DEGs	%	DEGs	%
Glucosinolate biosynthesis	7	2	28.6	1	14.3	2	28.6
Photosynthesis – antenna proteins	48	10	20.8	1	2.1	10	20.8
Cutin, suberin, and wax biosynthesis	77	7	9.1	2	2.6	9	11.7
Brassinosteroid biosynthesis (BRs)	31	2	6.5	0	0.0	1	3.2
Flavonoid biosynthesis	73	4	5.5	1	1.4	11	15.1
Phenylpropanoid biosynthesis	372	19	5.1	12	3.2	49	13.2
Stilbenoid, diarylheptanoid, and gingerol biosynthesis	61	3	4.9	0	0.0	10	16.4
Cyanoamino acid metabolism	161	7	4.4	6	3.7	13	8.1
Phenylalanine metabolism (SA)	123	5	4.1	1	0.8	17	13.8
Nitrogen metabolism	104	4	3.9	4	3.9	11	10.6
Arachidonic acid metabolism	57	2	3.5	1	1.8	5	8.8
Phenylalanine, tyrosine, and tryptophan biosynthesis	206	7	3.4	8	3.9	9	4.4
alpha-Linolenic acid metabolism (JA)	184	6	3.3	4	2.2	10	5.4
Tryptophan metabolism (IAA)	134	4	3.0	1	0.8	7	5.2
Alanine, aspartate and glutamate metabolism	241	7	2.9	4	1.7	22	9.1
Ascorbate and aldarate metabolism	214	6	2.8	7	3.3	19	8.9
Pentose phosphate pathway	215	6	2.8	7	3.3	15	7.0
Amino sugar and nucleotide sugar metabolism	408	11	2.7	6	1.5	24	5.9
Cysteine and methionine metabolism (ETH)	342	9	2.6	9	2.6	23	6.7
Terpenoid backbone biosynthesis	190	5	2.6	3	1.6	4	2.1
Taurine and hypotaurine metabolism	76	2	2.6	0	0.0	7	9.2
Steroid biosynthesis	123	3	2.4	1	0.8	8	6.5
Plant hormone signal transduction	633	15	2.4	11	1.7	30	4.7
Butanoate metabolism	128	3	2.3	1	0.8	2	1.6
Photosynthesis	131	3	2.3	3	2.3	12	9.2
Carbon fixation in photosynthetic organisms	306	7	2.3	8	2.6	28	9.2
Plant-pathogen interaction	482	11	2.3	13	2.7	20	4.2
Glutathione metabolism	264	6	2.3	8	3.0	18	6.8
Pentose and glucuronate interconversions	221	5	2.3	4	1.8	13	5.9
Sphingolipid metabolism	133	3	2.3	4	3.0	5	3.8

Hormone-related KEGG pathways have a gray background.

Under Cd stress, 7 DEGs were identified in the ABA-related pathway, including ABA β-glucosyltransferase, zeta-carotene desaturase (PLN02487), isoprenoid biosynthesis enzymes (IBe), carotene beta-ring hydroxylase (PLN02738), antheraxanthin epoxidase/zeaxanthin epoxidase (PLN02927) and cytochrome P450. Three of them were induced, especially *ThUGT* (ABA β-glucosyltransferase), which appeared at 24 and 72 h with 3.7 and 4.2-fold increased, respectively. In contrast, the expression of PLN02487, PLN02738, PLN02927, and cytochrome P450 showed reduced expression trend.

Thirty DEGs in ethylene biosynthetic pathway were detected. Among the metabolic processes, there were nine beta-eliminating lyase genes, four hypothetical proteins, two 1-aminocyclopropane-1-carboxylate oxidases and two 5-methyltetrahydropterin glutamate homocysteine methyltransferases. The expression of these 17 genes were upregulated. In contrast, the expression of genes involved in homocysteine S-methyltransferase, 1-aminocyclopropane-1-carboxylate synthase, ARD/ARD' family and S-adenosyl-methionine synthase were downregulated.

Hormones content, antioxidant activities, and downstream signals are induced in response to environmental stresses to control cell stability and mitigate the negative effects of ROS (Abdullah et al., 2021; Heidari et al., 2021). These results indicated that the key regulatory components of the biosynthetic pathways in ABA and ethylene changed significantly during the *T. hispida* response to Cd stress. In addition, there are also many genes involved in BRs, JA, GA, IAA, and SA, also responded to Cd stress.

ThUGT-Overexpressing *Tamarix hispida* Increased Abscisic Acid Content and Cadmium Tolerance

To further explore the genes involved in hormone pathways following Cd stress treatment, the ABA-related KEGG pathway after each stress treatment point was analyzed. The results showed that ABA pathway-related genes mainly participated in phytoene, lycopene, zeaxanthin, abscisate and lutein biosynthetic processes (**Supplementary Figure 1**). *ThUGT*, a predicted

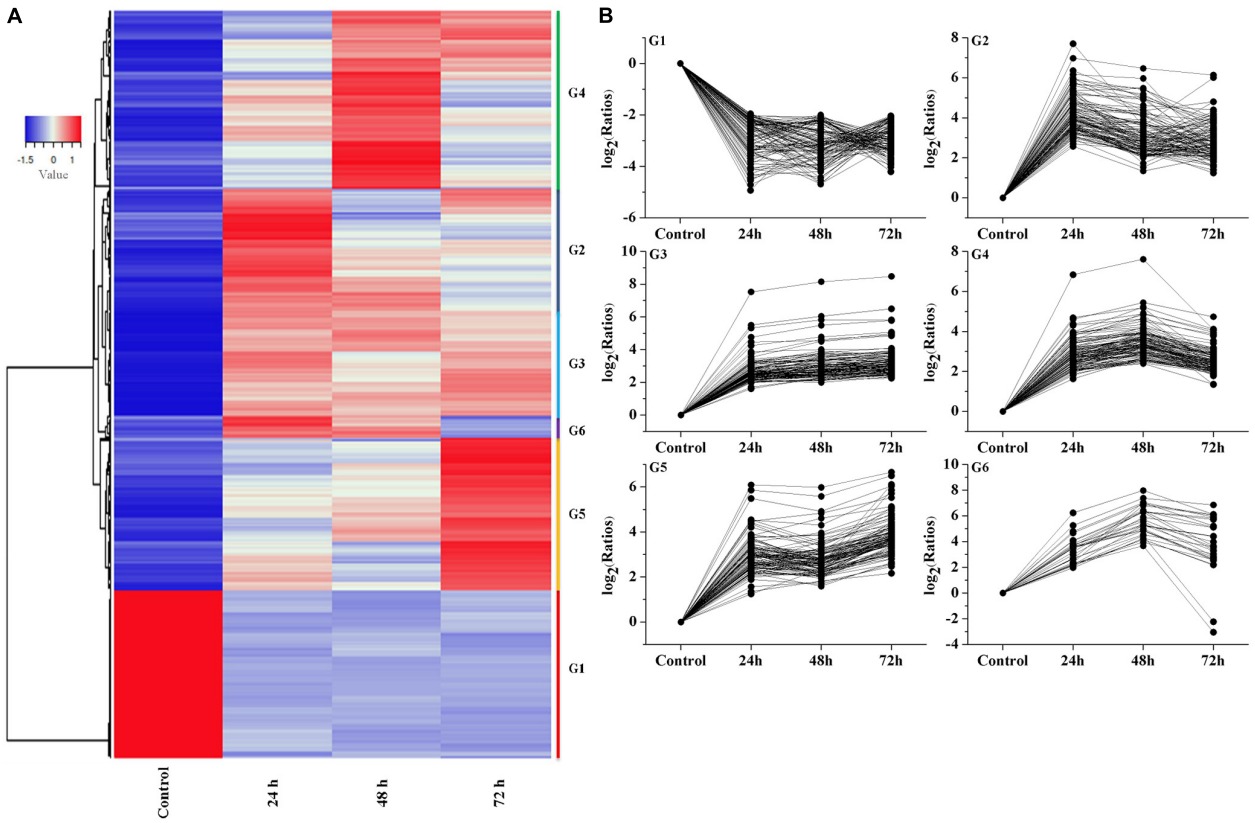


FIGURE 5 | Analysis of temporal expression pattern of genes. **(A)** Heat map analysis of DEGs. **(B)** Expressive trend pattern analysis of DEGs. The value of ratios take the base 2 logarithm were used to analysis each gene's expression trend in every group. For each gene, ratios = FPKM of the gene in sample/FPKM value of the gene in control.

ABA β-glucosyltransferase gene, was one of the DEGs in the ABA signaling pathway, and the expression level was induced (3.7 and 4.2 times of the control at 24 and 72 h, respectively). Therefore, *ThUGT* was transiently transformed into *T. hispida*. Pre-experimental results showed that in case of treatment with 100 μmol L^{-1} $CdCl_2$ after 24 h, the relative expression level of *ThUGT* gene had the largest difference with the normal conditions (**Figure 6**). The results of qRT-PCR showed that the expression level of *ThUGT* was significantly higher in the overexpressing plants than in the control plants (**Figure 7A**).

The staining analysis showed that the overexpressing plants were stained lighter after $CdCl_2$ stress treatment compared with the control plants (**Figure 7B**). The results obtained with physiological indicators showed that the ABA content in *ThUGT*-overexpressing *T. hispida* and the control plants was effectively increased under 100 μmol L^{-1} $CdCl_2$ treatment, while in the *ThUGT* transgenic plants, it increased less than in the control (**Figure 7C**). At the same time, the Cd ion content in both the overexpression plants and the control plants increased significantly after Cd treatment, but the Cd ion content in the overexpressing plants was significantly less than that of the control (**Figure 7C**). The H_2O_2 content in overexpression plants was lower than that of the control after Cd treatment (**Figure 7C**). These results suggested that the *ThUGT* transgenic plants eliminated more ROS and inhibited the absorption of Cd to a certain extent by *T. hispida* under cadmium stress, thereby enhancing the cadmium resistance of *T. hispida*.

DISCUSSION

When plants are under heavy metal stress, it can produce a series of responses to relieve the toxic effects of heavy metals. Some reports suggested that the production of glutathione, phytochelatin, and metal chelates or chaperones, which can bind to heavy metal ions and then transport them out of cells. In addition, antioxidant enzymes, such as superoxide dismutase (SOD), peroxidase (POD), glutathione reductase (GR), ascorbate peroxidase (APX), and glutathione peroxidase (GPX), were produced to scavenge oxygen free radicals generated by heavy metals (Emamverdian et al., 2015).

In this work, we used RNA-seq to explore the time course of the response mechanism in *T. hispida* under Cd stress. Our study showed that 114,292 unigenes were obtained after the transcriptome data were spliced and 36.3% of the unigenes were novel transcripts under Cd stress. Further analysis found that

TABLE 2 | The top 20 GO molecular function based on number of common DEGs.

Gene Ontology Molecular Function	DEG_item
protein binding	88
ATP binding	52
DNA binding	45
zinc ion binding	33
nucleic acid binding	31
oxidoreductase activity	28
protein kinase activity	24
metal ion binding	21
structural constituent of ribosome	20
RNA binding	18
transcription factor activity, sequence-specific DNA binding	17
calcium ion binding	14
transmembrane transporter activity	14
catalytic activity	13
hydrolase activity, hydrolyzing O-glycosyl compounds	12
GTP binding	10
electron carrier activity	9
GTPase activity	9
heme binding	9
nucleotide binding	8

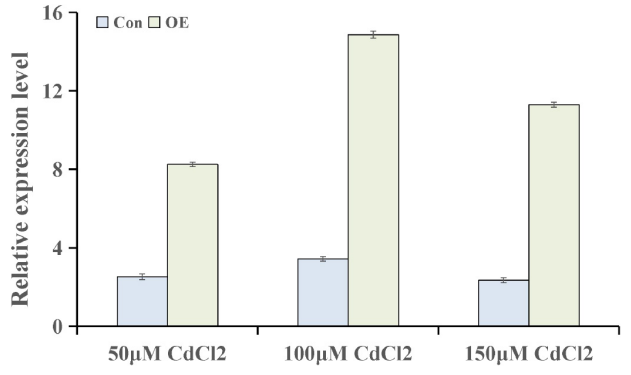


FIGURE 6 | The expression levels of *ThUGT* gene in transiently transformed *T. hispida* under different $CdCl_2$ stress conditions.

6,778 DEGs were detected at 24 h and then increased to 8,282 and 8,601 at 48 and 72 h, respectively (**Figure 2A**), which indicated that an increasing number of DEGs were induced or activated under Cd stress in *T. hispida*.

H_2O_2 is a product of aerobic metabolism of cells, and its production is increased under various stresses. It not only has the effect of damaging biological macromolecules and thereby harming cells, it is also an important signal molecule that induce the expression of a series of defense genes in the cell, improve the activity of protective enzymes to remove active oxygen, prevent its excessive accumulation under adversity conditions, and protect plants from damage (Moller, 2001; Mittler, 2002). Interestingly, Cd content increased with increasing treatment time (**Figure 1A**), while the H_2O_2 content peaked at 24 h (**Figures 1B,C**). At the same time, we also found that many studies reported H_2O_2 functions a secondary messenger during plant development and defense to abiotic stress (Avshalumov et al., 2003; Jiang et al., 2012; Khalili et al., 2014; Liu et al., 2016; Saxena et al., 2016). Therefore, we suggest that H_2O_2 may be serving as a secondary messenger to induce the expression of stress-related proteins in the early stage of cadmium stress and initiate the development of systemic acquired resistance in *T. hispida*. This plant synthesizes many ROS-clearing genes, thereby effectively eliminating excess H_2O_2 to help *T. hispida* tolerate Cd stress. Consistent with this property, the results of the DEGs at 24 h showed that 1/6 DEGs participated in the redox reaction, the transport of ions, and the synthesis of signal substances (**Supplementary Table 3**), which have important functions in plant stress resistance.

In previous studies, multiple transcription factors were reported to be involved in the scavenging of ROS or increasing SOD and POD activities in *T. hispida* to improve the salt tolerance or osmotic stress ability of plants (Zang et al., 2015; Ji et al., 2016; Qin et al., 2017; He et al., 2019). Many studies on transcriptome analyses of Cd-treated plants have found that it is mainly related to the pathways of "ROS-scavenging enzymes" (Guo et al., 2017; Gupta et al., 2017), "cell wall alternation and strengthening" (Fan et al., 2011; Wan and Zhang, 2012; Xu et al., 2015), "lipid oxidation" (Guo et al., 2017), "auxin biosynthesis and metabolism" (Yue et al., 2016), and "nitric oxide-mediated homeostasis" (Zuccarelli et al., 2017).

In *T. hispida*, the GO analysis showed in the biological pathway that "oxidation–reduction process" was the top enriched term, and "oxidoreductase activity" in the molecular function also ranked in the top 10 (**Figures 4A,B**). These results indicated that the expression of antioxidant and redox homeostasis-related genes play important role in response to Cd stress in *T. hispida*.

Abscisic acid, a widely known phytohormone involved in the plant response to abiotic stress, plays a vital role in mitigating Cd^{2+} toxicity in herbaceous species. Studies have found that when plants were exposed to Cd, endogenous ABA levels were increased in plant cells (Sharma and Kumar, 2002). Several other studies demonstrated that the application of ABA can reduce Cd accumulation in crops (Hsu and Kao, 2003, 2005; Uraguchi et al., 2009; Fan et al., 2014). Fan et al. (2014) reported that ABA treatment correlates with the downregulation of ABA-inhibited *IRON-REGULATED TRANSPORTER 1* (*IRT1*) to decrease Cd accumulation. Through interaction with MYB49, the ABI5 represses MYB49 binding to the downstream genes *bHLH38*, *bHLH101*, *HIPP22,* and *HIPP44*, which result in the inactivation of *IRT1* and reduced Cd uptake (Zhang et al., 2019a).

The results of KEGG pathway analysis showed that seven main phytohormones participate in Cd stress in *T. hispida.* Among the phytohormones, the input numbers vs. background gene numbers for ABA pathways increased over time (**Figure 4C**), with ratios of 0.9, 2.7, and 3.5% being observed at 24, 48, and 72 h, respectively. By GO function analysis, we found that seven DEGs were involved in the ABA signaling pathway, three of

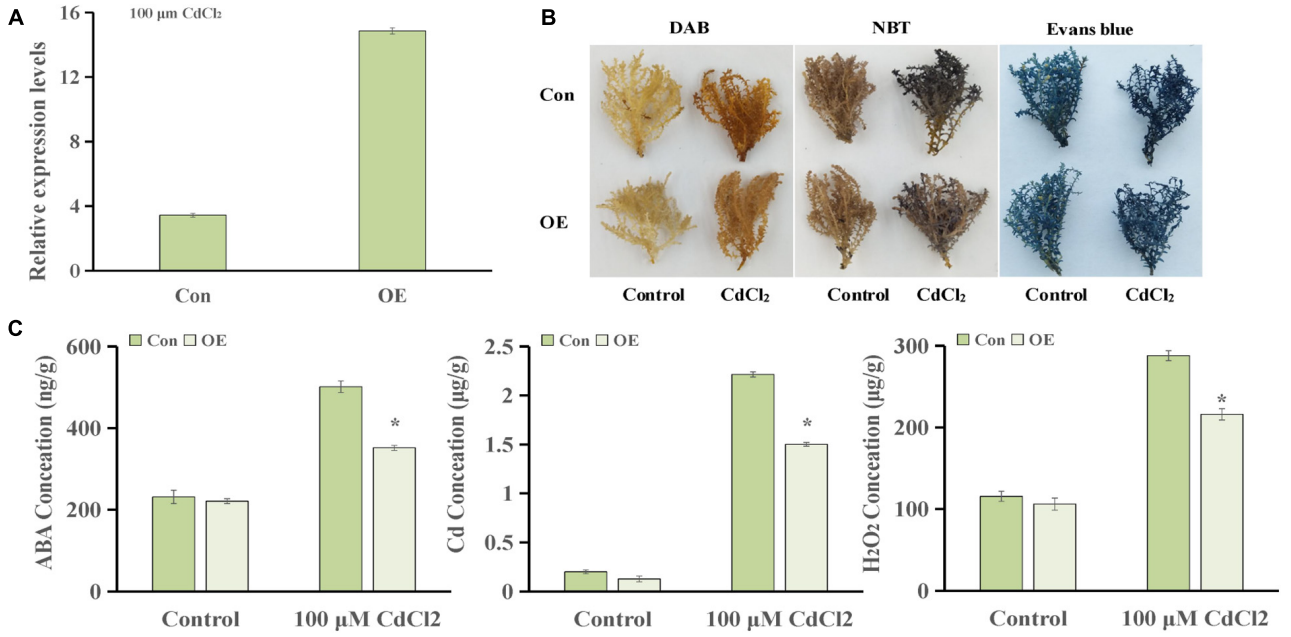


FIGURE 7 | Cadmium tolerance analysis of *ThUGT* gene. **(A)** qRT-PCR analysis **(B)** staining analysis and **(C)** physiological indicators determination of *ThUGT* transgenic *T. hispida* under 100 μmol L^{-1} $CdCl_2$ stress. * ($P < 0.05$) indicate signification difference compared with control.

which were upregulated (**Supplementary Figure 1**). In particular, the expression of *ThUGT* gene was clearly induced. The ABA content in transient overexpression *ThUGT T. hispida* plants was significantly increased after cadmium stress treatment, but it was less compared with the control (**Figure 7C**). Moreover, increase in the H_2O_2 content and Cd content was less in OE plants than the control after Cd stress (**Figure 7C**), which was consistent with the results reported in previous studies, indicating that the *ThUGT* gene has the ability to remove ROS under Cd stress and enhance the tolerance of *T. hispida* to Cd stress.

Uridine diphosphate-glucosyltransferases (UGTs) are a family of proteins involved in physiological responses to the inactivation of many glycosylation hormones (Kleczkowski et al., 1995). The overexpression of *UGT74E2* gene in *A. thaliana* can improve the resistance to drought and salt stress by regulating the ABA dynamic balance (Tognetti et al., 2010). ABA β-glucosyltransferase belongs to the UGT family and is a key enzyme in the ABA catabolism binding pathway. This enzyme plays an important role in maintaining the normal physiological level of ABA. The ABA β-glucosyltransferase gene *AtUGT71B6* in *A. thaliana* regulates intracellular ABA balance (Priest et al., 2006; Dong et al., 2014). The *Phaseolus vulgaris PvABAGT* gene regulates the ABA balance and stress response in adversity during bean development (Xu et al., 2002).

In *Beta vulgaris*, the ROS produced by a plasma membrane NADPH oxidase may act as a signal to induce *BvGT* (UGT family gene) expression after wounding and bacterial infiltration (Sepulveda-Jimenez et al., 2004). Oxidative stress and conditions that promote cell death could induce the expression of glucosyltransferase genes and produce transportable glucosides that function as ROS scavengers (Mazel and Levine, 2002). Therefore, we concluded that *ThUGT* may play a crucial role in the ABA conjugation pathway and in adaptation to Cd stress by inducing ROS scavengers to deduce H_2O_2.

CONCLUSION

The present study identified novel transcripts, gene structures, and DEGs in *T. hispida* under Cd stress. In total, 114,292 unigenes were identified. The transcripts identified in this study will serve as a valuable genomic resource for future studies. Among these genes, a large number of genes were related to ROS clearance and hormone signals, which may facilitate the analysis of ABA signaling pathways. The overexpression of related a gene *ThUGT* reduced the accumulation of Cd in *T. hispida* under high $CdCl_2$ stress. These results will help to establish a foundation for future research to improve cadmium tolerance in *T. hispida* and other plants.

RNA-Seq DATA SUBMITTED TO PUBLIC DATABASE

The data presented in the study are deposited in the NCBI repository, accession number PRJNA795701 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA795701).

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

P-LW and C-QG conceived and designed the experiments and wrote the manuscript. P-LW and Y-YW performed the experiments. P-LW and X-JL analyzed the data. All authors provided editorial advice.

FUNDING

This work was supported by the National Natural Science Foundation of China (No. 31370676), the Province in Heilongjiang Outstanding Youth Science Fund (JC2017004), and Heilongjiang Touyan Innovation Team Program (Tree Genetics and Breeding Innovation Team).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.843725/full#supplementary-material

Supplementary Figure 1 | Signaling pathways involved in ABA. The genes associated with ABA were marked in different colors. Red: up-regulated genes. Blue: down-regulated genes.

REFERENCES

Abdullah, Faraji, S., Mehmood, F., Malik, H. M. T., Ahmed, I., Heidari, P., et al. (2021). The GASA gene family in cacao (*Theobroma cacao*, Malvaceae): genome wide identification and expression analysis. *Agronomy* 11, 1–20.

Audic, S., and Claverie, J. M. (1997). The significance of digital gene expression profiles. *Genome Res.* 7, 986–995. doi: 10.1101/gr.7.10.986

Avshalumov, M. V., Chen, B. T., Marshall, S. P., Pena, D. M., and Rice, M. E. (2003). Glutamate-dependent inhibition of dopamine release in striatum is mediated by a new diffusible messenger, H_2O_2. *J. Neurosci.* 23, 2744–2750. doi: 10.1523/JNEUROSCI.23-07-02744.2003

Bączek-Kwinta, R., Juzoń, K., and Borek, M. (2019). Photosynthetic response of cabbage in cadmium-spiked soil. *Photosynthetica* 57, 731–739.

Baker, B. J. M., Reeves, R. D., and HAjar, A. S. M. (1994). Heavy metal accumulation and tolerance in British populations of the metallophyte *Thlaspi caerulescens* J. & C. Presl (Brassicaceae). *New Phytol.* 127, 6l–68. doi: 10.1111/j.1469-8137.1994.tb04259.x

Bert, V., Bonnin, I., Saumitou-Laprade, P., Patrick de Laguérie, P.d, and Daniel Petit, D. (2002). Do arabidopsis halleri from nonmetallicolous populations accumulate zinc and cadmium more effectively than those from metallicolous populations. *New Phytol.* 155, 47–57. doi: 10.1046/j.1469-8137.2002.00432.x

Cao, F. B., Chen, F., Sun, H. Y., Zhang, G. P., Chen, Z. H., and Wu, F. B. (2014). Genome-wide transcriptome and functional analysis of two contrasting genotypes reveals key genes for cadmium tolerance in barley. *BMC Genomics* 15:611. doi: 10.1186/1471-2164-15-611

Conesa, A., Götz, S., García-Goìmez, J. M., Terol, J., Taloìn, M., and Robles, M. (2005). Blast2GO, a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674–3676. doi: 10.1093/bioinformatics/bti610

Daudi, A., and O'Brien, J. S. (2012). Detection of hydrogen peroxide by DAB staining in *Arabidopsis*. *Bio Protoc.* 2:e263.

Dong, T., Xu, Z. Y., Park, Y., Kim, D. H., Lee, Y., and Hwang, I. (2014). Abscisic acid uridine diphosphate glucosyltransferases play a crucial role in abscisic acid homeostasis in *Arabidopsis*. *Plant Physiol.* 165, 277–289. doi: 10.1104/pp.114.239210

Emamverdian, A., Ding, Y. L., Mokhberdoran, F., and Xie, Y. F. (2015). Heavy metal stress and some mechanisms of plant defense response. *Sci. World J.* 2015:756120. doi: 10.1155/2015/756120

Ernst, J., and Bar-Joseph, Z. (2006). STEM, a tool for the analysis of short time series gene expression data. *BMC Bioinformatics* 7:191. doi: 10.1186/1471-2105-7-191

Fan, J. L., Wei, X. Z., Wan, L. C., Zhang, L. Y., Zhao, X. Q., Liu, W. Z., et al. (2011). Disarrangement of actin filaments and Ca^{2+} gradient by CdCl2 alters cell wall construction in *Arabidopsis thaliana* root hairs by inhibiting vesicular trafficking. *J Plant Physiol.* 168, 1157–1167.

Fan, S. K., Fang, X. Z., Guan, M. Y., Ye, Y. Q., Lin, X. Y., Du, S. T., et al. (2014). Exogenous abscisic acid application decreases cadmium accumulation in *Arabidopsis* plants, which is associated with the inhibition of IRT1-mediated cadmium uptake. *Front. Plant Sci.* 5:721. doi: 10.3389/fpls.2014.00721

Farinati, S., DalCorso, G., Varotto, S., and Furini, A. (2010). The *Brassica juncea* BjCdR15, an ortholog of *Arabidopsis* TGA3, is a regulator of cadmium uptake, transport and accumulation in shoots and confers cadmium tolerance in transgenic plants. *New Phytol.* 185, 964–978. doi: 10.1111/j.1469-8137.2009.03132.x

Gao, C. Q., Wang, Y. C., Jiang, B., Liu, G. F., Yu, L. L., Wei, Z. G., et al. (2011). A novel vacuolar membrane H^+-ATPase c subunit gene (ThVHAc1) from *Tamarix hispida* confers tolerance to severalabiotic stresses in *Saccharomyces cerevisiae*. *Mol. Biol. Rep.* 38, 957–963. doi: 10.1007/s11033-010-0189-9

Gao, J., Sun, L., Yang, X. E., and Liu, J. X. (2013). Transcriptomic analysis of cadmium stress response in the heavy metal hyperaccumulator *Sedum alfredii* hance. *PLoS One* 8:e64643. doi: 10.1371/journal.pone.0064643

Gdbrijel, O., Davor, R., Zed, R., Marija, R., and Monika, Z. (2009). Cadmium accumulation by muskmelon under salt stress in contaminated organic soil. *Sci. Total Environ.* 407, 2175–2182. doi: 10.1016/j.scitotenv.2008.12.032

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883

Guo, Q., Meng, L., Zhang, Y. N., Mao, P. C., Tian, X. X., Li, S. S., et al. (2017). Antioxidative systems, metal ion homeostasis and cadmium distribution in *Iris lactea* exposed to cadmium stress. *Ecotoxicol. Environ. Saf.* 139, 50–55. doi: 10.1016/j.ecoenv.2016.12.013

Gupta, D. K., Pena, L. B., Romero-Puertas, M. C., Hernández, A., Inouhe, M., and Sandalio, L. M. (2017). NADPH oxidases differentially regulate ROS metabolism and nutrient uptake under cadmium toxicity. *Plant Cell Environ.* 40, 509–526. doi: 10.1111/pce.12711

Halimaa, P., Lin, Y. F., Ahonen, V. H., Blande, D., Clemens, S., Gyenesei, A., et al. (2014). Gene expression differences between *Noccaea caerulescens* ecotypes help identifying candidate genes for metal phytoremediation. *Environ. Sci. Technol.* 48, 3344–3353. doi: 10.1021/es4042995

Han, Y., Wang, S., Zhao, N., Deng, S., Zhao, C., Li, N., et al. (2016). Exogenous abscisic acid alleviates cadmium toxicity by restricting Cd^{2+} influx in *Populus euphratica* cells. *J. Plant Growth Regul.* 35, 827–837.

Hamada, T., Nakano, S., Iwai, S., Tanimoto, A., Ariyoshi, K., and Koide, O. (1991). Pathological study on beagles after long-term oral administration of cadmium. *Toxicol. Pathol.* 19, 138–147. doi: 10.1177/019262331019010 0208

He, Z., Li, Z., Lu, H., Huo, L., Wang, Z., Wang, Y., et al. (2019). The NAC protein from *Tamarix hispida*, ThNAC7, confers salt and osmotic stress tolerance by increasing reactive oxygen species scavenging capability. *Plants (Basel)* 8:221. doi: 10.3390/plants8070221

Herbette, S., Taconnat, L., Hugouvieux, V., Piette, L., Magniette, M. L. M., Cuine, S., et al. (2006). Genome-wide transcriptome profiling of the early cadmium response of *Arabidopsis* roots and shoots. *Biochimie* 88, 1751–1765. doi: 10.1016/j.biochi.2006.04.018

Heidari, P., Amerian, M. R., and Barcaccia, G. (2021). Hormone profiles and antioxidant activity of cultivated and wild tomato seedlings under low-temperature stress. *Agronomy* 11, 1–16.

Hsu, Y. T., and Kao, C. H. (2003). Role of abscisic acid in cadmium tolerance of rice (*Oryza sativa* L.) seedlings. *Plant Cell Environ.* 26, 867–874. doi: 10.1046/j.1365-3040.2003.01018.x

Hsu, Y. T., and Kao, C. H. (2005). Abscisic acid accumulation and cadmium tolerance in rice seedlings. *Physiol. Plant.* 124, 71–80. doi: 10.1111/j.1399-3054.2005.00490.x

Ji, X., Nie, X., Liu, Y., Zheng, L., Zhao, H., Zhang, B., et al. (2016). A bHLH gene from *Tamarix hispida* improves abiotic stress tolerance by enhancing osmotic potential and decreasing reactive oxygen species accumulation. *Tree Physiol.* 36, 193–207. doi: 10.1093/treephys/tpv139

Ji, X. Y., Zheng, L., Liu, Y. J., Nie, X. G., Liu, S. N., and Wang, Y. C. (2014). A transient transformation system for the functional characterization of genes involved in stress response. *Plant Mol. Biol. Rep.* 32, 732–739. doi: 10.1007/s11105-013-0683-z

Jiang, J. X., and Zhang, T. Z. (2003). Extraction of total RNA in cotton tissues with ctab-acidic phenolic method. *Acta Gossypii Sin.* 15, 166–167.

Jiang, Y. P., Cheng, F., Zhou, Y. H., Xia, X. J., Mao, W. H., Shi, K., et al. (2012). Hydrogen peroxide functions as a secondary messenger for brassinosteroids-induced CO_2 assimilation and carbohydrate metabolism in *Cucumis sativus*. *J. Zhejiang Univ. Sci. B* 13, 811–823. doi: 10.1631/jzus.B1200130

Khalili, M., Hasanloo, T., and Safdari, Y. (2014). Hydrogen peroxide acts as a secondary messenger for production of silymarin in Ag^{+} elicited *Silybum marianum* hairy root cultures. *J. Med. Plants By Prod.* 1, 35–40.

Kleczkowski, K., Schell, J., and Bandur, R. (1995). Phytohormone conjugates, nature and function. *Crit. Rev. Plant Sci.* 14, 283–298.

Küpper, H., Lombi, E., Zhao, F. J., and McGrath, S. P. (2000). Cellular compartmentation of cadmium and zinc in relation to other elements in the hyperaccumulator *Arabidopsis halleri*. *Planta* 212, 75–84. doi: 10.1007/s004250000366

Li, R. Q., Yu, C., Li, Y. R., Lam, T. W., Yiu, S. M., Kristiansen, K., et al. (2009). SOAP2, an improved ultrafast tool for short read alignment. *Bioinformatics* 25, 1966–1967. doi: 10.1093/bioinformatics/btp336

Liu, J. P., Zhang, C. C., Wei, C. C., Liu, X., Wang, M. G., Yu, F. F., et al. (2016). The RING finger ubiquitin E3 ligase OsHTAS enhances heat tolerance by promoting H_2O_2-induced stomatal closure in rice. *Plant Physiol.* 170, 429–443. doi: 10.1104/pp.15.00879

Liu, T. M., Zhu, S. Y., Tang, Q. M., and Tang, S. W. (2015). Genome-wide transcriptomic profiling of ramie (*Boehmeria nivea* L. *Gene* 558, 131–137. doi: 10.1016/j.gene.2014.12.057

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Luo, J. S., Yang, Y., Gu, T. Y., Wu, Z. M., and Zhang, Z. H. (2019). The *Arabidopsis* defensin gene AtPDF2.5 mediates cadmium tolerance and accumulation. *Plant Cell Environ.* 42, 2681–2695. doi: 10.1111/pce.13592

Martin, F., Bovet, L., Cordier, A., Stanke, M., Gunduz, I., Peitsch, M. C., et al. (2012). Design of a tobacco exon array with application to investigate the differential cadmium accumulation property in two tobacco varieties. *BMC Genomics* 13:674–690. doi: 10.1186/1471-2164-13-674

Mazel, A., and Levine, L. (2002). Induction of glucosyltransferase transcription and activity during superoxide-dependent cell death in *Arabidopsis* plants. *Plant Physiol. Biochem.* 40, 133–140. doi: 10.1016/s0981-9428(01)01351-1

Mittler, R. (2002). Oxidative stress, antioxidants and stress tolerance. *Trends Plant Sci.* 7, 405–410. doi: 10.1016/s1360-1385(02)02312-9

Moller, I. M. (2001). Plant mitochondria and oxidative stredd: electron transport, NADPH turnover, and metabolism of reactive oxygen species. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 52, 561–591. doi: 10.1146/annurev.arplant.52.1.561

Milner, M. J., Mitani-Ueno, N., Yamaji, N., Yokosho, K., Craft, E., Fei, Z. J., et al. (2014). Root and shoot transcriptome analysis of two ecotypes of *Noccaea caerulescens* uncovers the role of NcNramp1 in Cd hyperaccumlation. *Plant J.* 78, 398–410. doi: 10.1111/tpj.12480

Oono, Y., Yazawa, T., Kawahara, Y., Kanamori, H., Kobayashi, F., Sasaki, H., et al. (2014). Genome-wide transcriptome analysis reveals that cadmium stress signaling controls the expression of genes in drought stress signal pathways in rice. *PLoS One* 9:e96946. doi: 10.1371/journal.pone.0096946

Priest, D. M., Ambrose, S. J., Vaistij, F. E., Elias, L., Higgins, G. S., Ross, A. R. S., et al. (2006). Use of the glucosyltransferase UGT71B6 to disturb abscisic acid homeostasis in *Arabidopsis thaliana*. *Plant J.* 46, 492–502. doi: 10.1111/j.1365-313X.2006.02701.x

Qin, L., Wang, L., Guo, Y., Li, Y., Ümüt, H., and Wang, Y. (2017). An ERF transcription factor from *Tamarix hispida*. *Plant Sci.* 265, 154–166. doi: 10.1016/j.plantsci.2017.10.006

Rajkumar, A. P., Qvist, P., Lazarus, R., Lescai, F., Ju, J., Nyegaard, M., et al. (2015). Experimental validation of methods for differential gene expression analysis and sample pooling in RNA-seq. *BMC Genomics* 16:548–556. doi: 10.1186/s12864-015-1767-y

Rodriguez-Serrano, M., Romero-Puertas, M. C., Pazmino, D. M., Testillano, P. S., Risueno, M. C., Del Río, L. A., et al. (2009). Cellular response of pea plants to cadmium toxicity, cross talk between reactive oxygen species, nitric oxide, and calcium. *Plant Physiol.* 150, 229–243. doi: 10.1104/pp.108.131524

Salt, D. E., and Wagner, G. J. (1993). Cadmium transport across tonoplast of vesicles from oat roots. evidence for a Cd^{2+}/H^{+} antiport activity. *J. Biol. Chem.* 268, 12297–12302. doi: 10.1016/s0021-9258(18)31388-7

Saxena, I., Srikanth, S., and Chen, Z. (2016). Cross talk between H_2O_2 and interacting signal molecules under plant stress response. *Front. Plant Sci.* 7:570. doi: 10.3389/fpls.2016.00570

Sepulveda-Jimenez, G., Rueda-Benıtez, P., Porta, H., and Rocha-Sosa, M. (2004). A red beet (*Beta vulgaris*) UDP-glucosyltransferase gene induced by wounding, bacterial infiltration and oxidative stress. *J. Eep. Bot.* 56, 605–611. doi: 10.1093/jxb/eri036

Sharma, S. S., and Kumar, V. (2002). Responses of wild type and abscisic acid mutants of *Arabidopsis thaliana* to cadmium. *J. Plant Physiol.* 159, 1323–1327. doi: 10.1078/0176-1617-00601

Suwazono, Y., Kido, T., Nakagawa, H., Nishijo, M., and Nogawa, K. (2009). Biological half-life of cadmium in the urine of inhabitants after cessation of cadmium exposure. *Biomarkers* 14, 77–81. doi: 10.1080/13547500902730698

Tognetti, V. B., Van Aken, O., Morreel, K., Vandenbroucke, K., Van de Cotte, B., De Clercq, I., et al. (2010). Perturbation of indole-3-butyric acid homeostasis by the UDP- glucosyltransferase UGT74E2 modulates *Arabidopsis* architecture and water stress tolerance. *Plant Cell* 22, 2660–2679. doi: 10.1105/tpc.109.071316

Uraguchi, S., Mori, S., Kuramata, M., Kawasaki, A., Arao, T., and Ishikawa, S. (2009). Root-to-shoot Cd translocation *via* the xylem is the major process determining shoot and grain cadmium accumulation in rice. *J. Eep. Bot.* 60, 2677–2688. doi: 10.1093/jxb/erp119

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Roy, N. V., Paepe, A. D., et al. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol. Res.* 3, 1–12. doi: 10.1186/gb-2002-3-7-research0034

Vogel-Mikus, K., Pongrac, P., Kump, P., Necemer, M., and Regvar, M. (2005). Colonisation of a Zn, Cd and Pb hyperaccumulator *Thlaspi praecox* wulfen with indigenous arbuscular mycorrhizal fungal mixture induces changes in heavy metal and nutrient uptake. *Environ. Pollut.* 139, 362–371. doi: 10.1016/j.envpol.2005.05.005

Wan, L. C., and Zhang, H. Y. (2012). Cadmium toxicity, effects on cytoskeleton, vesicular trafficking and cell wall construction. *Plant Signal. Behav.* 7, 345–348. doi: 10.4161/psb.18992

Wang, C., Gao, C., Wang, L., Zheng, L., Yang, C., and Wang, Y. (2013). Comprehensive transcriptional profiling of NaHCO3-stressed *Tamarix hispida* roots reveals networks of responsive genes. *Plant Mol. Biol.* 84, 145–157. doi: 10.1007/s11103-013-0124-2

Wang, P. L., Lei, X., Lü, J. X., and Gao, C. Q. (2020). Overexpression of the ThTPS gene enhanced salt and osmotic stress tolerance in *Tamarix hispida*. *J. For. Res.* 1–10.

Wei, S. H., Zhou, Q. X., and Wang, X. (2005). Cadmium Hyperaccumulator *Solanum nigrum* L. and its accumulating characteristics. *Environ. sci.* 26, 167–171.

Wu, F. B., Zhang, G. P., Dominy, P., Wu, H. X., Bachir, and Dango, M. L. (2007). Differences in yield components and kernel Cd accumulation in response to Cd toxicity in four barley genotypes. *Chemosphere* 70, 83–92. doi: 10.1016/j.chemosphere.2007.06.051

Xu, S. S., Lin, S. Z., and Lai, Z. X. (2015). Cadmium impairs iron homeostasis in Arabidopsis thaliana by increasing the polysaccharide contents and the iron-binding capacity of root cell walls. *Plant Soil* 392, 71–85. doi: 10.1007/s11104-015-2443-3

Xu, Z. J., Nakajima, M., Suzuki, Y., and Yamaguchi, I. (2002). Cloning and characterization of the abscisic acid-specific glucosyltransferase gene from Adzuki bean seedings. *J. Plant Physiol.* 129, 1285–1295. doi: 10.1104/pp.001784

Yang, G. Y., Wang, C., Wang, Y. C., Guo, Y. C., Zhao, Y. L., Yang, C. P., et al. (2016). Overexpression of ThVHAc1 and its potential upstream regulator, ThWRKY7, improved plant tolerance of cadmium stress. *Sci. Rep.* 6:18752. doi: 10.1038/srep18752

Yang, J. L., Wang, Y. C., Liu, G. F., Yang, C. P., and Li, C. H. (2011). Tamarix hispida metallothionein-like ThMT3, a reactive oxygen species scavenger, increases tolerance against Cd2+, Zn2+, Cu2+, and NaCl in transgenic yeast. *Mol Biol Rep.* 38, 1567–1574. doi: 10.1007/s11033-010-0265-1

Yang, X. E., Long, X. X., Ye, H. B., He, Z. L., Calver, D. V., and Stoffl la, P. J. (2004). Cadmium tolerance and hyperaccumulation in a new Zn-hyperaccumulating plant species (*Sedum alfredii Hance*). *Plant Soil.* 259, 181–189. doi: 10.1023/b:plso.0000020956.24027.f2

Ye, J., Fang, L., Zheng, H. K., Zhang, Y., Chen, J., Zhang, Z. J., et al. (2006). WEGO, a web tool for plotting go annotations. *Nucleic Acids Res.* 34, 293–297.

Yue, R. Q., Lu, C. X., Qi, J. S., Han, X. H., Yan, S. F., Guo, S. L., et al. (2016). Transcriptome analysis of cadmium-treated roots in maize (*Zea mays* L.). *Front. Plant Sci.* 7:1298. doi: 10.3389/fpls.2016.01298

Zang, D., Wang, C., Ji, X., and Wang, Y. (2015). *Tamarix hispida* zinc finger protein ThZFP1 participates in salt and osmotic stress tolerance by increasing proline content and SOD and POD activities. *Plant Sci.* 235, 111–121. doi: 10.1016/j.plantsci.2015.02.016

Zhang, Z. H., Zhou, T., Tang, T. J., Song, H. Q., Guan, C. Y., Huang, J. S., et al. (2019b). Multiomics landscapes uncover the pivotal role of subcellular reallocation of cadmium in regulating rapeseed resistance to cadmium toxicity. *J. Exp. Bot.* 2019:erz295.

Zhang, P., Wang, R. L., Ju, Q., Li, W. Q., Tran, L. P., and Jin, X. (2019a). The R2R3-MYB transcription factor MYB49 regulates cadmium accumulation. *Plant Physiol.* 180, 529–542. doi: 10.1104/pp.18.01380

Zhang, X., Wang, L., Meng, H., Wen, H., Fan, Y., and Zhao, J. (2011). Maize ABP9 enhances tolerance to multiple stresses in transgenic *Arabidopsis* by modulating ABA signaling and cellular levels of reactive oxygen species. *Plant Mol. Biol.* 75, 365–378. doi: 10.1007/s11103-011-9732-x

Zhou, B. R., Yao, W. J., Wang, S. J., Wang, X. W., and Jiang, T. B. (2014). The metallothionein gene, tamt3, from tamarix androssowii confers cd2+ tolerance in tobacco. *Int. J. Mol. Sci.* 15, 10398–10409. doi: 10.3390/ijms150610398

Zhou, Q., Guo, J. J., He, C. S., Shen, C., Huang, Y. Y., Chen, J. X., et al. (2016). Comparative transcriptome analysis between low- and high- cadmium -accumulating genotypes of pakchoi (*Brassica chinensis* L.) in response to cadmium stress. *Environ. Sci. Technol.* 50, 6485–6494. doi: 10.1021/acs.est.5b06326

Zuccarelli, R., Coelho, A. C. P., Peres, L. E. P., and Freschi, L. (2017). Shedding light on no homeostasis, light as a key regulator of glutathione and nitric oxide metabolisms during seedling deetiolation. *Nitric Oxide* 68, 77–90. doi: 10.1016/j.niox.2017.01.006

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 25 April 2022
doi: 10.3389/fpls.2022.873071

Comprehensive Evolutionary Analysis of *CPP* Genes in *Brassica napus* L. and Its Two Diploid Progenitors Revealing the Potential Molecular Basis of Allopolyploid Adaptive Advantage Under Salt Stress

*Mengdi Li[1,2], Fan Wang[2], Jiayu Ma[1], Hengzhao Liu[1], Hang Ye[1], Peng Zhao[1] and Jianbo Wang[2]**

[1] Key Laboratory of Resource Biology and Biotechnology in Western China, Ministry of Education, College of Life Sciences, Northwest University, Xi'an, China, [2] State Key Laboratory of Hybrid Rice, College of Life Sciences, Wuhan University, Wuhan, China

OPEN ACCESS

Edited by:
Rajeev K. Varshney, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), India

Reviewed by:
Jinpeng Wang, Institute of Botany (CAS), China
Peijian Cao, Zhengzhou Tobacco Research Institute of CNTC, China

***Correspondence:**
Jianbo Wang
jbwang@whu.edu.cn

Specialty section:
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

Received: 10 February 2022
Accepted: 29 March 2022
Published: 25 April 2022

Citation:
Li M, Wang F, Ma J, Liu H, Ye H, Zhao P and Wang J (2022) Comprehensive Evolutionary Analysis of CPP Genes in Brassica napus L. and Its Two Diploid Progenitors Revealing the Potential Molecular Basis of Allopolyploid Adaptive Advantage Under Salt Stress. Front. Plant Sci. 13:873071. doi: 10.3389/fpls.2022.873071

Allopolyploids exist widely in nature and have strong environmental adaptability. The typical allopolyploid *Brassica napus* L. is a widely cultivated crop, but whether it is superior to its diploid progenitors in abiotic stress resistance and the key genes that may be involved are not fully understood. *Cystein-rich polycomb-like protein* (*CPP*) genes encode critical transcription factors involved in the response of abiotic stress, including salt stress. To explore the potential molecular basis of allopolyploid adaptation to salt stress, we comprehensively analyzed the characteristics and salt stress response of the *CPP* genes in *B. napus* and its two diploid progenitors in this study. We found some molecular basis that might be associated with the adaptability of *B. napus*, including the expansion of the CPP gene family, the acquisition of introns by some *BnCPPs*, and abundant *cis*-acting elements upstream of *BnCPPs*. We found two duplication modes (whole genome duplication and transposed duplication) might be the main reasons for the expansion of CPP gene family in *B. napus* during allopolyploidization. *CPP* gene expression levels and several physiological indexes were changed in *B. napus* and its diploid progenitors after salt stress, suggesting that *CPP* genes might play important roles in the response of salt stress. We found that some *BnCPPs* might undergo new functionalization or subfunctionalization, and some *BnCPPs* also show biased expression, which might contribute to the adaptation of *B. napus* under saline environment. Compared with diploid progenitors, *B. napus* showed stronger

physiological responses, and *BnCPP* gene expression also showed higher changes after salt stress, indicating that the allopolyploid *B. napus* had an adaptive advantage under salt stress. This study could provide evidence for the adaptability of polyploid and provide important clues for the study of the molecular mechanism of salt stress resistance in *B. napus*.

Keywords: adaptive advantage, *Brassica napus*, *CPP* gene family, gene expression, physiological response, salt stress, allopolyploidization

INTRODUCTION

Allopolyploidization is a powerful driving force for plant speciation and species diversity (Abbott et al., 2013; Soltis et al., 2016). This process makes the plants undergo a series of genetic and epigenetic changes, which may lead to the formation of new phenotypes, thus promoting the successful establishment of new allopolyploids and enabling it to adapt to the diverse ecological environment (Adams and Wendel, 2005; Chen, 2007; Jackson and Chen, 2010; Madlung, 2013; Yoo and Wendel, 2014). Natural allotetraploid *Brassica napus* (2n = 4x = 38, $A_nA_nC_nC_n$) is one of the most widely cultivated oil crops, which was formed by hybridization and genome doubling of diploid *B. rapa* (2n = 20, A_r) and *B. oleracea* (2n = 18, C_o) about 7,500 years ago, and its domestication and cultivation history were relatively short (Chalhoub et al., 2014). Therefore, natural *B. napus* and its diploid progenitors, *B. rapa* and *B. oleracea*, are a group of ideal material for investigating allopolyploidization event (Prakash et al., 2009). As an allotetraploid, *B. napus* has advantages than the diploid *Brassica* species in many aspects such as environmental adaptability (Segraves, 2017). Studies have found that rapeseed can accumulate erucic acid through erucic acid metabolic pathways under salt stress, and high erucic acid genes are related to the stress resistant genes (Ashraf and McNeilly, 2004; Wang et al., 2020), which implied that rapeseed might have the ability to resistant the salt stress. However, the potential molecular basis of the adaptive advantages in allopolyploids was still largely unknown.

Cystein-rich polycomb-like protein (CPP) contains one or two cysteine-rich (CXC) domains and a short sequence R that connects two CXC domains (Riechmann et al., 2000; Andersen et al., 2007; Zhang et al., 2015). Many studies reported *CPP* genes showed response to a variety of abiotic stresses in plants. For example, the expression of 15 *CPP* genes was significantly increased after 24 h of drought treatment in soybean (Zhang et al., 2015). *CsCPP* genes engaged in rapid stress responses under abiotic stresses (such as low temperature, drought, salt and abscisic acid) in cucumber (Zhou et al., 2018). The expression of *ZmCPP* genes in maize changed under extreme temperature, salt and drought stresses (Song et al., 2016). Studies have also reported the specific function of *CPP* genes. *TSO1*, the first *CPP* gene identified in *Arabidopsis thaliana*, highly expressed in flowers, developing ovules and microspores of Arabidopsis, and the lack of *TSO1* affected the karyokinesis and cytokinesis of cells (Hauser et al., 1998, 2000; Song et al., 2000). Further studies by Wang et al. (2018) found that *TSO1* played a key role in regulating cell cycle and cell fate, and was involved in regulating the development of plant stems and roots. In soybean, transcription factor CPP1 was involved in the regulation of the symbiotic nodules development by regulating the expression of *GMLBC3* (Cvitanich et al., 2000). In addition, Yang et al. (2008) first identified 8 *CPP* genes in Arabidopsis and 11 *CPP* genes in rice, and they found that the CXC domain and the R sequence might have coevolved during evolution. With the release and improvement of genomic information of many species, more CPP gene families in different species have been studied. For example, 20, 13, and 5 *CPP* genes were identified in soybean, maize, cucumber and wheat, respectively (Zhang et al., 2015; Song et al., 2016; Zhou et al., 2018). In general, the CPP gene family has only been systematically studied in a few species, and most of them are diploid species. Considering that the comparative analysis of CPP gene family between allopolyploid and its diploid parents/progenitors can provide new insights into the molecular basis of environmental adaptation in allopolyploids, it is necessary to identify and analyze CPP gene families in allopolyploid plants.

In this study, we comprehensively analyzed the characteristics and salt stress response (including physiological and gene expression changes) of the CPP gene family in *B. napus* and its two diploid progenitors, *B. rapa* and *B. oleracea*. We also explored the evolutionary changes of the CPP gene family in the allotetraploid *B. napus* compared with its progenitors, and uncover the potential molecular basis of the advantage of salt stress resistance in allotetraploid *B. napus*. Our results provide important clues for researchers to explore the roles of *CPP* genes in salt stress tolerance, and will deepen our understanding of the evolution of the molecular basis for polyploid stress resistance.

MATERIALS AND METHODS

Identification and Duplication Mode of *CPP* Genes

All genome data of *B. rapa*, *B. oleracea* and *B. napus* used in this study were obtained from *Brassica* BRAD database (Cheng et al., 2011). Three methods were used to obtain candidate CPP gene family members. First, 8 AtCPP proteins were used as query sequences (downloaded from TAIR database[1]) to perform Blastp alignment (*E*-value < $1e^{-5}$) with all protein sequences of three *Brassica* species. Second, the HMM file of the CXC domain (ID: PF03638) was downloaded from the Pfam database[2], and then

[1]https://www.Arabidopsis.org/

[2]http://pfam.xfam.org/

the software HMMER was used to search this domain in all protein sequences of three species (*E*-value $< 1e^{-5}$). Third, the syntenic genes of 8 *AtCPP* were searched in three *Brassica* species using the syntenic gene search function of BRAD database. All the members obtained by these three methods (i.e., the union of the members) were candidate members of the CPP gene family. Furthermore, Smart database[3], Pfam database and CDD database[4] were used to detect the integrity of the CXC domain of all the candidate CPP proteins, and the candidate members with sequence deletion in the domain were removed to obtain the final CPP gene family members. The duplication modes of all *CPP* genes were analyzed by the online website PlantDGD[5]. CDS sequences of three *Brassica* species were downloaded from BRAD database, and the CDS sequences of *CPP* genes of each species were compared using Blastn program. If the consistency and coverage of two CDS were both greater than 80%, they were considered as segmental duplicated gene pairs (Zhou et al., 2004). The software DNAsp5 (Librado and Rozas, 2009) was used to estimate the non-synonymous (*K*a) and synonymous substitution (*K*s) values between these gene pairs, and then the selection pressure of *CPP* genes in *B. napus* and its progenitors was determined based on Ka/Ks value.

Phylogenetic Tree Construction and Characterization Analysis of CPP Proteins

CPP proteins from five species (*B. rapa*, *B. oleracea*, *B. napus*, Arabidopsis and rice) were aligned using the Clustal X program in MEGA software (v7.0; Kumar et al., 2016). According to the alignment results, the MEGA software was used to perform phylogenetic analysis of these CPP proteins utilizing the neighbor-joining (NJ) method with 1000 bootstrap replicates. The phylogenetic tree was decorated by the online website Tree of Life (iTOL[6]). According to the information provided by Smart database, IBS software (v1.0.3) was used to show the distribution of CXC domain on CPP proteins. The software Clustal W was used to align the CXC domains of all CPP proteins from the three *Brassica* species. The online website MEME[7] was used to predict the characteristic motifs of CPP proteins of three species, and the number and length of the predicted motifs were set as 10 and 6–50 aa, respectively. The structure of the *CPP* genes was analyzed using the online tool GSDS[8].

Chromosomal Localization and Synteny Analysis of *CPP* Genes

The location information of all identified *CPP* genes was obtained from BRAD database, and their positions on chromosomes were visualized using the MapInspector software[9,10]. The syntenic *CPP* genes were searched by the syntenic gene search function of BRAD database, and the synteny relationship between *CPP* genes was showed by Circos software (Krzywinski et al., 2009).

Physicochemical Properties and Subcellular Localization Prediction of CPP Proteins

Online website ExPASy[11] was used to predicate the physicochemical properties of CPP proteins, including the number of amino acids, molecular weight (MW), isoelectric point (pI), the grand average of hydropathy (GRAVY) and instability index (II). Subcellular localization of CPP proteins in three *Brassica* species was predicted using the online tool WoLF PSORT[12].

The *Cis*-Acting Elements in the Promoter of *CPP* Genes and Their Expression Patterns

The 2000 bp upstream of the transcription start site of *CPP* genes were considered to be the promoter region of *CPP* genes. The *cis*-acting elements in the promoter region of *CPP* genes were predicted using website PlantCARE[13]. Gene expression data from four organs in three *Brassica* species were derived from our previous studies (Li et al., 2020), and the raw data of which were deposited in the NCBI database (accession number: SRR7816633-SRR7816668). The heatmap was created using the software HemI.

Plant Materials and Treatments

The experimental materials used in this study were *B. napus* (cv. Darmor), *B. rapa* (cv. Chiifu) and *B. oleracea* (cv. Jinzaosheng). Seeds were placed on wet double filter paper in a culture dish for germination in the light incubator (temperature: 22°C, light/dark: 16 h/8 h). When cotyledons of the seedlings were fully unfolded, placed them in black pots in which an equal amount of vermiculite and vegetative soil were mixed. The seedlings then grown in a light incubator under the same conditions, during which 1/2 Hoagland's nutrient solution were regularly applied. Seedlings growing for 4 weeks were subjected to salt stress. The materials in the experimental group were irrigated with NaCl solution with a concentration of 300 mM (once a day, 30 ml), and the materials in the control group were irrigated with the same amount of distilled water for 4 days. Take the second to fifth leaves of the seedlings and freeze them immediately in liquid nitrogen for further use. Each species/treatment has three biological replicates. The experimental materials used in this study were shown in the **Supplementary Figure 1**.

Measurement of Physiological Indexes

Three enzyme activities, including peroxidase (POD), superoxide dismutase (SOD) and catalase (CAT), were quantified according to the method of Cao et al. (2017), with minor modifications as described by Xu et al. (2020). The contents of proline (Pro),

[3]http://smart.embl-heidelberg.de/

[4]http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi/

[5]http://pdgd.njau.edu.cn:8080/

[6]http://itol.embl.de/

[7]http://meme-suite.org/

[8]http://gsds.cbi.pku.edu.cn/

[9]https://gaow.github.io/genetic-analysis-software/m/mapinspect/

[10]https://mapinspect.software.informer.com/download/

[11]https://web.expasy.org/tools/

[12]https://wolfpsort.hgc.jp/

[13]http://bioinformatics.psb.ugent.be/webtools/plantcare/html/

soluble sugar, soluble protein, and chlorophyll (Chl a + Chl b) were measured according to Quinet et al. (2012).

RNA Isolation and Quantitative Real-Time PCR (qRT-PCR) Analysis

Total RNAs from each sample was isolated using the Trizol reagent (Invitrogen, Carlsbad, CA, United States) according to the instructions. The quality of RNA samples was checked on 1% agarose gel. The first strand complementary DNA (cDNA) was synthesized using Moloney Murine Leukemia Virus Reverse Transcriptase (M-MLV RT, Promega, Madison, WI, United States) following the manufacturer's instructions. Then, cDNA was diluted 10-fold as the templates of qRT-PCR. The qRT-PCR reactions were performed using the ABI Step One Plus Real-Time PCR System (Applied Biosystems, Carlsbad, CA, United States) with the SYBR kit. The program was set to 94°C for 5 min; 38 cycles of 94°C for 30 s, 55°C for 30 s and 72°C for 20 s; and 72°C for 10 min. Three biological and three technical repeats were used for each sample. The *ACT2/7* gene was used as internal control to standardize the results and the relative expression level of selected genes was calculated using the delta-delta threshold cycle ($2^{-\Delta\Delta Ct}$) method. Gene-specific primers for qRT-PCR were designed by Primer 5 software and the primer sequences were listed in **Supplementary Table 1**.

RESULTS

CPP Gene Family Expanded in Allotetraploid *Brassica napus*

To ensure that CPP gene families in *B. napus* and its two diploid progenitors are fully identified, we used three methods to obtain candidate members. The three methods include Blastp alignment, HMMER scanning, and syntenic gene searching (see section "Identification and duplication mode of CPP genes" in Materials and Methods section). A total of 18, 19, and 37 candidate *CPP* genes were initially obtained in *B. rapa*, *B. oleracea* and *B. napus*. The integrity of the CXC domain of all candidate proteins was examined, and which with complete domain were identified as the final CPP gene family members. Finally, 15, 10, and 34 members were identified in *B. rapa*, *B. oleracea,* and *B. napus* (**Table 1**). The identified members were named according to the previous studies (Li et al., 2019b; Wang et al., 2019; **Table 1**).

As we all know, *B. napus* is derived from natural hybridization and chromosome doubling of its two diploid progenitors, *B. rapa* and *B. oleracea*. Therefore, theoretically, the number of *CPP* genes in *B. napus* should be equal to the sum of the number of *CPP* genes in *B. rapa* and *B. oleracea*. In this study, 15 and 10 *CPP* genes were identified in *B. rapa* and *B. oleracea*, respectively, and the sum of *CPP* genes in the two species was much smaller than the number of *CPP* genes in *B. napus* (34). Specifically, the number of each *BnCPP* genes was greater than or equal to the sum of the corresponding homoeologous genes in the two diploid ancestral species (with one exception, *BnCPP8*; **Supplementary Table 2**). Compared with diploid species, the number of *CPP1-3* genes in *B. napus* did not change and the number of *CPP4-6* genes in both subgenomes of *B. napus* increased during allopolyploidization process. Statistically, the number of lost genes (2) was less than the number of increased genes (11) in *B. napus* (**Supplementary Table 2**). Therefore, the CPP gene family in *B. napus* expanded during the allopolyploidization process, although a small number of gene loss events occurred.

The Duplication Mode and Synteny Analysis of *CPP* Genes

To explore the reasons for the expansion of CPP gene family in *B. napus*, we analyzed the duplication modes of *CPP* genes (**Table 2** and **Supplementary Figure 2**). Gene duplication modes include whole genome duplication (WGD) and various single gene duplication, and the latter consists of four models, including tandem duplication (TD), proximal duplication (PD), transposed duplication (TRD) and dispersed duplication (DSD; Qiao et al., 2019). We found that *CPP* genes in *B. rapa*, *B. oleracea* and *B. napus* showed a large number of genome wide duplication, accounting for 60, 50, and 38.2%, respectively, which might be related to the WGD and WGT events of *Brassica* progenitor. For the four single gene duplication modes, the mechanism of the DSD is unclear (Qiao et al., 2019), so it was excluded from the following comparative analysis. We found that TRD was the most frequent mode (about 50%) in all three species, suggesting that this mode was important for the expansion of the CPP gene family. In addition, although WGD and TRD were the most frequent duplication modes of *CPP* genes in the three species, the frequency of TRD was higher than that of WGD only in the *CPP* genes of *B. napus*. Therefore, we speculated that TRD might greatly promote the expansion of CPP gene family in *B. napus* during allopolyploidization. Moreover, PD was observed in only one *CPP* gene in *B. rapa*, indicating that this duplication mode is not the main source of CPP gene family expansion in *B. rapa*, *B. oleracea* and *B. napus*. In addition, a total of 29 segmental duplicated gene pairs were detected in three *Brassica* species, and their *K*a/*K*s values were calculated to explore the selection pressure on *CPP* genes during evolution (**Supplementary Table 3**). The results showed that all the values were less than 1, indicating that this family was subjected to purification selection during evolution. The mean value of all *K*a/*K*s values was 0.42, which might indicate that the purification selection on *CPP* genes was not strong. One study suggested that *CPP* genes in plants might be subjected to purification selection only in the conserved domain (Yang et al., 2008). We also visualized these segmental duplication pairs and five duplication modes in **Supplementary Figure 3**.

To better investigate the evolution of *CPP* genes among three *Brassica* species, we visualized the syntenic relationship of identified *CPP* genes. We found that *CPP* genes in these three *Brassica* species have a good syntenic relationship, and they were closely related in evolution (**Figure 1**). There were 8 *AtCPP* genes, 13 *BrCPP* genes, 7 *BoCPP* genes and 13 *BnCPP* genes, respectively, and these syntenic genes were mainly

TABLE 1 | The *CPP* gene family information in *B. napus* and its diploid progenitors.

Gene name	Gene ID	Chromosome	Gene position		Intron number	Orthologous gene in *Arabidopsis*
			Start	End		
BrCPP1	Bra041042	Scaffold000410	7858	9837	6	AT2G20110/*AtCPP1*
BrCPP2	Bra001128	A03	14899606	14901982	7	AT3G04850/*AtCPP2*
BrCPP3	Bra021165	A01	22801607	22803170	5	AT3G16160/*AtCPP3*
BrCPP4a	Bra023810	A01	19285539	19288393	8	AT3G22760/*AtCPP4*
BrCPP4b	Bra033920	A05	16207472	16209714	6	AT3G22760/*AtCPP4*
BrCPP5a	Bra033919	A05	16191831	16195460	10	At3G22780/*AtCPP5*
BrCPP5b	Bra001884	A03	19146128	19152768	16	At3G22780/*AtCPP5*
BrCPP6	Bra036900	A01	12049577	12052924	8	AT4G14770/*AtCPP6*
BrCPP7a	Bra010348	A08	13239578	13242200	7	AT4G29000/*AtCPP7*
BrCPP7b	Bra011075	A01	4078631	4081317	6	AT4G29000/*AtCPP7*
BrCPP7c	Bra024171	A03	27178351	27180638	6	AT4G29000/*AtCPP7*
BrCPP7d	Bra025590	A04	8011846	8012463	1	AT4G29000/*AtCPP7*
BrCPP8a	Bra009859	A06	18346154	18348620	7	AT5G25790/*AtCPP8*
BrCPP8b	Bra036551	A09	2851635	2853619	4	AT5G25790/*AtCPP8*
BrCPP8c	Bra020513	A02	24310423	24316725	18	AT5G25790/*AtCPP8*
BoCPP1	Bol020837	C02	22859403	22861436	6	AT2G20110/*AtCPP1*
BoCPP2	Bol034129	Scaffold000040	1607390	1608218	3	AT3G04850/*AtCPP2*
BoCPP3	Bol034761	C01	32915941	32917598	5	AT3G16160/*AtCPP3*
BoCPP4	Bol023280	C01	27868376	27869528	3	AT3G22760/*AtCPP4*
BoCPP6	Bol008191	C01	38274359	38274904	2	AT4G14770/*AtCPP6*
BoCPP7a	Bol019616	C01	5441257	5443907	7	AT4G29000/*AtCPP7*
BoCPP7b	Bol021116	C08	16093123	16095690	6	AT4G29000/*AtCPP7*
BoCPP8a	Bol022330	C07	35441538	35444006	7	AT5G25790/*AtCPP8*
BoCPP8b	Bol032486	C09	3715173	3717272	5	AT5G25790/*AtCPP8*
BoCPP8c	Bol016381	C02	42997976	43000106	5	AT5G25790/*AtCPP8*
BnA.CPP1b	BnaA07g00040D	chrA07	21055	23400	6	AT2G20110/*AtCPP1*
BnA.CPP2a	BnaA03g28830D	chrA03	14033704	14036289	7	AT3G04850/*AtCPP2*
BnA.CPP3b	BnaA01g28180D	chrA01	19681341	19684544	6	AT3G16160/*AtCPP3*
BnA.CPP4c	BnaA05g16940D	chrA05	11718332	11720970	8	AT3G22760/*AtCPP4*
BnA.CPP5b	BnaA05g16930D	chrA05	11709193	11712979	10	At3G22780/*AtCPP5*
BnA.CPP5d	BnaA03g36790D	chrA03	18065426	18075292	16	At3G22780/*AtCPP5*
BnA.CPP5e	BnaA01g24520D	chrA01	16916341	16917772	5	At3G22780/*AtCPP5*
BnA.CPP6a	BnaA01g18960D	chrA01	10416162	10419574	8	AT4G14770/*AtCPP6*
BnA.CPP6e	BnaA03g15030D	chrA03	6929076	6937274	31	AT4G14770/*AtCPP6*
BnA.CPP7a	BnaA01g07920D	chrA01	3741875	3745039	7	AT4G29000/*AtCPP7*
BnA.CPP7d	BnaA08g30730D	chrA08_random	1678620	1681241	7	AT4G29000/*AtCPP7*
BnA.CPP7f	BnaAnng24750D	chrAnn_random	28633500	28634835	3	AT4G29000/*AtCPP7*
BnA.CPP7g	BnaAnng37830D	chrAnn_random	42872263	42873422	3	AT4G29000/*AtCPP7*
BnA.CPP7h	BnaA04g10170D	chrA04	8987674	8988291	1	AT4G29000/*AtCPP7*
BnA.CPP8b	BnaA06g27870D	chrA06	19167716	19170265	7	AT5G25790/*AtCPP8*
BnA.CPP8d	BnaA09g04480D	chrA09	2203990	2205985	4	AT5G25790/*AtCPP8*
BnC.CPP1a	BnaC07g00810D	chrC07	1116628	1119040	6	AT2G20110/*AtCPP1*
BnC.CPP2b	BnaC03g33970D	chrC03	20676845	20679238	6	AT3G04850/*AtCPP2*
BnC.CPP3a	BnaC01g35470D	chrC01	34864388	34866026	5	AT3G16160/*AtCPP3*
BnC.CPP4a	BnaCnng56700D	chrCnn_random	56576169	56579283	8	AT3G22760/*AtCPP4*
BnC.CPP4b	BnaCnng43510D	chrCnn_random	42545612	42548303	8	AT3G22760/*AtCPP4*
BnC.CPP4d	BnaCnng27640D	chrCnn_random	26215960	26218588	8	AT3G22760/*AtCPP4*
BnC.CPP5a	BnaCnng27630D	chrCnn_random	26211360	26214466	9	At3G22780/*AtCPP5*
BnC.CPP5c	BnaC03g42960D	chrC03	27663590	27671331	17	At3G22780/*AtCPP5*
BnC.CPP5f	BnaCnng43520D	chrCnn_random	42564973	42566124	4	At3G22780/*AtCPP5*
BnC.CPP6b	BnaC01g22810D	chrC01	16334181	16337785	8	AT4G14770/*AtCPP6*
BnC.CPP6c	BnaC03g18250D	chrC03	9349757	9353091	10	AT4G14770/*AtCPP6*
BnC.CPP6d	BnaC03g18050D	chrC03	9243783	9253263	31	AT4G14770/*AtCPP6*
BnC.CPP7b	BnaC01g09520D	chrC01	5562615	5570008	14	AT4G29000/*AtCPP7*
BnC.CPP7c	BnaC08g13260D	chrC08	18279478	18282187	7	AT4G29000/*AtCPP7*
BnC.CPP7e	BnaC07g41650D	chrC07	41430430	41432917	8	AT4G29000/*AtCPP7*
BnC.CPP8a	BnaC07g29080D	chrC07	33999748	34002214	7	AT5G25790/*AtCPP8*
BnC.CPP8c	BnaCnng48380D	chrCnn_random	47677091	47679547	5	AT5G25790/*AtCPP8*
BnC.CPP8e	BnaC02g40480D	chrC02	43513384	43515513	6	AT5G25790/*AtCPP8*

TABLE 2 | Five modes of *CPP* gene duplication in *B. napus* and its diploid progenitors.

Gene name	Gene ID	WGD	Single gene duplication			
			TRD	TD	PD	DSD
BrCPP1	Bra041042		√			√
BrCPP2	Bra001128		√			√
BrCPP3	Bra021165					√
BrCPP4a	Bra023810				√	√
BrCPP4b	Bra033920			√		√
BrCPP5a	Bra033919	√	√	√		
BrCPP5b	Bra001884	√	√			√
BrCPP6	Bra036900	√				√
BrCPP7a	Bra010348	√	√			√
BrCPP7b	Bra011075	√	√			√
BrCPP7c	Bra024171	√				√
BrCPP7d	Bra025590		√			√
BrCPP8a	Bra009859	√				√
BrCPP8b	Bra036551	√				√
BrCPP8c	Bra020513	√	√			√
BoCPP1	Bol020837		√			√
BoCPP2	Bol034129					
BoCPP3	Bol034761					√
BoCPP4	Bol023280	√	√			
BoCPP6	Bol008191		√			√
BoCPP7a	Bol019616	√	√			√
BoCPP7b	Bol021116	√				√
BoCPP8a	Bol022330	√				√
BoCPP8b	Bol032486	√				√
BoCPP8c	Bol016381	√	√			√
BnC.CPP1a	BnaC07g00810D		√			√
BnA.CPP1b	BnaA07g00040D		√			√
BnA.CPP2a	BnaA03g28830D					√
BnC.CPP2b	BnaC03g33970D					√
BnC.CPP3a	BnaC01g35470D		√			√
BnA.CPP3b	BnaA01g28180D					√
BnC.CPP4a	BnaCnng56700D		√			√
BnC.CPP4b	BnaCnng43510D			√		√
BnA.CPP4c	BnaA05g16940D			√		√
BnC.CPP4d	BnaCnng27640D		√			√
BnC.CPP5a	BnaCnng27630D		√			√
BnA.CPP5b	BnaA05g16930D	√		√		
BnC.CPP5c	BnaC03g42960D		√			√
BnA.CPP5d	BnaA03g36790D	√				√
BnA.CPP5e	BnaA01g24520D	√				√
BnC.CPP5f	BnaCnng43520D			√		√
BnA.CPP6a	BnaA01g18960D	√				√
BnC.CPP6b	BnaC01g22810D					√
BnC.CPP6c	BnaC03g18250D		√			√
BnC.CPP6d	BnaC03g18050D	√	√			√
BnA.CPP6e	BnaA03g15030D	√				√
BnA.CPP7a	BnaA01g07920D		√			√
BnC.CPP7b	BnaC01g09520D	√	√			√
BnC.CPP7c	BnaC08g13260D	√				√
BnA.CPP7d	BnaA08g30730D		√			√
BnC.CPP7e	BnaC07g41650D	√				√
BnA.CPP7f	BnaAnng24750D		√			√
BnA.CPP7g	BnaAnng37830D		√			√
BnA.CPP7h	BnaA04g10170D		√			√
BnC.CPP8a	BnaC07g29080D	√	√			√
BnA.CPP8b	BnaA06g27870D	√				√
BnC.CPP8c	BnaCnng48380D		√			√
BnA.CPP8d	BnaA09g04480D	√				√
BnC.CPP8e	BnaC02g40480D	√	√			√

located in five conserved genome blocks F, H, Q, T, and U (**Supplementary Table 4**). In the three *Brassica* species, about 55.9% (33/59) of *CPP* genes retained some syntenic relationships between species. In addition, we found that the increase or loss of syntenic genes occurred in *B. napus* after allopolyploidization. For example, the number of *AtCPP2* syntenic genes in *B. rapa* and *B. oleracea* was 1 and 0 respectively, but the number of *AtCPP2* syntenic genes in *B. napus* increased to two. *AtCPP1* had one syntenic gene in *B. rapa* and *B. oleracea* respectively, but it was not found in *B. napus*. *AtCPP3* had one syntenic gene in both *B. rapa* and *B. oleracea*, which was all located in the LF subgenome (note: LF subgenome stands for less fractioned subgenome; Cheng et al., 2011, 2012), and these two genes were retained in *B. napus* and were also located in the two LF subgenomes. Therefore, we speculated that *AtCPP3* syntenic genes might play an irreplaceable role in the evolution and adaptation of these three *Brassica* species.

Evolutionary Changes in Chromosomal Localization of *CPP* Genes

In addition to focusing on evolutionary changes in the number of *CPP* genes in *B. napus*, we also focused on the changes in the relative position of *CPP* genes on chromosomes. The number of *CPP* genes in A_n subgenome (16) was similar to that in the A_r genome (15), but the number of *CPP* genes in the C_n subgenome (18) was more than that in the C_o genome (10). Moreover, we noticed that most of *CPP* genes in the C_n subgenome (12 out of 18) experienced the TRD duplication event (**Table 2**), indicating that the number of *CPP* genes in C_n subgenome might expend after allopolyploidization, which might be related to the fact that more transposable elements were distributed in the C_n subgenome (Chalhoub et al., 2014). In addition, another reason for this phenomenon might be the species-specific loss of *CPP* genes in C_o genome during evolution. We further analyzed the chromosomal location of *CPP* genes, and a total of 25 *BnCPP* genes, 14 *BrCPP* genes and 9 *BoCPP* genes have been accurately mapped to the specific location of the chromosome (**Figure 2**). The remaining *CPP* genes have not been located in the chromosome due to the relatively low degree assembly of the reference genome. *CPP* genes were not found on 7 chromosomes (A_n02, A_n08, A_n10, C_n04-06, C_n09), 2 chromosomes (A_r07, A_r10) and 4 chromosomes (C_o03-06) in *B. napus*, *B. rapa* and *B. oleracea*, respectively. By comparing the distribution of *CPP* genes on A_n and A_r subgenome/genome, we found a similar pattern, that is, 4, 3, and 2 *CPP* genes were distributed on chromosome 1, 3 and 5, respectively, while the remaining *CPP* genes were distributed on other chromosomes as a single form. However, the distribution of *CPP* genes on C_n and C_o subgenome/genome was significantly different. In addition, about 71.4% (10/14) of *CPP* genes in A_r genome of *B. rapa* did not change their relative positions in chromosomes after allopolyploidization, while up to 44.4% (4/9) of *CPP* genes in C_o genome of *B. oleracea* maintained their original relative positions in chromosomes after allopolyploidization. This phenomenon was consistent with previous studies about the other gene families in *B. napus* (Li et al., 2019a,b), which

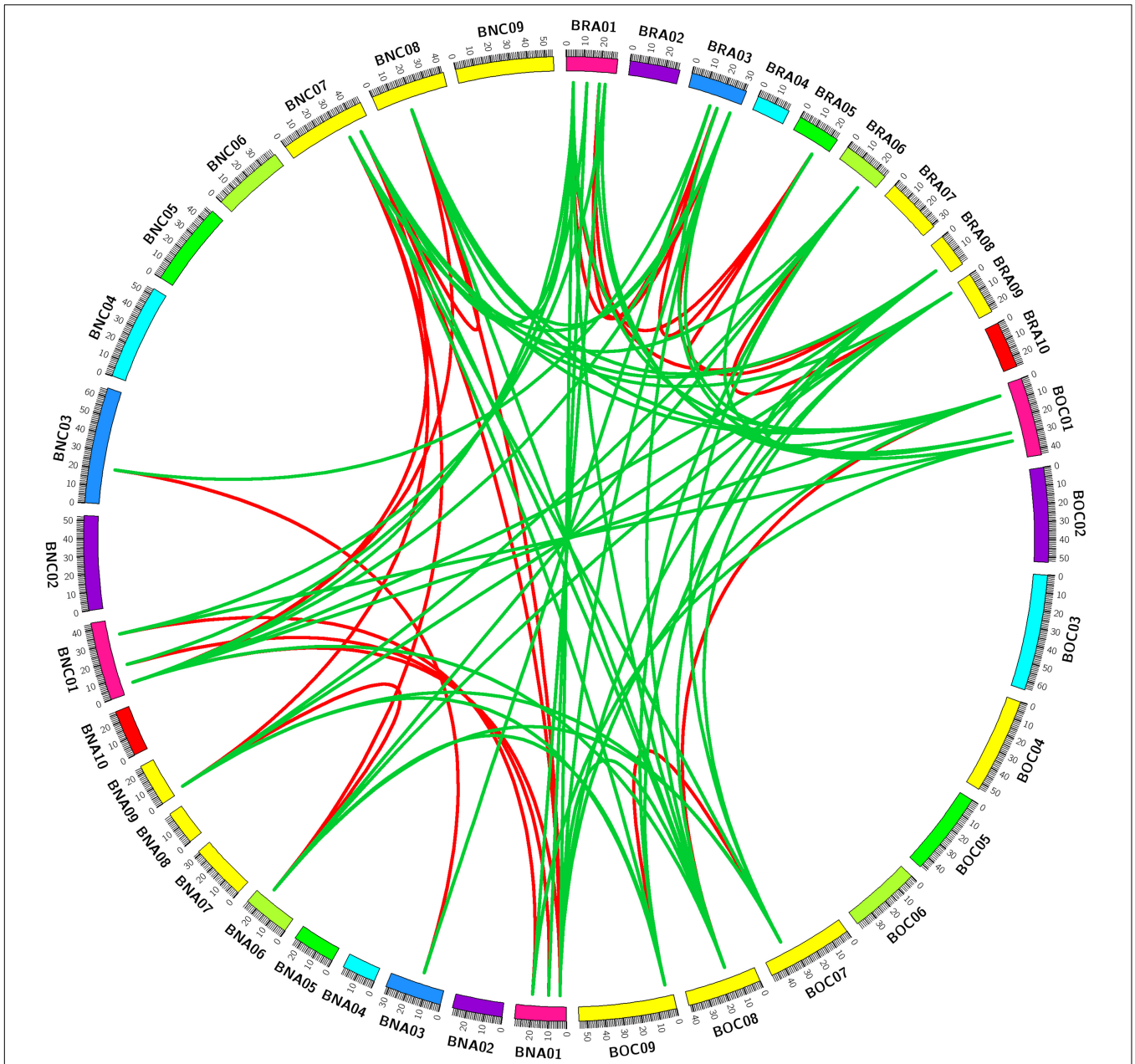


FIGURE 1 | Synteny analysis of *CPP* genes in genomes of *B. rapa*, *B. oleracea,* and *B. napus*. Red lines and green lines separately linked the syntenic paralogs and orthologs. BRA01–10, ten chromosomes in *B. rapa*; BOC01–09, nine chromosomes in *B. oleracea*; BNA01–10, ten chromosomes in A_n sub-genomes in *B. napus*; BNC01–09, nine chromosomes in C_n sub-genomes in *B. napus*.

implied that there might be more chromosomal rearrangement or homologous exchange events occurring in regions associated with these gene family members in the C_n subgenome than in the A_n subgenome.

Evolutionary Changes in Gene Structure of *CPP* Genes

We wanted to know what evolutionary changes had taken place in the gene structure of *CPP* genes after allopolyploidization. To better analyze these changes in terms of the evolutionary relationship, we first conducted phylogenetic analysis of the CPP gene family. CPP proteins of typical monocotyledon rice and dicotyledon Arabidopsis were used as reference proteins, and a phylogenetic tree was constructed based on 78 CPP protein sequences, among which there were 8, 11, 15, 10, and 34 CPP proteins in Arabidopsis, rice, *B. rapa*, *B. oleracea* and *B. napus*, respectively (**Figure 3**). These CPP proteins can be clearly divided into two clades (I, II) and six subclades (Ia-c, IIa-c), which was consistent with CPP phylogenetic tree in

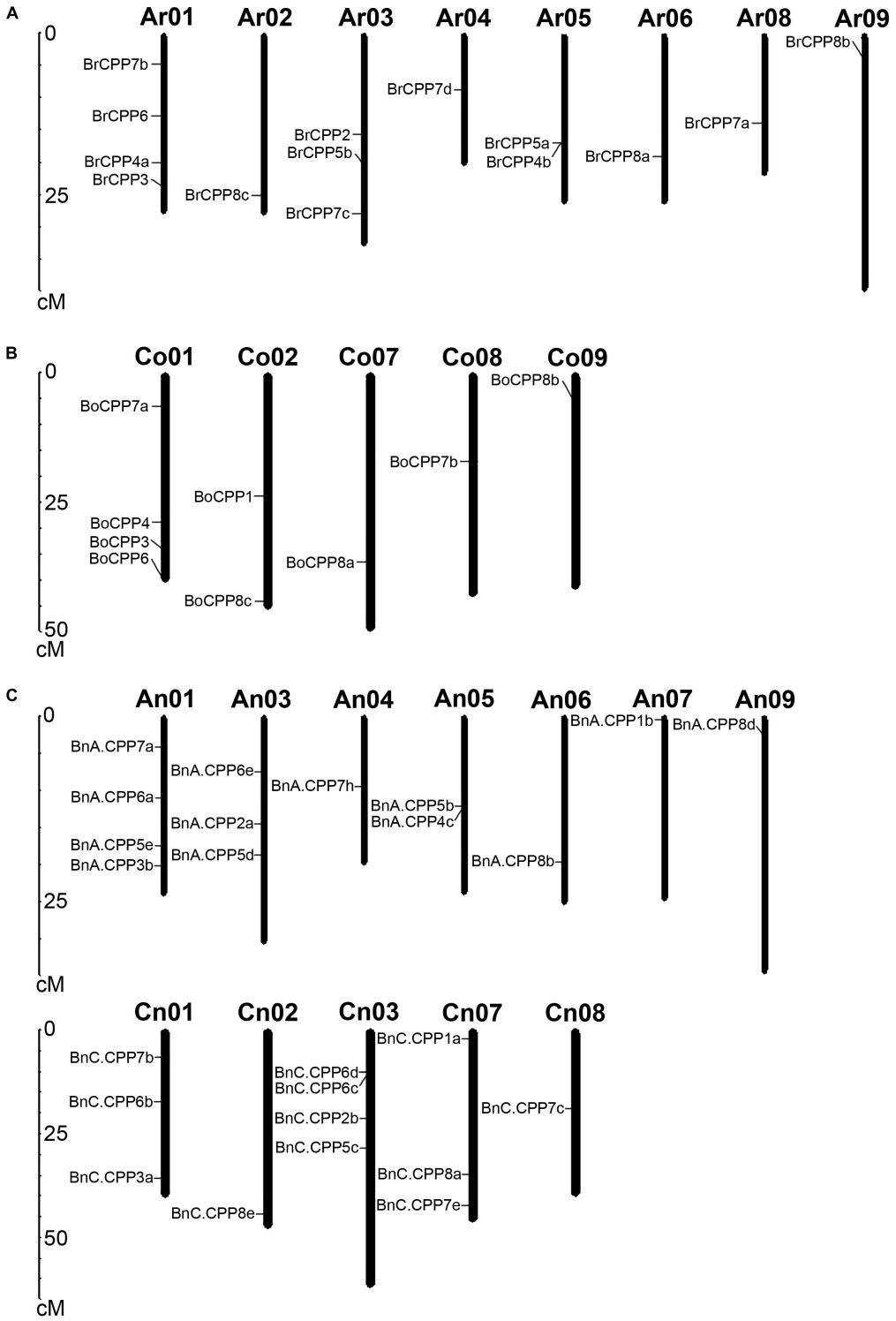


FIGURE 2 | Chromosome distribution of CPP genes in *B. rapa* **(A)** and *B. oleracea* **(B)** and *B. napus* **(C)**. All genes located in unassembled scaffolds were not shown in this figure.

other species (Zhou et al., 2018). We found that the CPP protein of Arabidopsis was present in all six subclades, whereas the CPP protein of rice was found in only three subclades (Ia, Ic, IIb), which might indicate that the evolution of CPP family has gradually diverged after the differentiation of monocotyledon and dicotyledon.

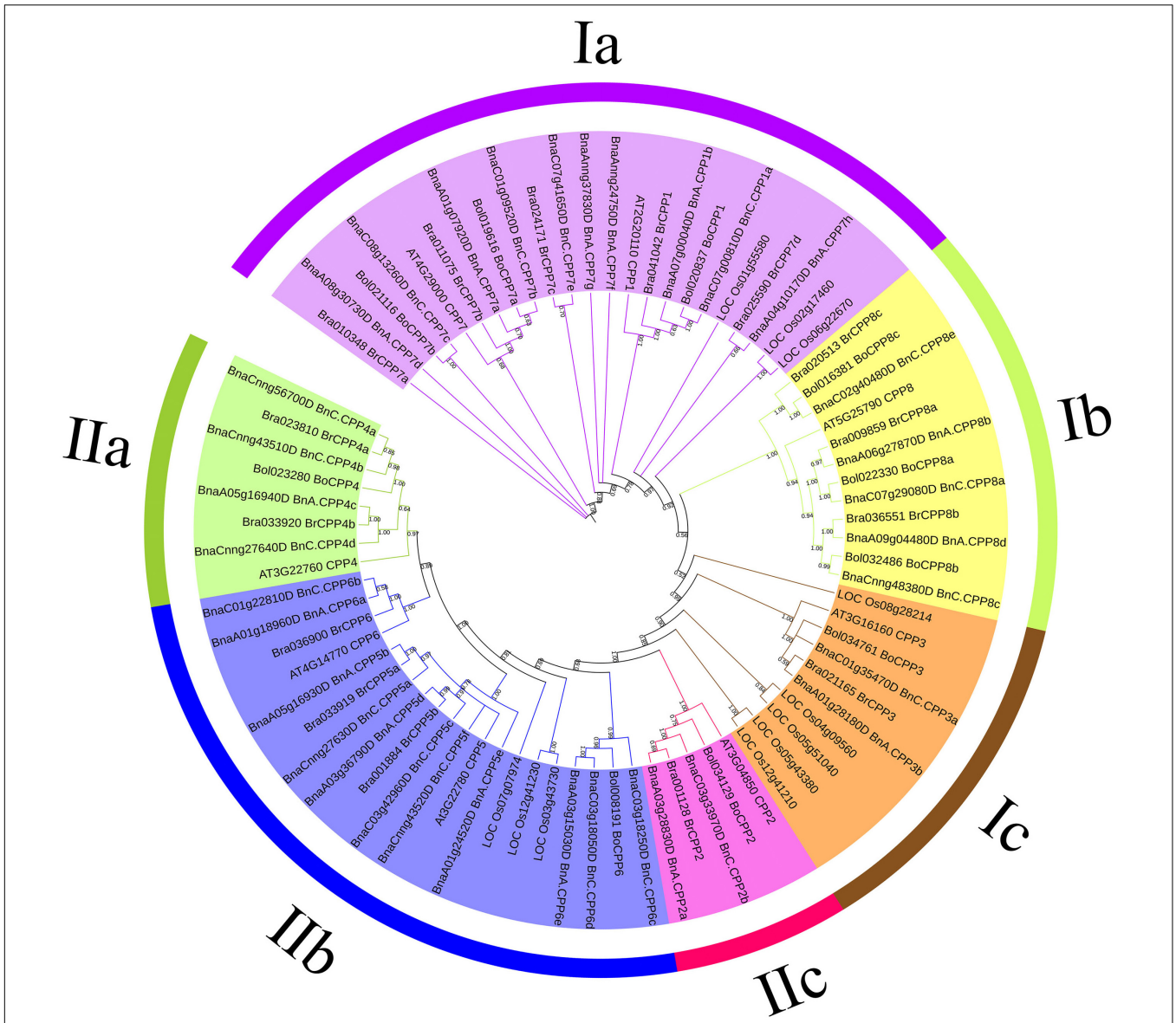


FIGURE 3 | Phylogenetic tree of CPP proteins in *B. napus, B. rapa, B. oleracea, Arabidopsis thaliana,* and *Oryza sativa*. The prefixes Bra, Bol, Bna, AT, and LOC stand for *B. rapa*, *B. oleracea*, *B. napus*, *A. thaliana* and *O. sativa*, respectively. Only bootstrap values greater than 50% were displayed. All CPP proteins were divided into two clades (I,II) and six subclades (Ia-c,IIa-c).

Next, we analyzed the structure of *CPP* genes in *B. napus* and its two diploid progenitors. Compared with many other gene families, members of the CPP gene family have a high diversity of gene structure (**Figure 4**). We compared the structural diversity of genes in two clades of the phylogenetic tree, and found that both the structural diversity and the mean number of introns of *CPP* gene in clade II were higher than those in clade I, indicating that the structure of *CPP* genes in clade I may be more conserved than that in clade II. In addition, we found that *BnCPP* genes had the highest intron number and structural diversity, while *BoCPP* genes had the lowest intron number and structural diversity, indicating that *CPP* genes in *B. napus* evolved toward increasing intron number and structural diversity after allopolyploidization. To further explore the evolutionary changes in the structure of *CPP* genes after allopolyploidization, we identified 20 homoeologous gene pairs with potential direct evolutionary relationship and conducted comparative analysis of their gene structure according to Li et al. (2019a; 2019b). Among the 20 homoeologous gene pairs, the intron number of 13 gene pairs (65%) was the same between *B. napus* and its diploid progenitors, which suggested that *CPP* genes are conserved to a certain extent at the DNA level after allopolyploidization. Moreover, among the remaining 7 gene pairs, there were 6 gene pairs showed the number of introns in *BnCPP* gene was greater than that in *BrCPP* or *BoCPP* genes. These results indicated that intron increase or loss occurred in some *CPP* genes after

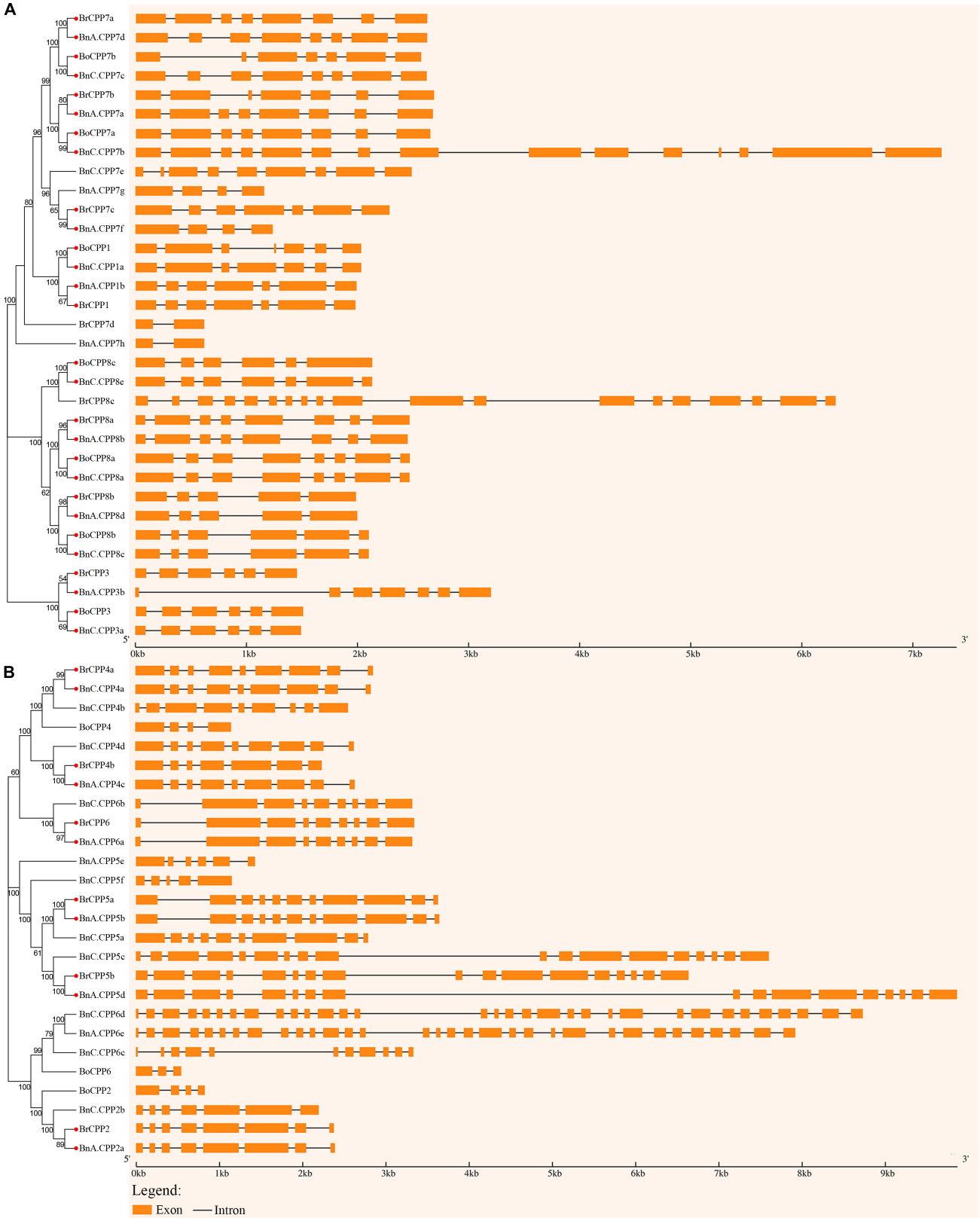


FIGURE 4 | Intron-exon structure of the *CPP* genes in *B. napus* and diploid progenitors. The genes on two clades, I **(A)** and II **(B)**, of the phylogenetic tree were shown separately. The red circle indicates orthologs with potential direct evolutionary relationship.

allopolyploidization, and the phenomenon of intron increasing was more common.

The CXC Domain and Motif Analysis of CPP Proteins

Analysis of CXC domains in *B. napus* and its progenitors showed that most CPP proteins in the three *Brassica* species contained one or two conserved CXC domains, just like CPP proteins in other species (**Supplementary Figure 4**). However, there were four CPP proteins (BnC.CPP5c, BrCPP5b, BnA.CPP5d, BnC.CPP6c) had four CXC domains, which have not been observed in CPP proteins of other species. This phenomenon might be related to the occurrence of homoeologous exchange events during allopolyploidization (Chalhoub et al., 2014). The alignment results of CXC domains in all identified CPP proteins showed that the two CXC domains and the short-conserved sequence RNPXAFXPK between them were very conserved, and only a few CPP proteins had amino acid deletion or replacement at these locations (**Supplementary Figure 5**). The motif analysis showed that 6 motifs existed in almost all CPP proteins, which were motifs 1–5 and 9 (**Supplementary Figure 6**). By analyzing the sequences of these motifs, we found that motifs 2, 4, and 9 were associated with the N-terminal CXC domain of CPP proteins, motif 5 was associated with a short-conserved sequence between two CXC domains, and motifs 1 and 3 were associated with the C-terminal CXC domain of CPP proteins. The motifs of proteins encoded by 20 homoeologous gene pairs with potential direct evolutionary relationship were compared, and we found 15 (75%) BnCPP proteins maintained the similar motif type and arrangement pattern of BrCPP/BoCPP proteins, indicating that the motif of CPP proteins were relatively conservative after allopolyploidization.

Physicochemical Properties and Subcellular Localization Prediction of CPP Protein

We predicted the physicochemical properties of CPP proteins in three *Brassica* species (**Table 3**), and the results were as follows. The average lengths of BnCPP, BrCPP and BoCPP proteins were 603, 592, and 381 aa, respectively, and the average molecular weights were 66.2, 65.0, and 41.5 kDa, respectively, which means that the average length and average molecular weight of BnCPP proteins were higher than those of its two progenitors. Furthermore, BnC.CPP7b and BoCPP7a were homoeologous gene pairs with potential direct evolutionary relationship, and we found that the protein sequence length of BnC.CPP7b (1408 aa) is much longer than that of BoCPP7a (578 aa), indicating that the encoding region of *CPP* gene might have increased after allopolyploidization. It was noteworthy that the length of 6 CPP proteins over 1000 aa were found in three *Brassica* species, and 5 of them were belonging to clade II in phylogenetic tree (**Figure 3**). In the three *Brassica* species, the instability index of all CPP proteins was higher than 40, and the GRAVY of all CPP proteins (except BoCPP2) was lower than 0, indicating that almost all CPP proteins were unstable hydrophilic proteins. The results showed that 44.12% of BnCPP, 53.33% of BrCPP and 30% of BoCPP proteins were acidic proteins (pI < 7), indicating that the ratio of acidic CPP protein in *B. napus* was between the two diploids. In addition, except for subclade Ic, acidic CPP proteins were found in all the other clades, suggesting that this clade might have diverged from other clades in the process of evolution. Subsequently, we predicted the subcellular localization of CPP proteins in the three *Brassica* species (**Table 3**). The results showed that the subcellular localization of CPP protein in three *Brassica* species was the same as that in other species (Song et al., 2016; Zhou et al., 2018). Most CPP proteins in *B. napus* and its diploid progenitors were predicted to be located in the nucleus, which may be consistent with the function of CPP proteins as transcription factors.

Prediction of *Cis*-Acting Elements in the Promoter Region of *CPP* Genes

We predicted the *cis*-acting elements in the upstream promoter region of *CPP* genes in three *Brassica* species (**Figure 5**). Fifty-nine *cis*-acting elements related to plant development and growth, phytohormone responses, stress responses and light responses were identified in the promoter region of *CPP* genes in *B. napus* and its progenitors. Among them, the largest number of *cis*-elements were associated with light responses. Seven identical stress response elements were identified in each of the three *Brassica* species, and the antioxidant response element (ARE) was found in the promoter region of 52 *CPP* genes, which was the most common type of *cis*-acting elements. We hypothesized that these stress-related elements, especially the antioxidant elements, existed in two diploid *Brassica* progenitors and were entirely preserved during allopolyploidization. We also found some elements related to development and growth that could regulate cell cycle (MSA-like) and affect expression of meristem (CAT-box), and meanwhile, *CPP* genes have been shown to participate in the coordination of cell cycle and fate (Wang et al., 2018). In addition, a large number of hormone-responsive elements have been found, including methyl jasmonate response elements (CGTCA-motif), abscisic acid response elements (ABRE) and salicylic acid response elements (TCA-element). We further analyzed and compared the *cis*-acting elements of the promoter regions of 20 homoeologous gene pairs mentioned above, and the results showed that there were four gene pairs (*BrCPP7b* and *BnA.CPP7a*, *BoCPP1* and *BnC.CPP1a*, *BrCPP8a* and *BnA.CPP8b*, *BoCPP8a* and *BnC.CPP8a*) have the same *cis*-acting element type, indicating that the *cis*-acting elements of the promoter regions in these *CPP* genes were conserved during allopolyploidization. Moreover, eight *BnCPP* genes were found to have more *cis*-element types or numbers than their diploid progenitors, which might contribute to the regulatory role of *CPP* gens in the genetics, development and evolution of *B. napus*.

Analysis of the Expression Patterns of *CPP* Genes

To explore the expression patterns of *CPP* genes, we analyzed the expression patterns of *CPP* genes in four organs (stems, leaves, flowers and siliques) in *B. napus* and its diploid progenitors (**Figure 6**). *CPP* genes were widely expressed in the four

TABLE 3 | The physicochemical parameters and subcellular localization prediction of CPP proteins in *B. napus* and its diploid progenitors.

Protein name	Protein length (aa)	MW (kDa)	pI	Instability index	Gravy	Subcellular localization
BnA.CPP1b	506	54.25883	7.13	52.07	–0.618	Nucleus
BnA.CPP2a	522	57.31421	4.95	60.01	–0.435	Nucleus
BnA.CPP3b	340	38.12622	7.93	53.21	–0.607	Nucleus
BnA.CPP4c	574	62.93197	8.31	67	–0.624	Nucleus
BnA.CPP5b	748	82.16356	8.04	60.97	–0.629	Nucleus
BnA.CPP5d	1216	133.79573	7.8	65.78	–0.657	Nucleus
BnA.CPP5e	286	31.9237	8.62	58.3	–0.329	Extracell
BnA.CPP6a	647	71.22856	5.5	68.84	–0.719	Nucleus
BnA.CPP6e	1300	142.52083	6.86	46.9	–0.595	Nucleus
BnA.CPP7a	583	62.11057	8.63	53.37	–0.635	Nucleus
BnA.CPP7d	573	61.27621	7.72	58.92	–0.667	Nucleus
BnA.CPP7f	287	31.24956	9.17	53.6	–0.614	Nucleus
BnA.CPP7g	265	28.59231	9.1	53.52	–0.689	Nucleus
BnA.CPP7h	142	15.87176	6.06	45.94	–0.585	Cytoplasm
BnA.CPP8b	510	56.64173	6.02	54.58	–0.922	Nucleus
BnA.CPP8d	455	50.42734	6.55	53.2	–0.853	Nucleus
BnC.CPP1a	495	53.42396	7.44	50.38	–0.623	Nucleus
BnC.CPP2b	543	60.0745	5.07	59.62	–0.357	Chloroplast
BnC.CPP3a	328	36.63148	7.68	53.09	–0.626	Nucleus
BnC.CPP4a	604	66.34732	6.9	67.44	–0.683	Nucleus
BnC.CPP4b	596	65.46938	7.44	66.98	–0.626	Nucleus
BnC.CPP4d	549	60.17399	8.33	67.36	–0.576	Nucleus
BnC.CPP5a	679	74.87134	6.67	58.09	–0.654	Nucleus
BnC.CPP5c	1241	137.16818	6.84	62.82	–0.69	Chloroplast
BnC.CPP5f	264	29.9765	9.58	58.62	–0.671	Nucleus
BnC.CPP6b	657	71.95252	5.49	65.58	–0.656	Nucleus
BnC.CPP6c	323	35.73944	8.67	58.07	–0.827	Nucleus
BnC.CPP6d	1296	141.3965	6.87	46.71	–0.585	Nucleus
BnC.CPP7b	1408	155.07617	6.42	58.2	–0.744	Vacuole
BnC.CPP7c	566	60.36419	7.44	56.95	–0.639	Nucleus
BnC.CPP7e	558	60.71411	8.35	64.69	–0.745	Nucleus
BnC.CPP8a	519	57.41975	6.54	51.48	–0.888	Nucleus
BnC.CPP8c	458	50.79463	7.09	51.89	–0.898	Nucleus
BnC.CPP8e	478	52.64484	6.22	58.45	–1.014	Nucleus
BoCPP1	385	41.55645	5.7	51.38	–0.677	Nucleus
BoCPP2	175	19.54062	7.36	40.42	0.066	Cytoplasm
BoCPP3	330	36.89163	8.1	52.75	–0.693	Nucleus
BoCPP4	265	29.24992	7.94	69.21	–0.631	Nucleus
BoCPP6	130	14.11321	8.94	50.85	–0.606	Nucleus
BoCPP7a	578	61.33569	8.56	55.56	–0.635	Nucleus
BoCPP7b	464	49.00643	7.17	59.94	–0.597	Nucleus
BoCPP8a	518	57.36174	6.66	50.39	–0.895	Nucleus
BoCPP8b	457	50.72355	7.09	51.8	–0.904	Nucleus
BoCPP8c	504	55.66243	6.27	57.74	–0.95	Nucleus
BrCPP1	503	53.9774	7.14	52.58	–0.639	Nucleus
BrCPP2	521	57.32656	5.06	62.28	–0.369	Nucleus
BrCPP3	326	36.54725	7.7	51.68	–0.708	Nucleus
BrCPP4a	592	64.90159	6.17	66.52	–0.651	Nucleus
BrCPP4b	575	63.43283	8.66	70.18	–0.539	Chloroplast
BrCPP5a	748	82.23362	7.88	61.8	–0.628	Nucleus
BrCPP5b	1223	134.46711	6.74	63.78	–0.662	Nucleus
BrCPP6	645	71.09954	5.51	68.35	–0.693	Nucleus
BrCPP7a	574	61.44543	7.93	59.66	–0.664	Nucleus
BrCPP7b	528	56.58224	8.51	54.86	–0.729	Nucleus
BrCPP7c	552	60.40711	8.84	52.84	–0.77	Nucleus
BrCPP7d	142	15.87176	6.06	45.94	–0.585	Cytoplasm
BrCPP8a	509	56.50368	5.81	53.9	–0.892	Nucleus
BrCPP8b	456	50.6566	6.68	52.3	–0.861	Nucleus
BrCPP8c	990	110.14645	5.73	50.22	–0.592	Nucleus

organs, suggesting that the CPP gene family might have multiple biological functions and might play roles in different organs. Two *BnCPP* genes (*BnA.CPP5e* and *BnC.CPP5f*) were not expressed in the four organs, suggesting that they might not play a role in these organs or might be expressed in specific time and space. The expression patterns of different *CPP* genes were various in the four organs. For example, the expression level of *BrCPP1* was relatively high in flowers, while the expression level of *BrCPP4a* and *BnC.CPP4a* was relatively high in siliques. Based on the expression data of 20 homoeologous gene pairs with potential direct evolutionary relationship, we compared the expression patterns between *BnCPP* and *BrCPP*/*BoCPP* genes to explore the evolutionary changes of *CPP* gene expression during allopolyploidization. We found most of these gene pairs showed different expression patterns in these organs. The expression of some *BnCPP* genes and their ancestral homoeologous genes in some organs was completely opposite. For example, the expression levels of *BrCPP1* and *BoCPP8a* were the highest in flowers, while the expression levels of *BnA.CPP1b* and *BnC.CPP8a* were the lowest in flowers. Similarly, the expression level of *BoCPP8c* was the lowest in leaves, while the expression level of *BnC.CPP8e* was the highest in leaves. In addition, some *CPP* genes were expressed at roughly the same level in different organs, while the expression of their homoeologous genes were significantly different in these organs. For example, the expression level of *BoCPP7a* in leaves and siliques was identical, while the expression level of *BnC.CPP7b* in leaves was significantly lower than that in siliques. Moreover, one gene of the homoeologous pairs was expressed and the other was not expressed in some organs. For example, *BnA.CPP3b* was expressed only in stems, while its homoeologous gene *BrCPP3* was expressed in flowers and leaves, but not in stems. These results indicated that the expression patterns of *CPP* genes might be changed in the process of allopolyploidization. By comparing the expression levels of 13 homoeologous gene pairs with the same gene structure in the four organs, it was found that the majority of them had different expression patterns, that is, the gene expression might change though there was no significant change at the DNA level, which might be related to the regulation of gene expression during allopolyploidization.

To study the expression bias of *CPP* genes in different organs of allotetraploid *B. napus*, 20 homoeologous gene pairs were classified into eight groups (*CPP1*–*CPP8*) based on their homology. There were four groups (*CPP2*, *CPP4*–*CPP6*) in which the ancestor genes were not present in both *B. rapa* and *B. oleracea*, so they could not be analyzed in this way. Finally, four groups of genes (*CPP1*, *CPP3*, *CPP7,* and *CPP8*) were used for the gene expression bias analysis. The results showed that the expression bias of *BnCPP* genes was inconsistent in different organs (**Supplementary Table 5**). Specifically, the expression of *BnCPP8* was biased toward *B. rapa* in all four organs. The expression of *BnCPP3* was biased toward *B. rapa* in flowers, leaves and siliques, while which was biased toward *B. oleracea* in stems. The expression of *BnCPP7* was also biased toward *B. rapa* in three organs (flowers, stems and siliques), and was biased toward *B. oleracea* in leaves. The expression of *BnCPP1* was biased toward *B. rapa* only in leaves, and was biased toward

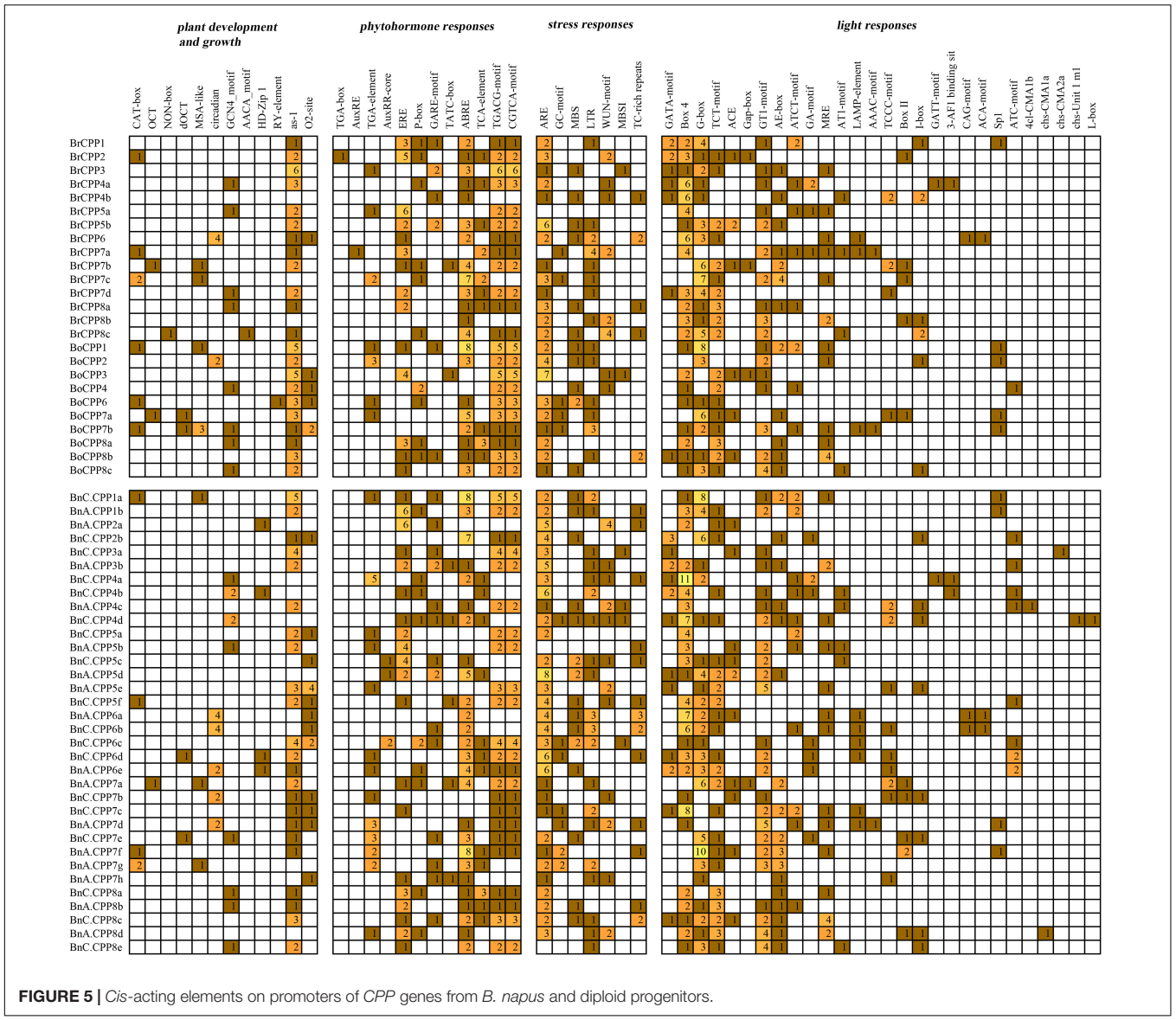


FIGURE 5 | *Cis*-acting elements on promoters of *CPP* genes from *B. napus* and diploid progenitors.

B. oleracea in the other three organs (flowers, stems and siliques). We found that 75% of *BnCPP* gene expression was biased toward *B. rapa* in flowers, leaves and siliques, while this proportion was only 50% in stems.

CPP Gene Expression Changes of Three *Brassica* Species Under Salt Stress

Previous studies have shown that CPP gene family can respond to a variety of abiotic stresses, such as low temperature and drought stress. To investigate whether the CPP gene family in three *Brassica* species responds to salt stress, we analyzed the expression of some *CPP* genes under salt stress by quantitative real-time PCR (qRT-PCR). *CPP* genes whose expression was not detected in the experimental material were not analyzed in further study, and we finally selected 12 *CPP* genes for the qRT-PCR experiment (**Supplementary Table 1**). The expression levels of 11 *CPP* genes were significantly changed after salt stress, suggesting that they might be involved in the response to salt stress (**Figure 7**). Further analysis showed that the expression levels of 8 *CPP* genes increased significantly after salt stress, while the expression levels of 3 *CPP* genes decreased significantly. These results suggest that the CPP gene family might adapt to salt stress through two different response mechanisms (positive and negative), and positive regulation might be the main regulatory mode. It was noted that the expression levels of some *CPP* genes increased sharply after salt stress. For example, the expression levels of *BoCPP1* and *BoCPP4* were upregulated by 140 and 80 times, respectively, and their expression levels were much higher than those of other up-regulated genes. Therefore, it was speculated that these two *BoCPP* genes might play dominant roles in salt stress response.

In addition, there were three homoeologous gene pairs that have potential direct evolutionary relationship (*BrCPP1*

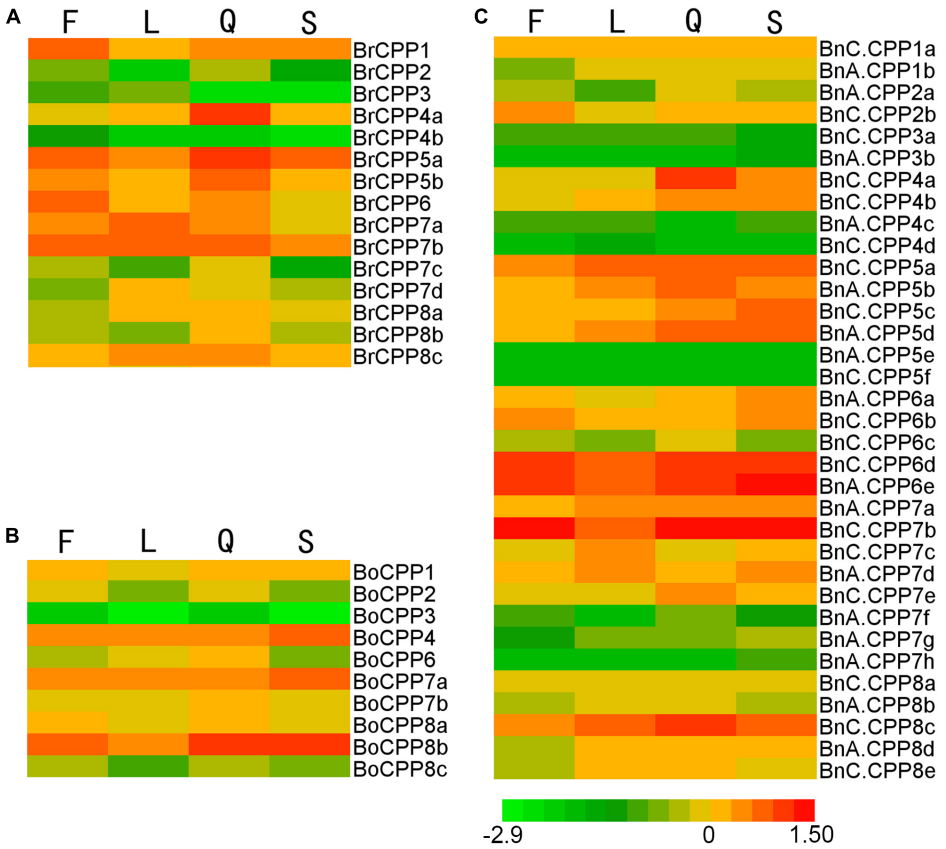


FIGURE 6 | Expression patterns of CPP genes in flowers (F), leaves (L), siliques (Q) and stems (S) from *B. rapa* **(A)**, *B. oleracea* **(B)**, and *B. napus* **(C)**.

and *BnA.CPP1b*, *BrCPP4a* and *BnC.CPP4a*, *BoCPP8c* and *BnC.CPP8e*) among the 11 *CPP* genes whose expression levels changed significantly after salt stress, and the changes of their expression levels after salt stress were compared and analyzed. We found that the expression of the three *BnCPP* genes and their homoeologous genes in two diploid progenitors showed the same trend after salt stress. In particular, the expression of *BrCPP1* and *BnA.CPP1b* was down-regulated after salt stress, and the expressions of *BrCPP4a* and *BnC.CPP4a* and *BoCPP8c* and *BnC.CPP8e* were up-regulated after salt stress. However, the expression levels of these *BnCPP* genes and their corresponding homoeologous genes in two diploid progenitors were different in degree. For example, the expression of *BnC.CPP4a* was up-regulated about 12.5 times, while the expression of *BrCPP4a* was up-regulated only about 4.8 times after salt stress, and the expression of *BnC.CPP8e* increased 2.8 times, while the expression of *BoCPP8c* increased 10.3 times after salt stress. These results suggested that the response degree of *CPP* genes to salt stress might change after allopolyploidization.

Physiological Changes of Three *Brassica* Species Under Salt Stress

The superoxide dismutase (SOD) activity of the three *Brassica* species increased after salt stress, among which the SOD activity of *B. rapa* increased the most, which was 1.7 times of the control, and the SOD activity of *B. oleracea* and *B. napus* was 1.18 times and 1.12 times of the control, respectively (**Figure 8**). After salt stress, peroxidase (POD) activity of *B. oleracea* and *B. napus* increased by 1.62 and 1.43 times, respectively, while POD activity of *B. rapa* decreased. Similarly, after salt stress, POD activity of *B. oleracea* and *B. napus* increased by 1.63 and 1.35 times, respectively, while catalase (CAT) activity of *B. rapa* decreased. Therefore, the activities of all three enzymes were increased to remove reactive oxygen in *B. oleracea* and *B. napus* after salt stress, while only the activity of SOD was increased in *B. rapa*, and the activities of the other two enzymes were restricted. These results suggested that the response of antioxidant enzyme systems in three *Brassica* species to salt stress might be different.

Plants can also respond to adverse conditions by accumulating osmotic substances. Soluble protein accumulation occurred in all three *Brassica* species after salt stress, which was most obvious in *B. rapa* (**Figure 8**). For details, after salt stress, the soluble protein content of *B. rapa* was 1.65 times that of the control, while which of *B. oleracea* and *B. napus* was 1.08 and 1.05 times that of the control, respectively. The soluble sugar content of three *Brassica* species increased after salt stress, and which of *B. oleracea* increased most. The soluble sugar content of *B. oleracea* was 1.23 times that of the control, and which of *B. rapa* and *B. napus* was 1.09 and 1.04 times that of the

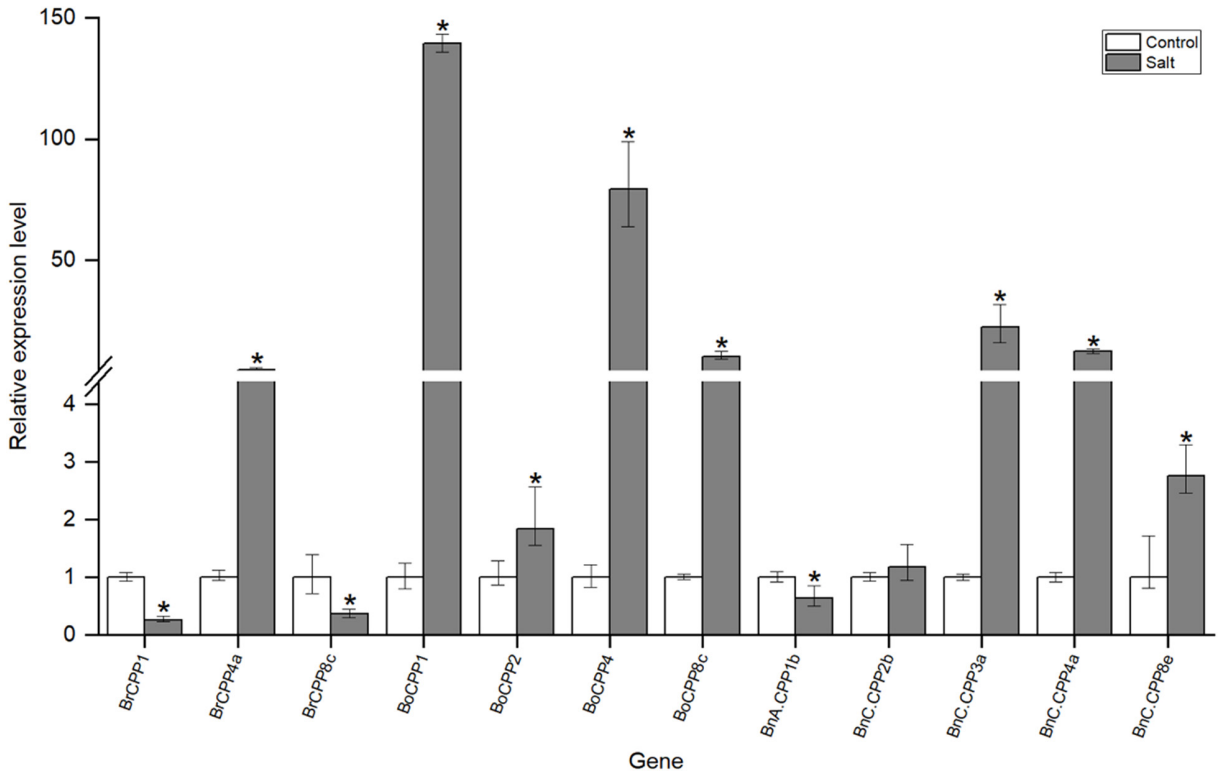


FIGURE 7 | Expression of *CPP* genes in leaves of *B. rapa, B. oleracea,* and *B. napus*. Control: the control group; Salt: salt treatment group; "*" represents a significant difference between the two groups.

control, respectively. In addition, we observed an interesting phenomenon that the content of proline in the three *Brassica* species increased sharply after salt stress, and the content of proline in *B. rapa*, *B. oleracea,* and *B. napus* were 4.7, 40.9 and 26.3 times higher than that of the control, respectively. These results indicated that these substances played the corresponding osmotic regulation functions to alleviate the salt stress injury of *B. napus* and its progenitors. Moreover, after salt stress, chlorophyll content in leaves of *B. rapa* decreased, but which of *B. oleracea* and *B. napus* increased, suggesting that the three species might have different salt-stress response mechanisms.

DISCUSSION

Polyploidization of plants can lead to the enhancement of their environmental adaptability (Segraves, 2017), and the formation and evolution of polyploids are often accompanied by series changes in plant morphology, physiology, genetics and epigenetics (Soltis et al., 2016). A large number of studies have been conducted on the large-scale genetic and epigenetic changes of plants after allopolyploidization, but there were few studies focused on the molecular basis of environmental adaptability on allopolyploids from the perspective of gene family. The allopolyploid *B. napus* and its two diploid progenitors (*B. rapa* and *B. oleracea*) was an appropriate system to study the scientific questions about allopolyploidization. As the transcription factor gene family, CPP gene family played a role in stress response of plants, and this gene family was not reported in *Brassica* species. In this study, the characteristics and the salt response of CPP gene family in *B. napus* and its progenitors were compared, which was conducive to exploring the genetic evolution characteristics of CPP gene family and detecting the potential advantageous molecular basis of stress response in allopolyploid *B. napus*.

Expansion of *CPP* Gene Family in Allopolyploid *Brassica napus*

The phenomenon of gene family expansion is widespread in the process of plant evolution. For example, the majority of gene families (about 80%) in *Arabidopsis thaliana* showed expansion in their evolutionary course (Lespinet et al., 2002). In this study, 15, 10, and 34 *CPP* genes were identified in *B. rapa*, *B. oleracea,* and *B. napus*, and the number of genes in this three *Brassica* species can be analyzed as follows. *Brassica* and *Arabidopsis* diverged from a common ancestor about 14 million years ago (MYA; Song et al., 2020). All 8 *CPP* genes in Arabidopsis had corresponding homoeologous genes in *B. rapa*, *B. oleracea,* and *B. napus*, indicating that the progenitors of *Brassica* retained these *CPP* genes completely after being differentiated from the evolutionary lineage of *Arabidopsis*. After separating from the evolutionary lineage of *Arabidopsis*, the progenitors of *Brassica* underwent a unique whole genome triplication (WGT) event and divided into two diploid species *B. rapa* and *B. oleracea* about 3

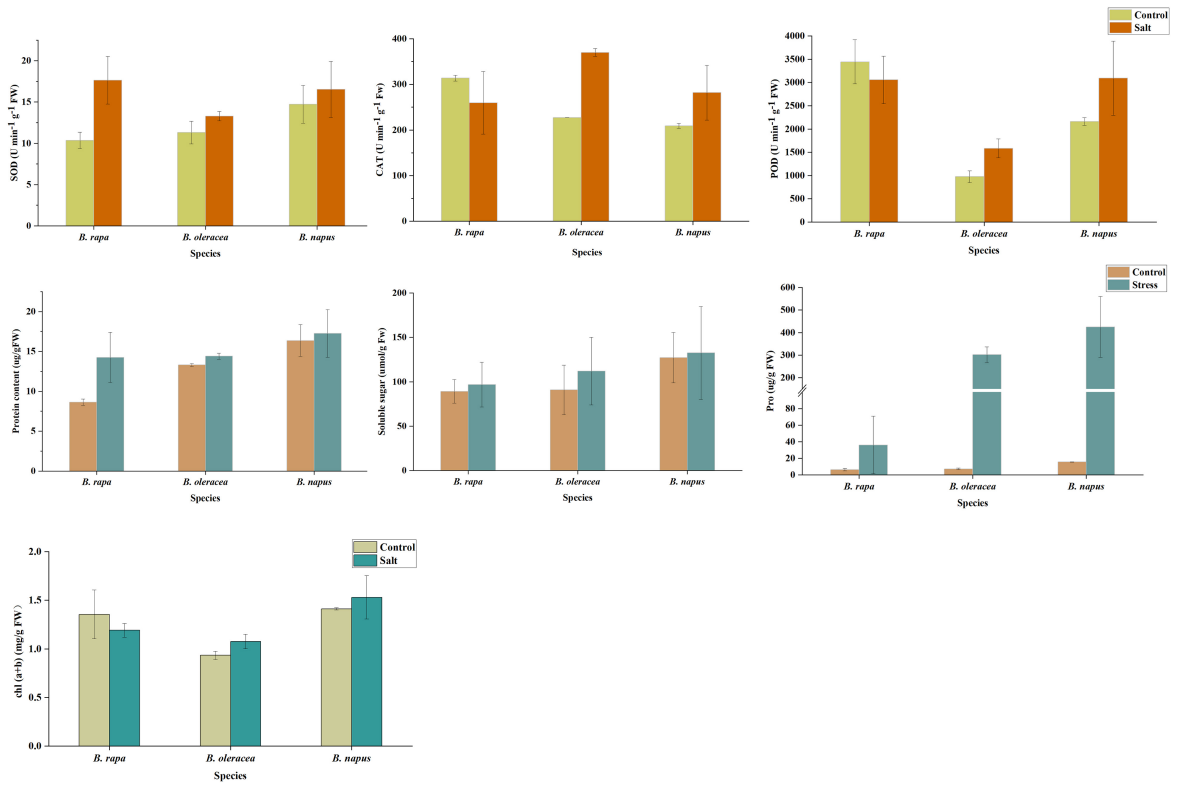


FIGURE 8 | The physiological responses of *B. napus* and diploid progenitors to salt stress.

MYA (Song et al., 2020). Therefore, in theory, if the progenitors of *Brassica* had not lost genes after the WGT event, there should be three times as many *CPP* genes in *B. rapa* and *B. oleracea* as in Arabidopsis. However, we found that many *CPP* genes in Arabidopsis have less than three homoeologous genes in diploid *B. rapa* and *B. oleracea* (**Table 1**). Therefore, due to the WGT events, the CPP gene family in *Brassica* expands compared to *Arabidopsis*, and at the same time, gene loss events occur in large numbers. In addition, the allotetraploid *B. napus* was formed by hybridization and genome doubling (i.e., allopolyploidization) of two diploid *B. rapa* and *B. oleracea* about 7,500 years ago (Chalhoub et al., 2014). Theoretically, the number of CPP gene family members in *B. napus* should be equal to the sum of the number of CPP members in the diploid *B. rapa* and *B. oleracea*. In this study, we found that the total number of CPP gene family members in *B. napus* was more than the sum of the two diploid progenitor species ($34 > 15 + 10$), that is, the CPP gene family also expanded in *B. napus* during allopolyploidization. Moreover, loss of *CPP* genes, such as *CPP8* homoeologues gene ($5 < 3 + 3$), was also found during allopolyploidization. From the above analysis, it can be concluded that the CPP gene family expanded in *B. napus* during evolution, despite the occurrence of gene loss events.

The main duplication modes that cause gene family expansion might be different. Analysis of gene duplication modes in gene families can provide explanations for the expansion of these gene families (Liang et al., 2017). Cannon et al. (2004) studied the duplication modes of 50 gene families in *Arabidopsis thaliana*, and found that the tandem duplication and segmental duplication were dominant in many gene families. In this study, the duplication modes of the *CPP* genes were studied, and we found that in addition to a large number of WGD, many TRD were detected in *CPP* genes of three *Brassica* species. In addition, unlike the two diploid *Brassica* species, the proportion of TRD in *CPP* genes was higher than that of WGD in *B. napus*. Therefore, it was speculated that TRD may be an important driving force for the expansion of CPP gene family in *B. napus* during allopolyploidization.

Intron Acquisition Events Were Observed in *BnCPP* Genes After Allopolyploidization

As an important part of structural genes in eukaryotes, the evolution and function of introns are particularly important. In the study of gene families, analyzing the gene structural changes (changes of the number of exon/intron) may help us to understand the evolutionary direction of gene families. In this study, a comparative analysis of the gene structure of 20 homoeologous gene pairs with potential direct evolutionary relationship was conducted. Our results showed that except for 13 homoeologous gene pairs that were highly conserved in intron number after allopolyploidization, the remaining 7 gene pairs showed differences in intron number, and 6 of them

showed more introns in *BnCPP* than in the corresponding ancestral species *BrCPP*/*BoCPP*, indicating that most of the *CPP* genes with structural changes acquired new introns after allopolyploidization. Similar phenomena have been observed in other gene families of *B. napus* (Li et al., 2019a). Earlier studies have suggested that introns might bring heavy burdens to the organism. For example, on the one hand, intron splicing required a high biological cost (Jiang and Goertzen, 2011). On the other hand, because introns were not subjected to the pressure of natural selection, a large number of mutations were often accumulated in introns, and accumulation of harmful mutations might adversely affect the adaptation of the organism (Lynch, 2006). However, recent studies have consistently suggested that the presence of introns might provide various benefits to the genome and organisms (Mukherjee et al., 2018), of which several have been discovered. First, introns could be used to improve protein diversity through alternative splicing and the fractionation of exons (Lamond, 1999). Second, introns might lead to the production of non-coding RNAs that were involved in some regulatory processes (Ying and Lin, 2005). Third, introns could also play a positive mediating role to enhance/promote the expression of some genes, as researchers have found that genes with higher expression levels have more and longer introns than genes with lower expression levels in rice and *Arabidopsis thaliana* (Ren et al., 2006). In addition, introns might also play a positive role in transcriptional coupling and mRNA output (Maniatis and Reed, 2002). Therefore, the increase of introns in *BnCPP* genes after allopolyploidization might be beneficial to the genome evolution of *B. napus*, but the specific effects still need further studies to prove.

The Expression Characteristics of *BnCPP* Genes Might Indicate the Evolutionary Advantage of *B. napus*

A series of genetic and epigenetic changes, occurring rapidly when polyploids were established, which not only help stabilize the plant genome, but also regulate the expression of genes in the plant (Pikaard, 2001). To understand whether gene expression changes occurred during the formation and evolution of allotetraploid *B. napus*, the expression pattern of *CPP* genes in three *Brassica* species was investigated in this study. The results showed that, on the one hand, except for two *BnCPP* genes, *CPP* genes were widely expressed in four organs of the three *Brassica* species, suggesting that the *CPP* genes might be indispensable in the genetic, developmental and evolutionary process of three species. On the other hand, we also observed various changes in *CPP* gene expression during allopolyploidization. For example, most gene pairs, despite having the same gene structure, showed divergence in their expression patterns. Therefore, it was speculated that there might be neofunctionalization or subfunctionalization of *CPP* genes during allopolyploidization, and these changes might have beneficial effects on the evolution of CPP gene family in *B. napus*. Moreover, the differentiation of gene function might also reduce the risk of redundant genes being eliminated. In addition, we found that the *BnCPP* gene showed different bias expression, which was different in four organs. For example, the expression of *BnCPP3* and *BnCPP7* in three organs were biased toward *B. rapa*, while the expression of *BnCPP1* in three organs was biased toward *B. oleracea*. It was speculated that such a variety of biased gene expression might help allopolyploid *B. napus* adapt better to the environment (Samans et al., 2017).

Previous studies have suggested that *CPP* genes may play a role in response to various abiotic stresses. For example, most *CPP* genes in the roots of soybean can respond to heat stress, and many *CPP* genes in the leaves of cucumber show different response patterns and can respond to various stresses such as low temperature, drought, ABA and salt (Zhang et al., 2015; Zhou et al., 2018). *B. napus*, as an important crop, may be affected by various adverse environment during its growth, among which salting environment is one of the main adverse environments. In this study, we investigated the expression of some *CPP* genes in *B. napus* and its progenitors under salt stress. The results showed that *CPP* gene expression levels were significantly changed in *B. napus* and its two progenitors after salt stress, suggesting that *CPP* genes in the three *Brassica* species might play important roles in the response of salt stress. In addition, *CPP* genes might be involved in the regulation of salt stress related gene expression through positive or negative regulation, and the former might be dominant. By comparing the expression changes of three homoeologous gene pairs with potential direct evolutionary relationship, we found that the expression trend of *CPP* genes in *B. napus* was consistent with that in the progenitors, but the degree of change was different. It was speculated that the function of *CPP* genes in response to salt stress might be well preserved after allopolyploidization, and the change in response degree might help *BnCPP* genes to respond better and faster to salt stress. Our study can provide reference for the study of molecular resistance mechanism of *B. napus*.

The Physiological Response of *B. napus* to Salt Stress Showed Its Resistance Advantage

Polyploids may have better stress resistance than diploids, and studies have shown that polyploids are more capable of producing beneficial substances and reducing harmful substances than diploids when subjected to abiotic stress. For example, more proline and less malondialdehyde accumulate in tetraploid rice than diploid rice under salt stress, and polyploid beet had lower Na^+/K^+ ratio than diploid beet in response to salt stress (Tu et al., 2014; Zhu, 2016; Wu et al., 2019). When plants are in an unfavorable environment, reactive oxygen species (ROS; such as O_2, H_2O_2, OH^-) accumulate in cells, which gradually damage cells and plants. In order to reduce this damage, plants have built a set of enzymes to maintain the balance of reactive oxygen species in cells. These enzymes include SOD, POD, and CAT. In this study, we systematically studied the physiological responses of *B. napus* and its progenitors to salt stress, including the activity of SOD, POD, and CAT, and the content of soluble protein, soluble sugar, proline, and chlorophyll. Plants can help them adapt to stress by scavenging ROS accumulated inside cells when they are stressed. Therefore, the increased activities

of SOD, POD, and CAT may be related to the response of plants to stress. Proline can be used as an osmotic regulator and hydroxyl radical scavenging agent to help stabilize the subcellular structure of plants, and the accumulation of proline contributes to the resistance of plants (Lv et al., 2011). In addition, plants produce a series of proteins related to stress resistance to avoid injury, and soluble proteins can help plants adapt to adversity as functional proteins or maintainers of osmotic potential (Wu et al., 2004). The increase of soluble sugar content is also a strategy for plants to adapt to adversity, as it not only acts as an osmotic regulator, but also provides carbon framework and energy for the synthesis of other organic solutes (Parankusam et al., 2017). As the basis for material and energy exchange with the outside world, chlorophyll content in plant leaves also changes under stress. The decrease of chlorophyll content was observed in *B. rapa* after salt stress in this study, which is consistent with most of the previous studies, however, the opposite phenomenon was observed in *B. oleracea* and *B. napus*, and the reason may be that the salt stress in our study does not cause rapid decomposition of chlorophyll in the leaves, and this phenomenon may also be related to chloroplast shrinkage (Chaves et al., 2009). In addition, compared with the two diploid progenitors, SOD activity and the accumulation of three osmotic regulatory substances were all the highest in *B. napus* before salt stress, and SOD activity was still relatively high and the accumulation of three osmotic regulatory substances was still the highest in *B. napus* after salt stress. Moreover, POD and CAT activities were inhibited and chlorophyll content decreased in *B. rapa*, while POD and CAT activities and chlorophyll content increased in *B. napus* under salt stress. These results suggest that not all physiological response pathways in allotetraploid *B. napus* were superior to their diploid progenitors *B. rapa* and *B. oleracea* in salt stress, but overall, *B. napus* showed a stronger and more stable physiological response in this study.

CONCLUSION

The CPP gene family in *B. napus* expanded after allopolyploidization, and the genome doubling and transposon mediated gene replication might be the main reasons for the expansion. All identified CPP members can be clearly divided into two clades (I, II), and the gene structure of *CPP* in clade I was more conserved than that in clade II. The intron increasing phenomenon was observed in some *CPP* genes after allopolyploidization. Although the gene structure of some *CPP* genes were not changed, the expression patterns of them were changed in *B. napus* after allopolyploidization. Some molecular basis might be associated with the adaptability of *B. napus*, including the expansion of the CPP gene family, the acquisition of introns by some *BnCPPs*, and abundant *cis*-acting elements upstream of *BnCPPs*. A total of 11 *CPP* genes potentially involved in the response to salt stress, and the response degree of some *CPP* genes to salt stress might change in *B. napus* after allopolyploidization. The physiological response of salt stress was stronger and more stable in *B. napus* than its two progenitors. Our study provides critical reference for exploring the potential advantageous molecular basis of various stress response in allopolyploids.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

ML conceived and drafted the manuscript, performed the data analysis, software analysis, and correction. FW performed the software analysis, surveyed the literature, validated experiments, and visualized the results. JM conducted software analysis, data analysis, and edited and revised the writing. HL performed the software analysis, surveyed the literature, and visualized the results. HY provided the methodology, performed the software analysis, and edited and revised the writing. PZ provided the methodology, guided the writing and editing, and prodded them. JW conceived the manuscript, did project management and data management, directed writing and editing, and provided funding. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (31970241).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.873071/full#supplementary-material

Supplementary Figure 1 | Phenotype figure of *B. rapa*, *B. napus,* and *B. oleracea* under salt stress.

Supplementary Figure 2 | Bar chart of the number of genes for five duplication modes.

Supplementary Figure 3 | Visual display of segmental duplicated gene pairs and five duplication modes. Blue lines, red lines, and gray lines represented the segmental duplicated gene pairs in *B. rapa*, *B. napus,* and *B. oleracea*, respectively. Red, yellow, green, blue, and pink circular markers represented WGD, TRD, TD, PD, and DSD, respectively.

Supplementary Figure 4 | The CXC domain location of CPP proteins in clade I **(A)** and clade II **(B)** from *B. napus* and its diploid progenitors.

Supplementary Figure 5 | Alignment of the conserved CXC domain sequences of CPP proteins from *B. napus* and its diploid progenitors.

Supplementary Figure 6 | Conserved motif characterization of CPP proteins in clade I (A) and clade II **(B)** from *B. napus* and diploid ancestors. The red circle indicates orthologs with potential direct evolutionary relationship.

REFERENCES

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J., Bierne, N., et al. (2013). Hybridization and speciation. *J. Evol. Biol.* 26, 229–246. doi: 10.1111/j.1420-9101.2012.02599.x

Adams, K. L., and Wendel, J. F. (2005). Polyploidy and genome evolution in plants. *Curr. Opin. Plant Biol.* 8, 135–141. doi: 10.1016/j.pbi.2005.01.001

Andersen, S. U., Algreen-Petersen, R. G., Hoedl, M., Jurkiewicz, A., Cvitanich, C., Braunschweig, U., et al. (2007). The conserved cysteine-rich domain of a tesmin/TSO1-like protein binds zinc *in vitro* and TSO1 is required for both male and female fertility in *Arabidopsis thaliana*. *J. Exp. Bot.* 58, 3657–3670. doi: 10.1093/jxb/erm215

Ashraf, M., and McNeilly, T. (2004). Salinity tolerance in *Brassica* oilseeds. *Crit. Rev. Plant Sci.* 23, 157–174. doi: 10.1080/07352680490433286

Cannon, S. B., Mitra, A., Baumgarten, A., Young, N. D., and May, G. (2004). The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. *BMC Plant Biol.* 4:10. doi: 10.1186/1471-2229-4-10

Cao, Y., Luo, Q., Tian, Y., and Meng, F. (2017). Physiological and proteomic analyses of the drought stress response in *Amygdalus Mira* (Koehne) Yü et Lu roots. *BMC Plant Biol.* 17:53. doi: 10.1186/s12870-017-1000-z

Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A., Tang, H., Wang, X., et al. (2014). Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science* 345, 950–953. doi: 10.1126/science.1253435

Chaves, M. M., Flexas, J., and Pinheiro, C. (2009). Photosynthesis under drought and salt stress: regulation mechanisms from whole plant to cell. *Ann. Bot.* 103, 551–560. doi: 10.1093/aob/mcn125

Chen, Z. J. (2007). Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. *Annu. Rev. Plant Biol.* 58, 377–406. doi: 10.1146/annurev.arplant.58.032806.103835

Cheng, F., Liu, S., Wu, J., Fang, L., Sun, S., Liu, B., et al. (2011). BRAD, the genetics and genomics database for *Brassica* plants. *BMC Plant Biol.* 11:136. doi: 10.1186/1471-2229-11-136

Cheng, F., Wu, J., Fang, L., Sun, S., Liu, B., Lin, K., et al. (2012). Biased gene fractionation and dominant gene expression among the subgenomes of *Brassica rapa*. *PLoS One* 7:e36442. doi: 10.1371/journal.pone.0036442

Cvitanich, C., Pallisgaard, N., Nielsen, K. A., Hansen, A. C., Larsen, K., Pihakaski-Maunsbach, K., et al. (2000). CPP1, a DNA-binding protein involved in the expression of a soybean *leghemoglobin c3* gene. *Proc. Natl. Acad. Sci. U.S.A.* 97, 8163–8168. doi: 10.1073/pnas.090468497

Hauser, B. A., He, J. Q., Park, S. O., and Gasser, C. S. (2000). TSO1 is a novel protein that modulates cytokinesis and cell expansion in Arabidopsis. *Development* 127, 2219–2226. doi: 10.1242/dev.127.10.2219

Hauser, B. A., Villanueva, J. M., and Gasser, C. S. (1998). Arabidopsis TSO1 regulates directional processes in cells during floral organogenesis. *Genetics* 150, 411–423. doi: 10.1093/genetics/150.1.411

Jackson, S., and Chen, Z. J. (2010). Genomic and expression plasticity of polyploidy. *Curr. Opin. Plant Biol.* 13, 153–159. doi: 10.1016/j.pbi.2009.11.004

Jiang, K., and Goertzen, L. R. (2011). Spliceosomal intron size expansion in domesticated grapevine (*Vitis vinifera*). *BMC Res. Notes* 4:52. doi: 10.1186/1756-0500-4-52

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., et al. (2009). Circos: an information aesthetic for comparative genomics. *Genome Res.* 19, 1639–1645. doi: 10.1101/gr.092759.109

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Lamond, A. I. (1999). Running rings around RNA. *Nature* 397, 655–656. doi: 10.1038/17697

Lespinet, O., Wolf, Y. I., Koonin, E. V., and Aravind, L. (2002). The role of lineage-specific gene family expansion in the evolution of eukaryotes. *Genome Res.* 12, 1048–1059. doi: 10.1101/gr.174302

Li, M., Wang, R., Liang, Z., Wu, X., and Wang, J. (2019a). Genome-wide identification and analysis of the *EIN3/EIL* gene family in allotetraploid *Brassica napus* reveal its potential advantages during polyploidization. *BMC Plant Biol.* 19:110. doi: 10.1186/s12870-019-1716-z

Li, M., Wang, R., Liu, Z., Wu, X., and Wang, J. (2019b). Genome-wide identification and analysis of the *WUSCHEL*-related homeobox (*WOX*) gene family in allotetraploid *Brassica napus* reveals changes in *WOX* genes during polyploidization. *BMC Genomics* 20:317. doi: 10.1186/s12864-019-5684-3

Li, M., Wang, R., Wu, X., and Wang, J. (2020). Homoeolog expression bias and expression level dominance (ELD) in four tissues of natural allotetraploid *Brassica napus*. *BMC Genomics* 21:330. doi: 10.1186/s12864-020-6747-1

Liang, Y., Wan, N., Cheng, Z., Mo, Y., Liu, B., Liu, H., et al. (2017). Whole-genome identification and expression pattern of the vicinal oxygen chelate family in rapeseed (*Brassica napus* L.). *Front. Plant Sci.* 8:745. doi: 10.3389/fpls.2017.00745

Librado, P., and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451–1452. doi: 10.1093/bioinformatics/btp187

Lv, W. T., Lin, B., Zhang, M., and Hua, X. J. (2011). Proline accumulation is inhibitory to *Arabidopsis* seedlings during heat stress. *Plant Physiol.* 156, 1921–1933. doi: 10.1104/pp.111.175810

Lynch, M. (2006). The origins of eukaryotic gene structure. *Mol. Biol. Evol.* 23, 450–468. doi: 10.1093/molbev/msj050

Madlung, A. (2013). Polyploidy and its effect on evolutionary success: old questions revisited with new tools. *Heredity* 110, 99–104. doi: 10.1038/hdy.2012.79

Maniatis, T., and Reed, R. (2002). An extensive network of coupling among gene expression machines. *Nature* 416, 499–506. doi: 10.1038/416499a

Mukherjee, D., Saha, D., Acharya, D., Mukherjee, A., Chakraborty, S., and Ghosh, T. C. (2018). The role of introns in the conservation of the metabolic genes of *Arabidopsis thaliana*. *Genomics* 110, 310–317. doi: 10.1016/j.ygeno.2017.12.003

Parankusam, S., Bhatnagar-Mathur, P., and Sharma, K. K. (2017). Heat responsive proteome changes reveal molecular mechanisms underlying heat tolerance in chickpea. *Environ. Exp. Bot.* 141, 132–144. doi: 10.1016/j.envexpbot.2017.07.007

Pikaard, C. S. (2001). Genomic change and gene silencing in polyploids. *Trends Genet.* 17, 675–677. doi: 10.1016/s0168-9525(01)02545-8

Prakash, S., Bhat, S., Quiros, C., Kirti, P., and Chopra, V. (2009). *Brassica* and its close allies: cytogenetics and evolution. *Plant Breed Rev.* 31:21. doi: 10.1002/9780470593783.ch2

Qiao, X., Li, Q., Yin, H., Qi, K., Li, L., Wang, R., et al. (2019). Gene duplication and evolution in recurring polyploidization-diploidization cycles in plants. *Genome Biol.* 20:38. doi: 10.1186/s13059-019-1650-2

Quinet, M., Vromman, D., Clippe, A., Bertin, P., Lequeux, H., Dufey, I., et al. (2012). Combined transcriptomic and physiological approaches reveal strong differences between short- and long-term response of rice (*Oryza sativa*) to iron toxicity. *Plant Cell Environ.* 35, 1837–1859. doi: 10.1111/j.1365-3040.2012.02521.x

Ren, X. Y., Vorst, O., Fiers, M. W., Stiekema, W. J., and Nap, J. P. (2006). In plants, highly expressed genes are the least compact. *Trends Genet.* 22, 528–532. doi: 10.1016/j.tig.2006.08.008

Riechmann, J. L., Heard, J., Martin, G., Reuber, L., Jiang, C., Keddie, J., et al. (2000). Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. *Science* 290, 2105–2110. doi: 10.1126/science.290.5499.2105

Samans, B., Chalhoub, B., and Snowdon, R. J. (2017). Surviving a genome collision: genomic signatures of allopolyploidization in the recent crop species *Brassica napus*. *Plant Genome* 10, 1–15. doi: 10.3835/plantgenome2017.02.0013

Segraves, K. A. (2017). The effects of genome duplications in a community context. *New Phytol.* 215, 57–69. doi: 10.1111/nph.14564

Soltis, D. E., Visger, C. J., Marchant, D. B., and Soltis, P. S. (2016). Polyploidy: pitfalls and paths to a paradigm. *Am. J. Bot.* 103, 1146–1166. doi: 10.3732/ajb.1500501

Song, J. M., Guan, Z., Hu, J., Guo, C., Yang, Z., Wang, S., et al. (2020). Eight high-quality genomes reveal pan-genome architecture and ecotype differentiation of *Brassica napus*. *Nat. Plants* 6, 34–45. doi: 10.1038/s41477-019-0577-7

Song, J. Y., Leung, T., Ehler, L. K., Wang, C., and Liu, Z. (2000). Regulation of meristem organization and cell division by TSO1, an Arabidopsis gene with cysteine-rich repeats. *Development* 127, 2207–2217. doi: 10.1242/dev.127.10.2207

Song, X. Y., Zhang, Y. Y., Wu, F. C., and Zhang, L. (2016). Genome-wide analysis of the maize (*Zea may* L.) *CPP-like* gene family and expression profiling under abiotic stress. *Genet. Mol. Res.* 15:gmr.15038023. doi: 10.4238/gmr.15038023

Tu, Y., Jiang, A., Gan, L., Hossain, M., Zhang, J., Peng, B., et al. (2014). Genome duplication improves rice root resistance to salt stress. *Rice* 7:15. doi: 10.1186/s12284-014-0015-4

Wang, J., Jin, Z., Zhou, M., Yu, Y., and Liang, M. (2020). Characterization of NF-Y transcription factor families in industrial rapeseed (*Brassica napus* L.) and identification of BnNF-YA3, which functions in the abiotic stress response. *Ind. Crops Prod.* 148:112253. doi: 10.1016/j.indcrop.2020.112253

Wang, R., Li, M., Wu, X., and Wang, J. (2019). The gene structure and expression level changes of the *GH3* gene family in *Brassica napus* relative to its diploid ancestors. *Genes* 10:58. doi: 10.3390/genes10010058

Wang, W., Sijacic, P., Xu, P., Lian, H., and Liu, Z. (2018). Arabidopsis TSO1 and MYB3R1 form a regulatory module to coordinate cell proliferation with differentiation in shoot and root. *Proc. Natl. Acad. Sci. U.S.A.* 115, e3045–e3054. doi: 10.1073/pnas.1715903115

Wu, G., Lin, L., Jiao, Q., and Li, S. (2019). Tetraploid exhibits more tolerant to salinity than diploid in sugar beet (*Beta vulgaris* L.). *Acta Physiol. Plant.* 41:52. doi: 10.1007/s11738-019-2844-7

Wu, Z. H., Zeng, F. H., Ma, S. J., Lin, L. J., Xie, Y. J., and Lu, X. Y. (2004). Effects of exogenous ABA on soluble protein in *Cynodon dactylon* under PEG stress. *Acta Pratac. Sin.* 13, 75–78. doi: 10.1088/1009-0630/6/5/011

Xu, C., Li, Z., and Wang, J. (2020). Linking heat and adaptive responses across temporal proteo-transcriptome and physiological traits of *Solidago canadensis*. *Environ. Exp. Bot.* 175:104035. doi: 10.1016/j.envexpbot.2020.104035

Yang, Z., Gu, S., Wang, X., Li, W., Tang, Z., and Xu, C. (2008). Molecular evolution of the *CPP-like* gene family in plants: insights from comparative genomics of *Arabidopsis* and rice. *J. Mol. Evol.* 67, 266–277. doi: 10.1007/s00239-008-9143-z

Ying, S. Y., and Lin, S. L. (2005). Intronic microRNAs. *Biochem. Biophys. Res. Commun.* 326, 515–520. doi: 10.1016/j.bbrc.2004.10.215

Yoo, M. J., and Wendel, J. F. (2014). Comparative evolutionary and developmental dynamics of the cotton (*Gossypium hirsutum*) fiber transcriptome. *PLoS Genet.* 10:e1004073. doi: 10.1371/journal.pgen.1004073

Zhang, L., Zhao, H. K., Wang, Y. M., Yuan, C. P., Zhang, Y. Y., Li, H. Y., et al. (2015). Genome-wide identification and expression analysis of the *CPP-like* gene family in soybean. *Genet. Mol. Res.* 14, 1260–1268. doi: 10.4238/2015.February.13.4

Zhou, T., Wang, Y., Chen, J. Q., Araki, H., Jing, Z., Jiang, K., et al. (2004). Genome-wide identification of NBS genes in japonica rice reveals significant expansion of divergent non-TIR NBS-LRR genes. *Mol. Genet. Genomics* 271, 402–415. doi: 10.1007/s00438-004-0990-z

Zhou, Y., Hu, L., Ye, S., Jiang, L., and Liu, S. (2018). Genome-wide identification and characterization of cysteine-rich polycomb-like protein (CPP) family genes in cucumber (*Cucumis sativus*) and their roles in stress responses. *Biologia* 73, 425–435. doi: 10.2478/s11756-018-0049-y

Zhu, J. K. (2016). Abiotic stress signaling and responses in plants. *Cell* 167, 313–324. doi: 10.1016/j.cell.2016.08.029

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 25 April 2022
doi: 10.3389/fpls.2022.844545

Comprehensive Profiling of Tubby-Like Proteins in Soybean and Roles of the *GmTLP8* Gene in Abiotic Stress Responses

Hong-Ru Xu[1,2†], Ying Liu[2†], Tai-Fei Yu[2†], Ze-Hao Hou[2], Jia-Cheng Zheng[3], Jun Chen[2], Yong-Bin Zhou[2], Ming Chen[2], Jin-Dong Fu[2], You-Zhi Ma[2], Wen-Liang Wei[1] and Zhao-Shi Xu[2]**

[1] College of Agriculture, Yangtze University/Hubei Collaborative Innovation Center for Grain Industry/Engineering Research Center of Ecology and Agricultural Use of Wetland, Ministry of Education, Jingzhou, China, [2] Institute of Crop Science, Chinese Academy of Agricultural Sciences (CAAS)/National Key Facility for Crop Gene Resources and Genetic Improvement, Key Laboratory of Biology and Genetic Improvement of Triticeae Crops, Ministry of Agriculture, Beijing, China, [3] College of Agronomy, Anhui Science and Technology University, Fengyang, China

Tubby-like proteins (TLPs) are transcription factors that are widely present in eukaryotes and generally participate in growth and developmental processes. Using genome databases, a total of 22 putative *TLP* genes were identified in the soybean genome, and unevenly distributed across 13 chromosomes. Phylogenetic analysis demonstrated that the predicted GmTLP proteins were divided into five groups (I-V). Gene structure, protein motifs, and conserved domains were analyzed to identify differences and common features among the GmTLPs. A three-dimensional protein model was built to show the typical structure of TLPs. Analysis of publicly available gene expression data showed that *GmTLP* genes were differentially expressed in response to abiotic stresses. Based on those data, *GmTLP8* was selected to further explore the role of TLPs in soybean drought and salt stress responses. *GmTLP8* overexpressors had improved tolerance to drought and salt stresses, whereas the opposite was true of *GmTLP8*-RNAi lines. 3,3-diaminobenzidine and nitro blue tetrazolium staining and physiological indexes also showed that overexpression of *GmTLP8* enhanced the tolerance of soybean to drought and salt stresses; in addition, downstream stress-responsive genes were upregulated in response to drought and salt stresses. This study provides new insights into the function of GmTLPs in response to abiotic stresses.

Keywords: tubby-like protein, genome-wide analysis, abiotic stress, responsive mechanism, soybean

Abbreviations: TLP, tubby-like protein; ABA, abscisic acid; GFP, green fluorescent protein; MDA, malondialdehyde; PRO, proline; PEG, polyethylene glycol; DAB, 3, 3-diaminobenzidine; NBT, nitro blue tetrazolium; qRT-PCR, quantitative real-time PCR; ABRE, ABA-responsive element; MYC, drought and salt-responsive element; NCED, 9-*cis*-epoxycarotenoid dioxygenase; SCF-complex, Skp1-Cullin1-F-box complex; SKP1-like proteins, S-phase kinase-associated protein 1 like proteins; NHL, NDR1/HIN1-like gene.

OPEN ACCESS

Edited by:
Prasanta Kumar Subudhi, Louisiana State University, United States

Reviewed by:
Ajit Ghosh, Shahjalal University of Science and Technology, Bangladesh
Youxiong Que, Fujian Agriculture and Forestry University, China

***Correspondence:**
Wen-Liang Wei
whwenliang@163.com
Zhao-Shi Xu
xuzhaoshi@caas.cn

[†] *These authors have contributed equally to this work*

Specialty section:
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

Received: *28 December 2021*
Accepted: *15 March 2022*
Published: *25 April 2022*

Citation:
Xu H-R, Liu Y, Yu T-F, Hou Z-H, Zheng J-C, Chen J, Zhou Y-B, Chen M, Fu J-D, Ma Y-Z, Wei W-L and Xu Z-S (2022) Comprehensive Profiling of Tubby-Like Proteins in Soybean and Roles of the GmTLP8 Gene in Abiotic Stress Responses. Front. Plant Sci. 13:844545. doi: 10.3389/fpls.2022.844545

INTRODUCTION

The Tubby-like proteins (TLPs) are a class of eukaryotic transcription factors that were originally identified in obese mice (Kleyn et al., 1996; Liu, 2008). A typical TLP has a highly conserved tubular domain composed of 270 amino acids at the C-terminal, forming a β-barrel with a central hydrophobic α-helix and 12 antiparallel strands. It binds specific phosphatidylinositol 4,5-diphosphates to properly connect to the plasma membrane (Santagata et al., 2001; Mukhopadhyay and Jackson, 2011). TLPs have been widely studied in animals. For example, TULP3 is defined as a universal adapter for the transport of integral membrane proteins in the ciliary membrane to the cilia (Badgandi et al., 2017). Mutations of *TLPs* in humans lead to delayed obesity (Coleman and Eicher, 1990; Kleyn et al., 1996; Noben-Trauth et al., 1996; Kapeller et al., 1999; Borman et al., 2014), and mice with *TLP* mutations develop retinal degeneration, neurosensory hearing loss, and insulin resistance (Stretton et al., 2009).

In addition to the typical C-terminal tubular domain, plant TLPs have a conserved N-terminal F-box domain, which is not present in mammalian TLPs (Gagne et al., 2002; Lai et al., 2004). Previous studies have shown that TLPs have a variety of functions in plants, including growth, development, and disease resistance. TLPs may function in pollen grains, consistent with the fact that *AtTLP6*, *AtTLP7*, and *AtTLP2* are mainly expressed in pollen grains of *Arabidopsis* (Bao et al., 2014). *AtTLP2* is involved in the biosynthesis of homogalacturonic acid in *Arabidopsis* seed coat mucilage (Wang et al., 2019). Fourteen *TLP* genes have been identified in rice (*OsTLPs*), and differential expression analysis confirmed that the members of this group play important roles in processes related to physiological development (Liu, 2008). Expression of each *OsTLP* was induced by infection with *Xanthomonas oryzae* pv. *oryzae*, indicating that the *OsTLP* family is involved in host–pathogen interaction (Kou et al., 2009). *OsTLP2* can bind to the *OsWRKY13* promoter to regulate rice resistance to fungal plague and bacteria (Cai et al., 2008). Tomato *SlTLP1* and *SlTLP2* may have important roles in ethylene-dependent fruit ripening (Zhang et al., 2020); *SlTLFP8* regulates cell size and stomatal density through endoreduplication, reduces water loss, and enhances water use efficiency (Li et al., 2020).

Previous reports have demonstrated the responses of TLPs to various abiotic stresses. During seed germination and seedling growth, *AtTLP3* responds to abiotic stresses such as abscisic acid (ABA), NaCl, and mannitol (North et al., 1997; Bao et al., 2014). *AtTLP9* regulates ABA sensitivity during seed germination and early seedling development (Lai et al., 2004; Chen et al., 2020). Overexpression of *CaTLP1* in chickpeas can enhance tolerance to drought, salt stresses, and ABA (Bhushan et al., 2007). In apples, polyethylene glycol (PEG) treatment up-regulates expression of *MdTLP1-MdTLP5* and *MdTLP9* (Xu et al., 2016). Overexpression of apple *MdTLP7* enhances the tolerance of *Arabidopsis* to osmotic, salt, and temperature stresses (Xu et al., 2019). *ZmTLP2* and *ZmTLP11* are significantly up-regulated in maize under drought stress (Chen et al., 2016). *CsTLP8* plays a negative regulatory role in osmotic stress in cucumber, and its effects may be related to ABA (Li S. et al., 2021). Transcriptome analysis has shown that cotton *GhTULPs* are involved in abiotic stresses and tissue development. Overexpression of *GhTULP34* was shown to decrease the germination rate of *Arabidopsis* seeds under salt stress, inhibit root development under osmotic stress, and lead to the closure of plant stomata (Li Z. et al., 2021). In summary, TLPs play key roles in plant growth and development and in responses to biotic and abiotic stresses.

Soybean (*Glycine max*) is one of the most economically important crops in the world, often used as a source of food for humans and livestock because of its rich oil and protein (Papiernik et al., 2005). As global climate change occurs, the adaptability of soybean to its living environment is gradually reduced, causing a demand for stress-tolerant soybean varieties. Further studies are needed to improve soybean tolerance to extreme environments, including various abiotic stresses such as drought and salt (Le et al., 2012). There is little published information about TLPs and their relationship with abiotic stress mechanisms in soybean. In this study, 22 *TLP* genes were identified in the soybean genome, and bioinformatic analyses were conducted to determine their chromosomal locations, gene structures, protein domains, conserved motifs, three-dimensional structures, and *cis*-acting elements. Based on RNA-Seq and quantitative Real-Time Polymerase Chain Reaction (qRT-PCR), we further investigated the role of *GmTLP8* in drought and salt stress responses in soybean, and found that overexpression of *GmTLP8* enhanced tolerance to drought and salt stresses in soybean. These findings provide insights into the function of *GmTLP8*, specifically in abiotic stress responses, and into the importance of GmTLPs more broadly in plant abiotic stress responses.

MATERIALS AND METHODS

Identification of Tubby-Like Proteins in Soybean

Soybean genome, protein, complementary DNA (cDNA) sequences, and gene annotation files were obtained from NCBI[1] and the Phytozome database[2] (Finn et al., 2011; Fernandez-Pozo et al., 2015). The Hidden Markov Model (HMM) profile corresponding to the TLP Tub domain (PF01167) from the Pfam protein family database[3] was used to identify potential TLPs in the soybean genome (*G. max* Wm82. a2.v1) using HMMER v3 (Eddy, 1998; Mistry et al., 2021). Finally, the presence of the Tub domain in each TLP protein sequence was confirmed with the SMART tool[4] (Letunic et al., 2012) and Pfam database. The molecular weight and isoelectric point data for GmTLPs were calculated by ExPASY[5] (Artimo et al., 2012). Subcellular localization was predicted with WoLF PSORT[6].

[1]http://www.ncbi.nlm.nih.gov/
[2]https://phytozome.jgi.doe.gov/pz/portal.html
[3]http://pfam.xfam.org/
[4]http://smart.embl-heidelberg.de/
[5]http://web.expasy.org/
[6]https://wolfpsort.hgc.jp/

Phylogenetic Tree Construction and Multiple Sequence Alignment

The full-length amino acid sequences of TLP members in rice (OsTLPs), *Arabidopsis* (AtTLPs), cotton (GhTLPs), maize (ZmTLPs), apple (MdTLPs), poplar (PtTLPs), wheat (TaTLPs), tomato (SlTLPs), and the newly identified GmTLPs were obtained from NCBI and Phytozome, respectively, and aligned with default parameters using ClustalW (Chenna et al., 2003). An unrooted phylogenetic tree was constructed using the neighbor-joining (NJ) method in MEGAX (version 10.1.8) (Tang et al., 2021) with the following parameters: pairwise deletion; Poisson model; 1000 bootstrap replications.

The amino acid sequence of 22 TLP proteins of soybean aligned using DNAMAN (version 6.0.3).

Chromosomal Localization, Structural Characterization, and Conserved Motif Analysis

Chromosomal location data for *GmTLP*s were obtained from the Phytozome database. Intron insertion sites were identified by comparing the coding sequence of each *TLP* gene with the corresponding full-length sequence using the Gene Structure Display Server (GSDS) 2.0[7] (Hu et al., 2015). The conserved domain of the identified soybean GmTLP protein sequences were determined using MEME[8] with the maximum number of motifs set to 10 (Bailey et al., 2009).

Protein Domain Analysis and Homology Modeling

Protein sequences of the 22 GmTLPs were submitted to the SMART website[9] to obtain data related to conserved protein domains, and GSDS 2.0 was used for visual analysis. Three-dimensional models of the Tub domain were built with SWISS-MODEL[10] (Dong et al., 2019). Tub domain models were obtained for 20 GmTLPs with the protein sequence identity set to $\geq$30%.

Expression Patterns of *TLPs* in Soybean

Soybean gene expression files were downloaded from the Soybase website[11] to analyze the expression patterns of 22 GmTLPs members in different tissues at different developmental stages under normal conditions, including young-leaf, flower, pod, pod shell, seed, root, and nodule. In the database file provided by Soybase website, only 18 members' tissue differential expression information were found for further analysis. Transcriptome data for GmTLPs members under various abiotic stresses from our previous studies (Wang et al., 2020). 22 GmTLPs members were used for searching in transcriptome data, and their expression levels under normal condition, ABA treatment, drought and salt stresses were analyzed. Finally, the relevant information of 21 members was obtained. TBtools (version 1.075) (Chen et al., 2020) was used for visualization and cluster analysis of *GmTLP* expression patterns.

Analysis of *Cis*-Acting Elements in *GmTLP* Gene Promoters

GmTLP sequences obtained from the Phytozome database were extracted in batches with TBtools, and the 2000 bp upstream promoter sequences of the 22 *GmTLP* genes were obtained and submitted to the online program PlantCARE[12] to identify *cis*-acting elements. GSDS 2.0 was used for data visualization.

Plant Materials and Growth Conditions

The soybean variety Zhonghuang39 was used for analysis of *GmTLP* gene expression in this study. Soybeans were grown in 1: 1 vermiculite: humus in a greenhouse with a 16/8 h light/dark cycle, day/night temperatures of 28/20°C, and a relative humidity of 70%. At 14 days, the seedlings at the four-leaf stage were stressed with drought or salt. Referring to previous research methods (Wang et al., 2021), soybean seedlings were removed from soil. For drought stress, the seedlings were placed on filter paper; for salt stress, the seedlings were immersed in 200 mM NaCl solution. The sampling time of drought or salt stress was 0, 0.5, 1, 2, 4, 8, 12, and 24 h. There were three biological replicates per treatment. After treatment, the leaves were frozen in liquid nitrogen and stored at –80°C before further analysis (Xu et al., 2008). These samples were used for qRT-PCR analysis of subsequent *GmTLPs* members.

RNA Extraction and Quantitative Real-Time Polymerase Chain Reaction

Total RNA was extracted from soybean leaves using a plant RNA extraction kit following the manufacturer's instructions (TIANGEN, Beijing, China). cDNA was synthesized using the PrimeScript^{TM} RT Reagent Kit (TaKaRa, Shiga, Japan) following the manufacturer's protocol. Primers (**Supplementary Table 2**) were designed using Primer Premier 5.0. The soybean *Actin* gene (U60506) was used as the internal control for quantitative real-time PCR (qRT-PCR). There were three technical replicates for each sample. Differential expression was determined from the relative gene expression data using the $2^{-\Delta\ \Delta\ CT}$ method (Le et al., 2011).

Subcellular Localization of GmTLP8

We constructed an expression vector labeled with green fluorescent protein (hGFP) for subcellular localization analysis. The full-length cDNA sequence of *GmTLP8* was fused to the N-terminal hGFP protein driven by the CaMV35S promoter (Xu et al., 2007). The 35S:GFP vector was used as a control. A PEG4000-mediated method was used to transform the *GmTLP8*-GFP recombinant plasmid into *Arabidopsis* protoplasts (He et al., 2016). After incubation for 18–20 h in the dark at 22°C, the nucleus of *GmTLP8*-GFP protoplasts were specifically stained with 4', 6-diamidino-2-phenylindole (DAPI). The fluorescence signal was observed using a confocal laser scanning microscope

[7] http://gsds.cbi.pku.edu.cn

[8] https://meme-suite.org/meme/

[9] http://smart.embl-heidelberg.de/smart/batch.pl

[10] https://swissmodel.expasy.org

[11] http://www.SoyBase.org

[12] http://bioinformatics.psb.ugent.be/webtools/plantcare/html/

(Zeiss LSM 700, Oberkochen, Germany). There were three technical replicates for each group.

Agrobacterium rhizogenes-Mediated Transformation of Soybean Hairy Roots

The transformation was conducted to produce soybean hairy roots that were characterized by overexpression of *GmTLP8* (*GmTLP8*-OE), RNA interference of *GmTLP8* (*GmTLP8*-RNAi), or with the empty pCAMBIA3301 vector (EV-Control) (Chen et al., 2021). The CDS of *GmTLP8* was amplified without stop codon using gene-specific primer pairs, under the control of the CaMV35S promoter, *GmTLP8* cDNA was ligated into the plant transformation vector pCAMBIA3301 to generate *GmTLP8*-overexpressing (*GmTLP8*-OE) vector. In order to construct the RNAi vector, a 564 bp fragment including the first intron sequence and its reverse complement was synthesized (Biomed, Beijing, China) and inserted into pCAMBIA3301 to generate the pCAMBIA3301-*GmTLP8*-RNAi (*GmTLP8*-RNAi) vector. The recombinant construct and the empty pCAMBIA3301 (EV-Control) vector were transferred into *A. rhizogenes* strain K599, as previously described, then injected into soybean (*G. max cv.* Zhonghuang39) hypocotyl for *A. rhizogenes*-mediated transformation of soybean hairy roots (Wang et al., 2009; Du et al., 2018).

The injected plants were placed in a high-humidity greenhouse until hairy roots were generated at the infected site and had grown to ∼5 cm in length. After cutting off the original tap root 0.5 cm below the infected site, the seedlings were transplanted into fertilized soil and cultivated in a greenhouse at 25°C with a 16/8 h light/dark photoperiod for 7 days (Yu et al., 2021). The qRT-PCR analysis of *GmTLP8* expression in *GmTLP8*-OE, EV-control and *GmTLP8*-RNAi transgenic hairy root plants before processing (**Supplementary Figure 2G**). Each of the hairy root-related experiments was replicated at least three times independently. The primers of *GmTLP8*-3301-F and *GmTLP8*-3301-R were listed in **Supplementary Table 2**.

Drought and Salt Stress Assays of Soybean Hairy Root Composite Plants

Transgenic hairy root composite soybean plants were used in drought and salt stress assays after 7 days of normal growth. For drought treatment, soybean plants were grown for 7 days without watering; for NaCl treatment, soybean plants were treated with 150 mM NaCl for 3 days. Drought and salt treatment experiments were conducted a minimum of three times. Both the treated and untreated soybean hairy roots were washed with water prior to RNA isolation and physiological/biochemical experiments.

Measurements of Physiological Indexes

Several physiological parameters were measured in transgenic *GmTLP8*-OE, EV-Control, and *GmTLP8*-RNAi lines after the drought and NaCl treatments, namely levels of proline (Pro), malondialdehyde (MDA), hydrogen peroxide (H_2O_2), superoxide anion (O_2^-), and chlorophyll. Measurements were taken in soybean leaves using appropriate assay kits (Cominbio, Suzhou, China) following the manufacturer's instructions. All measurements were performed in three biological replicates.

Leaf Staining With 3,3-Diaminobenzidine and Nitro Blue Tetrazolium

Leaves from the three transgenic lines were stained with 3,3-diaminobenzidine and nitro blue tetrazolium after drought or salt stress treatment. The leaves were immersed in DAB solution or NBT staining solution (Solarbio, Beijing, China) for 18 or 14 h, respectively. Samples were then destained in a boiling solution of 3: 1 anhydrous ethanol: glycerol until the leaves were white (Du et al., 2018). Images were taken using a Canon 50D camera (Canon, Tokyo, Japan). There were three biological replicates for each plant line–treatment group combination.

RESULTS

Identification of Tubby-Like Proteins in Soybean Genome

Twenty-two GmTLP family members were identified in this study. The SMART and Pfam databases were used to confirm the presence of the conserved Tub domain in all of the putative TLP proteins. Twenty-two *GmTLP* genes were unevenly distributed across 13 chromosomes of soybean. According to their positions on chromosomes, we named them *GmTLP1* to *GmTLP22*. The details of TLPs in soybean, such as the coding sequence (CDS) length, amino acid length (aa), molecular weight (MW), isoelectric point (p*I*), and subcellular location are shown in **Table 1**.

Among the 22 GmTLPs, the protein length ranged from 183 (*GmTLP5*) to 454 amino acids (*GmTLP21*). The minimum protein MW was 20.8 kDa (*GmTLP5*), and the maximum was 50.6 kDa (*GmTLP21*). The p*I* ranged from 7.54 in *GmTLP10* to 9.65 in *GmTLP8*. Twelve of the proteins were predicted to be located in the nucleus, five in the cytosol, four in the mitochondria, and two in the chloroplast (**Table 1**), with *GmTLP1* predicted to be located in either the nucleus or cytosol.

Chromosome Distribution, Phylogenetic Analysis, and Multiple Sequence Alignment

A physical location map of the *GmTLPs* was drawn using physical location data from the soybean genome. The 22 *GmTLP* genes were distributed across 13 chromosomes, which were chromosome 1, 2, 7, 8, 10, 11, 12, 13, 14, 15, 16, 17, 20, respectively. There were three genes on chromosomes 2 and 13, two genes on chromosomes 7, 12, 15, 16, and 17, and only one gene each on chromosomes 1, 8, 10, 11, 14, and 20 (**Figure 1**).

To reveal the phylogenetic relationships between TLPs in different plant species, an unrooted phylogenetic tree was constructed by comparing the amino acid sequences for all of the known TLP members in several species, totaling 132 proteins. There were 11 from *Arabidopsis* (Lai et al., 2004), 15 from maize (Chen et al., 2016), 14 from rice (Liu, 2008), four from wheat (Hong et al., 2015), 11 from tomato (Zhang et al., 2020),

TABLE 1 | Basic information of TLPs in soybean.

Name	Gene ID	CDS(bp)	Chr	Protein(aa)	MW(Da)	p/	Subcellular localization
GmTLP1	Glyma.01G173700	1249	1	415	46341.13	9.37	Nucl/cyto
GmTLP2	Glyma.02G055300	1246	2	414	46043.62	9.18	Chlo
GmTLP3	Glyma.02G081800	1276	2	424	47548.48	9.46	Nucl
GmTLP4	Glyma.02G152700	1204	2	400	44763.46	9.47	Nucl
GmTLP5	Glyma.07G147700	553	7	183	20813.28	9.43	Cyto
GmTLP6	Glyma.07G251800	1081	7	359	40220.18	9.06	Cyto
GmTLP7	Glyma.08G183100	1285	8	427	48059.27	9.48	Nucl
GmTLP8	Glyma.10G224900	1294	10	430	48068.12	9.65	Nucl
GmTLP9	Glyma.11G069400	1246	11	414	46094.90	9.36	Cyto
GmTLP10	Glyma.12G115200	889	12	295	33619.95	7.54	Cyto
GmTLP11	Glyma.12G230000	1138	12	378	41678.40	9.37	Mito
GmTLP12	Glyma.13G214900	1171	13	389	43646.37	9.44	Mito
GmTLP13	Glyma.13G269600	1147	13	381	41987.79	9.32	Mito
GmTLP14	Glyma.13G371500	1072	13	356	40004.66	9.63	Nucl
GmTLP15	Glyma.14G073500	1273	14	423	47357.84	9.29	Nucl
GmTLP16	Glyma.15G049500	1285	15	427	47952.01	9.41	Nucl
GmTLP17	Glyma.15G098200	1159	15	385	43331.98	9.30	Nucl
GmTLP18	Glyma.16G138100	1246	16	414	46248.94	9.01	Chlo
GmTLP19	Glyma.16G167200	1276	16	424	47621.54	9.46	Nucl
GmTLP20	Glyma.17G022700	1108	17	392	44015.49	9.07	Nucl
GmTLP21	Glyma.17G251500	1366	17	454	50583.48	9.41	Mito
GmTLP22	Glyma.20G166900	1294	20	430	48203.31	9.61	Nucl

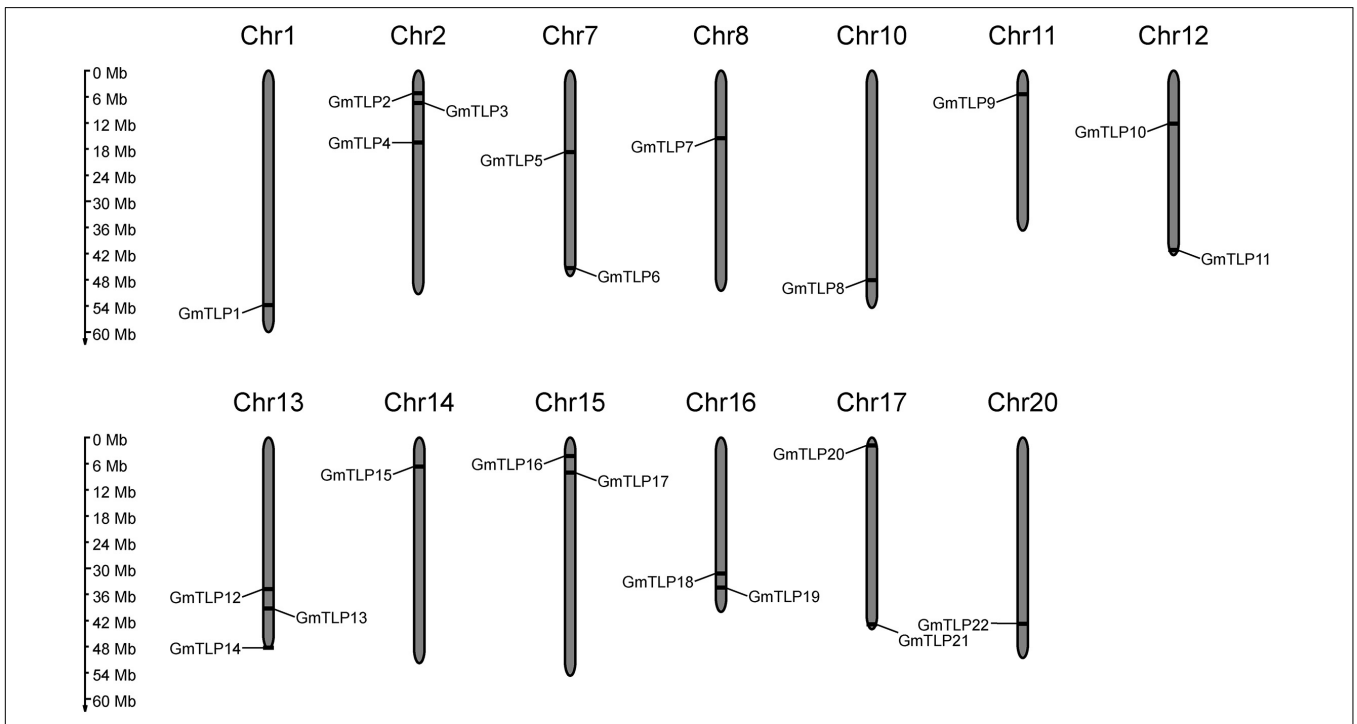


FIGURE 1 | Chromosomal distribution of the 22 putative *TLP* genes identified in soybean. The scale bar at left indicates the size of the chromosomes.

nine from apple (Xu et al., 2016), 11 from poplar (Yang et al., 2008), and 35 from cotton (Li Z. et al., 2021) in addition to the 22 putative TLPs identified in soybean. Phylogenetic tree was divided into five groups based on protein homology, and there were one, two, five, six, and eight GmTLP members in groups I, II, III, IV, and V, respectively (**Figure 2**).

The results of multiple sequence alignment showed that the positions of F-box domain and Tub domain in GmTLP protein

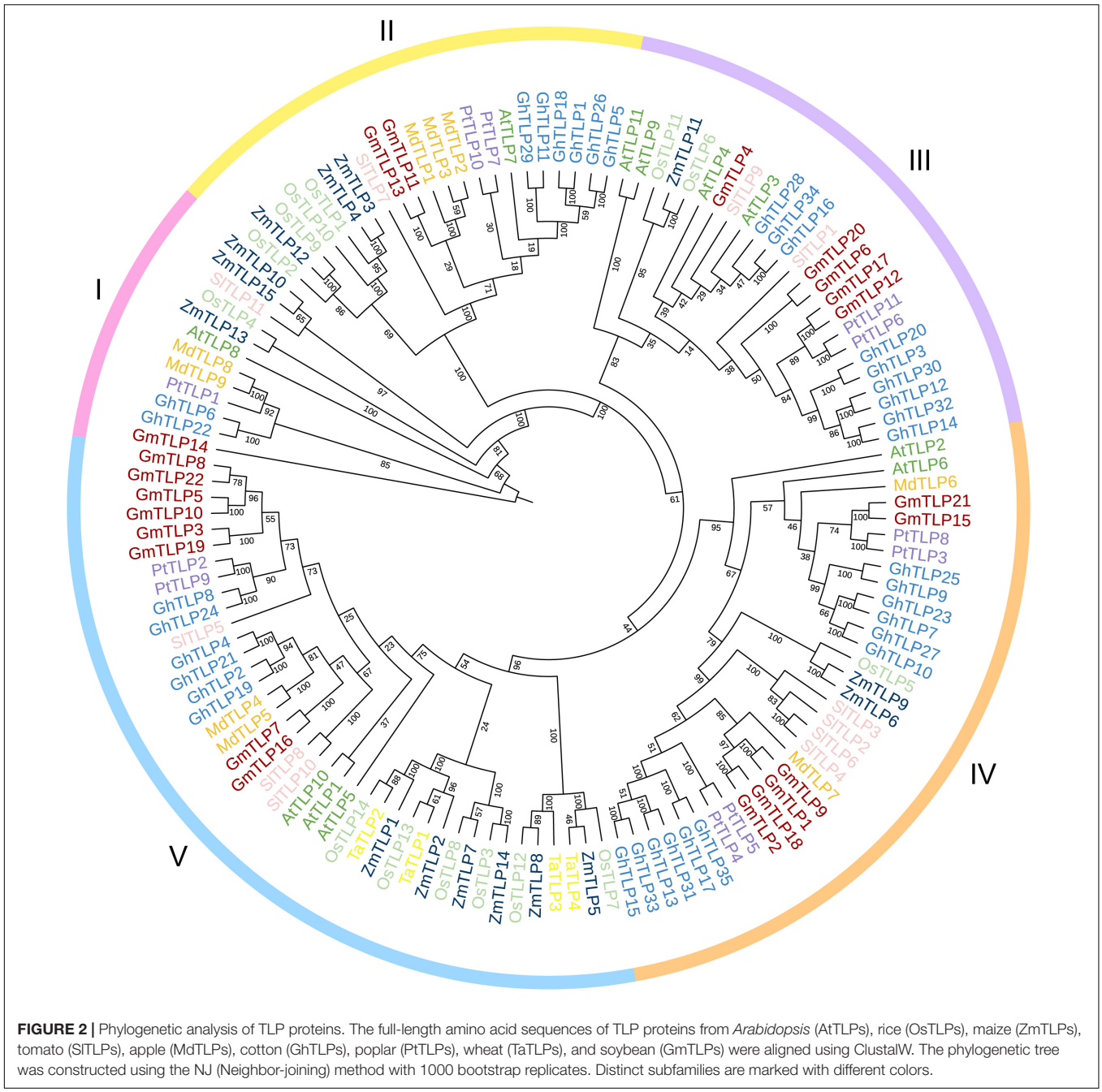


FIGURE 2 | Phylogenetic analysis of TLP proteins. The full-length amino acid sequences of TLP proteins from *Arabidopsis* (AtTLPs), rice (OsTLPs), maize (ZmTLPs), tomato (SlTLPs), apple (MdTLPs), cotton (GhTLPs), poplar (PtTLPs), wheat (TaTLPs), and soybean (GmTLPs) were aligned using ClustalW. The phylogenetic tree was constructed using the NJ (Neighbor-joining) method with 1000 bootstrap replicates. Distinct subfamilies are marked with different colors.

sequence were located in the front and rear segments of the sequence (**Figure 3**).

Gene Structure and Motifs in GmTLPs

We analyzed the gene structure of the 22 *GmTLPs* using GSDS 2.0 online to determine the intron and exon distribution of each (**Figure 4A**). A total of 15 genes contained three introns and four exons each, and the other seven genes contained four introns and five exons.

A total of 10 conserved motifs ($E \leq 0.01$) were analyzed using the MEME website to explore conservation and diversity of soybean TLPs. Among the 22 GmTLP family members, 19 contained all 10 motifs, with GmTLP4 containing two copies of motif 3 and GmTLP20 containing two copies of motif 5. GmTLP5, GmTLP10, and GmTLP14 contained six, five, and two motifs, respectively (**Figure 4C**). Consensus sequences for putative motifs are shown in **Supplementary Figure 1**.

Conserved Domain Analysis and Three-Dimensional Modeling

From the Pfam database, we found that two conserved domains in GmTLPs were Tub (PF01167) and F-box (PF00646). We

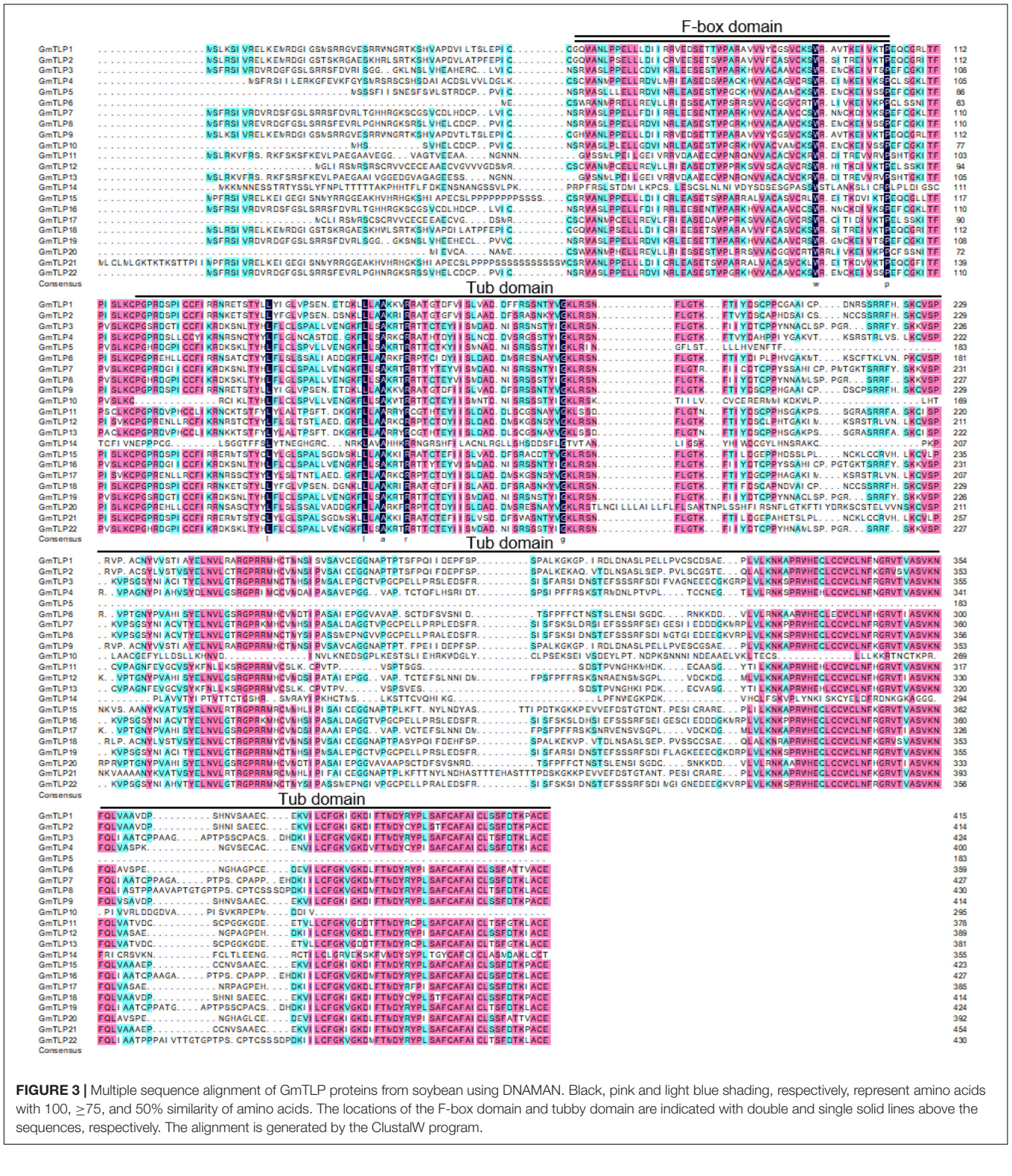


FIGURE 3 | Multiple sequence alignment of GmTLP proteins from soybean using DNAMAN. Black, pink and light blue shading, respectively, represent amino acids with 100, ≥75, and 50% similarity of amino acids. The locations of the F-box domain and tubby domain are indicated with double and single solid lines above the sequences, respectively. The alignment is generated by the ClustalW program.

then analyzed the conserved protein domains using both the SMART website (**Figure 4D**) and homology modeling in SWISS-MODEL (**Figure 5**).

The results of conserved domain analysis showed that 20 of the GmTLPs contained one F-box domain in the N-terminal region and one Tub domain in the C-terminal region. GmTLP4 had a Tub domain in the C-terminal region but no F-box domain in the N-terminal region, and GmTLP20 contained one F-box domain in the N-terminal region and two Tub domains in the C-terminal region. The analysis also revealed one ANK domain

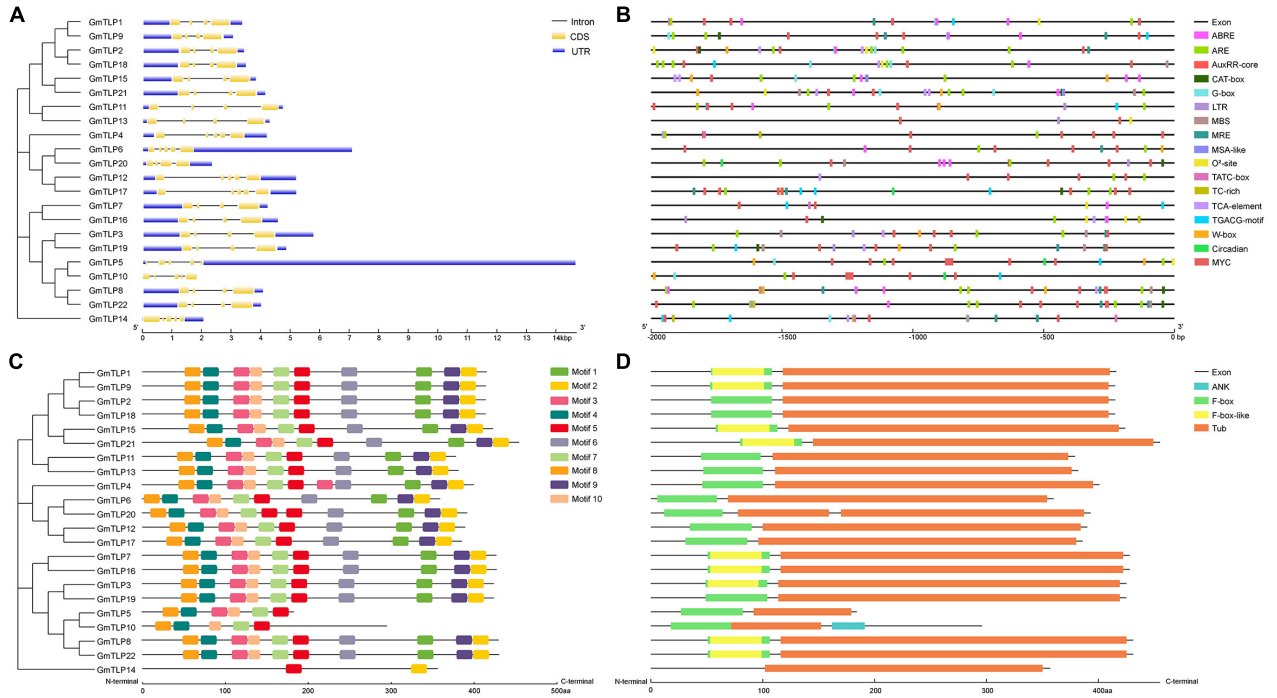


FIGURE 4 | Gene structures, *cis*-acting elements, motifs and conserved domains analysis of GmTLPs. **(A)** Phylogenetic relationships (left) and gene structures (right) of GmTLPs. The phylogenetic tree was constructed using MEGAX; the different classes of TLP proteins make up separate clades. The schematic diagram shows gene structure. Introns and exons are indicated by black lines and yellow boxes, respectively. The lengths of introns and exons in each gene are displayed proportionally. **(B)** Predicted *cis*-acting elements in the *GmTLP* promoters. Distinct color blocks indicate different *cis*-elements, including ABRE, ARE, LTR, TC-rich, MBS, MRE, G-box, TCA-element, TGACG-motif, W-box, circadian, MSA-like, O_2^--site, TATC-box, AuxRR-core, CAT-box, and MYC. The upstream distance from the translation start site can be estimated using the scale at the bottom. **(C)** Phylogenetic relationships (left) and putative motifs (right) of GmTLPs. Conserved motifs were identified using the MEME website and TBtools software. Ten putative motifs are indicated by colored boxes. **(D)** Conserved domain analyses of GmTLPs. The length of each protein can be estimated using the scale at the bottom.

in the C-terminal region of GmTLP10, the function of which was not clear (**Figure 4D**).

The homology modeling was a useful tool for the prediction of protein structure, and protein structural information was often more valuable than sequence data alone in determining protein function. We generated three-dimensional (3D) models of the Tub domains for 20 of the GmTLPs. These models showed that the Tub domain of each GmTLP was closed by a β-barrel with 12 anti-parallel strands and a central hydrophobic α-helix, which is a typical structure for a Tub domain (**Figure 5**).

Promoter Regions of *GmTLPs* Contain Various Stress Response Elements

Cis-acting regulatory elements play an important role in modulating gene expression. To understand transcriptional regulation of *GmTLPs*, we identified *cis*-acting elements within the promoter region of each *GmTLP* gene, defined as the 2000 bp region upstream of the start codon (**Figure 4B** and **Supplementary Table 1**). Results showed that most *cis*-acting elements in *GmTLP* promoters were involved in hormone or stress responses.

The main hormone-related *cis*-acting elements identified were ABA response element (ABRE), TATC-box (gibberellin), AuxRR core (auxin), TCA element (salicylic acid), and the TGACG motif (methyl jasmonate). Among the *GmTLP* promoters, 14 genes contained ABRE, four contained a TATC-box, six contained the AuxRR core, 11 contained a TCA element, and 12 contained a TGACG motif. This indicated that the *GmTLPs* may be involved in hormone-related responses.

The abiotic stress *cis*-acting elements identified were as follows: anaerobic inducing element (ARE), low temperature response element (LTR), a MYB binding site involved in drought induction (MBS), a drought and salt response element (MYC), and a defense and stress response element (TC-rich element). MYC, which was previously reported to be involved in drought and salt stress-induced responses, was revealed to be distributed in all of the *GmTLP* promoter sequences. In addition, 20 *GmTLP* promoters contained ARE, seven contained LTR, seven contained MBS, and six contained TC-rich elements. The presence of these *cis*-acting elements related to abiotic stresses indicated that *GmTLPs* are abiotic stress-responsive.

Among the 22 *GmTLPs*, 14 contained ABREs, indicating that these genes can be regulated by ABA. Drought and salt response element MYCs were present in all 22 *GmTLP* promoter sequences, strongly suggesting that *GmTLP* members are involved in the responses to those stressors (Zuo et al., 2020). In total, *cis*-acting element analysis indicated that most members

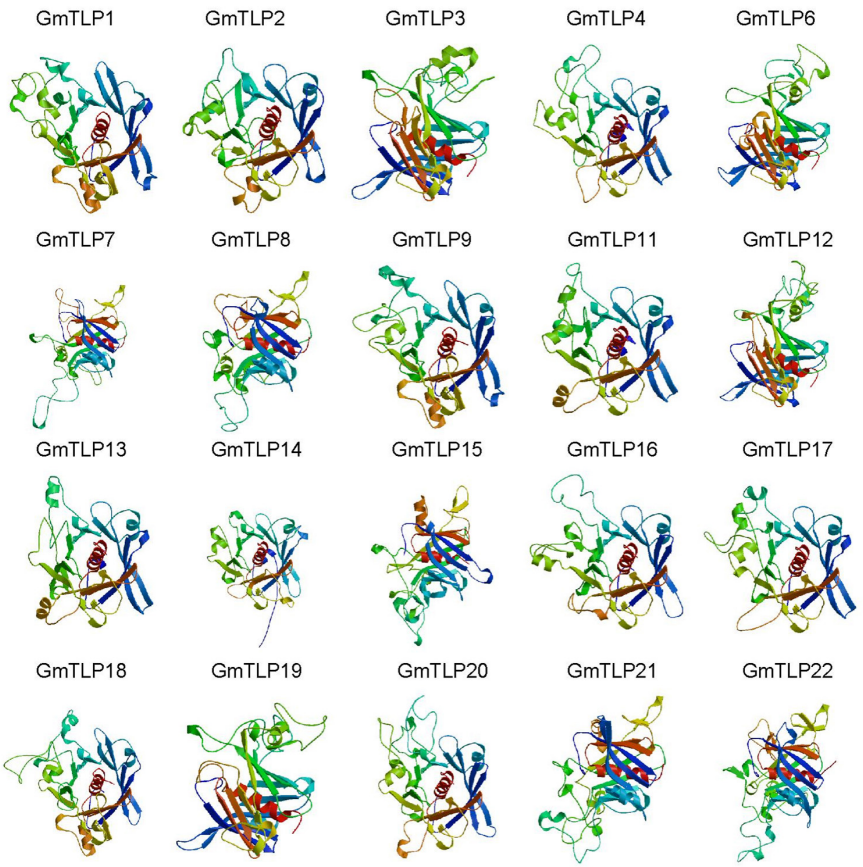


FIGURE 5 | Homology modeling of the 3D structure of GmTLP Tub domains. The α-helices are shown in red, and β-barrels are shown in different colors surround the α-helices. GmTLP, soybean tubby-like protein.

of the *GmTLP* family may be regulated by ABA in response to drought and salt stresses.

Tissue-Specific Expression Patterns of *TLPs* in Soybean

To understand expression patterns of *TLPs* during the growth and development of soybean and throughout different plant tissues, publicly available transcriptome sequencing data from the SoyBase database were analyzed. For 18 *GmTLP* members, we analyzed gene expression levels in different plant tissues including young-leaf, flower, pod, pod shell, seed, root, and nodule. The results showed that *GmTLP7* was extremely high expressed in all seven tissues. *GmTLP6*, *10*, *22*, *11*, and *21* were expressed at extremely low levels or not expressed in seven tissues. *GmTLP2*, *4*, *9*, *12*, and *14* were expressed in some tissues, but not in others. *GmTLP13*, *17*, *3*, *5*, *20*, *15*, and *16* were expressed in all seven tissues, with extremely high expression in some tissues and extremely low expression in others. Phylogenetic analysis divided the 18 *GmTLP* members into different groups, and members within each group shared similarities at the expression level in all seven tissues (**Figure 6**). The results showed a great deal of spatiotemporal difference in *GmTLP* expression levels.

Expression Pattern Analysis of *GmTLPs* Under Abiotic Stresses

We used a previously published transcriptome sequencing database to quantify the expression of *GmTLPs* under normal condition, ABA treatment, drought and salt stresses (Wang et al., 2020), and screened 21 *GmTLP* members (**Figure 7**). The results showed that *GmTLP21* was up-regulated and *GmTLP14* was down-regulated under ABA treatment. Under drought stress, six genes were up-regulated, and *GmTLP3*, *8*, *11*, *13*, *19,* and *22*, four genes were down-regulated, and *GmTLP14*, *15*, *18,* and *20*, respectively; Under salt stress, *GmTLP8* was up-regulated and *GmTLP14* was down-regulated. The *P*-value of data in abiotic stress expression profiles is shown in **Supplementary Table 3**.

Responses of *GmTLP8* to Various Treatments

According to the expression profiles of GmTLPs under different abiotic stresses, five genes (*GmTLP8*, *11*, *13*, *19*, and *22*) up-regulated expression under drought stress and one gene (*GmTLP8*) under salt stress were selected for qRT-PCR analysis to further verify their relative expression levels under drought and NaCl treatments. The selected genes were *GmTLP8*, *11*, *13*,

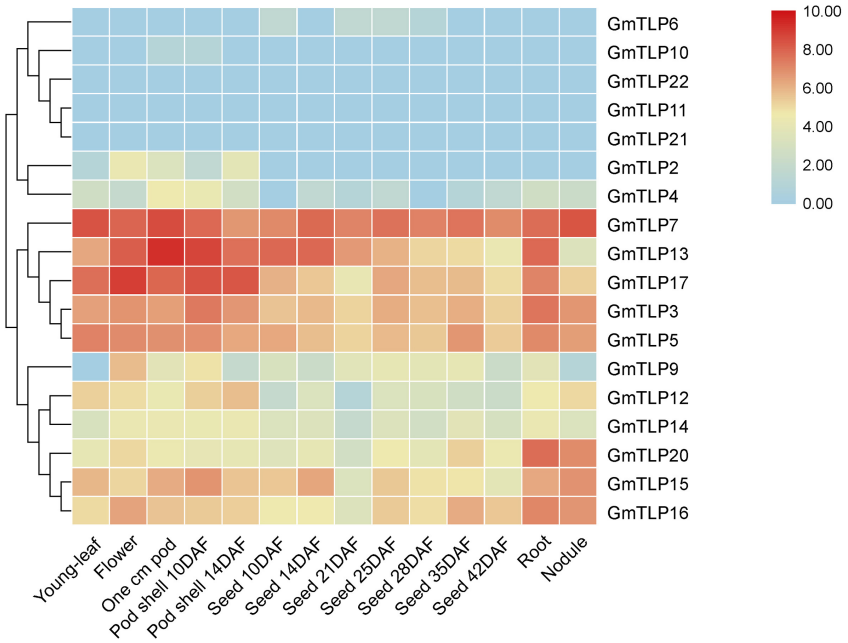


FIGURE 6 | Expression profiles of *GmTLPs* from six soybean tissues. Gene expression was analyzed in soybean young-leaf, flower, pod, pod shell, seed, root, and nodule. The abundance of each transcript (in $\log_{10}$-based FPKM) is represented by the color bar. Red indicates higher and blue indicates lower expression levels.

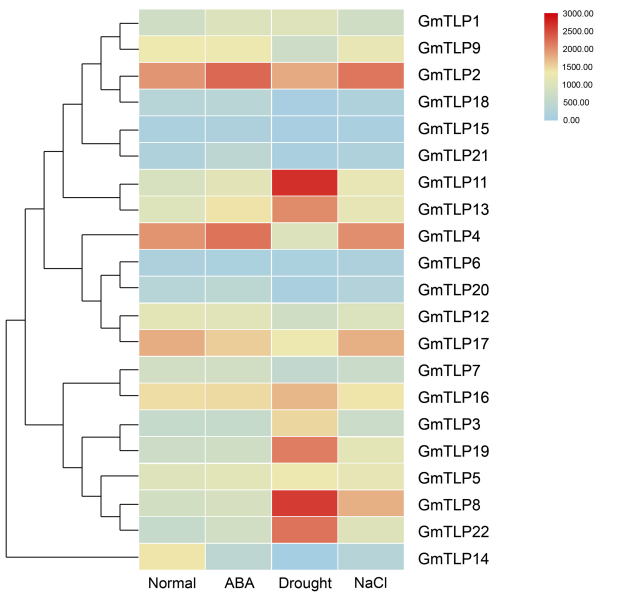


FIGURE 7 | Expression profiles of *GmTLPs* under normal condition (Normal), ABA treatment (ABA), drought (Drought) and salt (NaCl) stresses. The expression abundance of each transcript (in $\log_{10}$-based FPKM) is represented by the color. Red indicates higher and blue indicates lower expression levels.

19, and *22*. Under drought treatment (**Figure 8A**), *GmTLP8* expression peaked at 8 h (with an 8.5-fold increase compared to 0 h), *GmTLP11* at 24 h (3.7-fold), *GmTLP13* at 1 h (3.7-fold), *GmTLP19* at 24 h (4.9-fold), and *GmTLP22* at 8 h (4.8-fold). Under salt treatment (**Figure 8B**), *GmTLP8* expression peaked at 8 h (7.3-fold), *GmTLP11* at 2 h (4.5-fold), *GmTLP13* at 24 h (1.8-fold), *GmTLP19* at 12 h (4.4-fold), and *GmTLP22* at 12 h (4.5-fold). These results showed that *GmTLP8* was the most highly expressed in response to drought and salt treatments, and it was therefore selected for further study.

Subcellular Localization

To determine the subcellular localization of GmTLP8, the open reading frame (ORF) sequence (excluding the termination codon of *GmTLP8*) was fused with the N-terminal of the humanized green fluorescent protein (hGFP) reporter and co-transformed into *Arabidopsis* protoplasts. A 35S:hGFP as the control, the fluorescence signal in the cells was detected by confocal laser scanning microscopy. The fluorescence of GmTLP8 was detected in the nucleus and cytoplasm, while the fluorescence of the control 35S:hGFP was observed in the whole cell. DAPI staining also showed the localization of GmTLP8 in the nucleus (**Figure 9**). It suggests that GmTLP8 act as a transcription factor in the nucleus (Li S. et al., 2021).

GmTLP8 Improved Drought and Salt Tolerance in Soybean Transformants

The stress-tolerant effect of *GmTLP8* in soybean was explored using transgenic soybean hairy root composite plants. The hairy roots of *GmTLP8*-OE, EV-Control and *GmTLP8*-RNAi transgenic lines were used to analyze the relative expression level of *GmTLP8*. qRT-PCR analysis showed that the expression level of *GmTLP8*-OE transgenic hairy roots was significantly higher

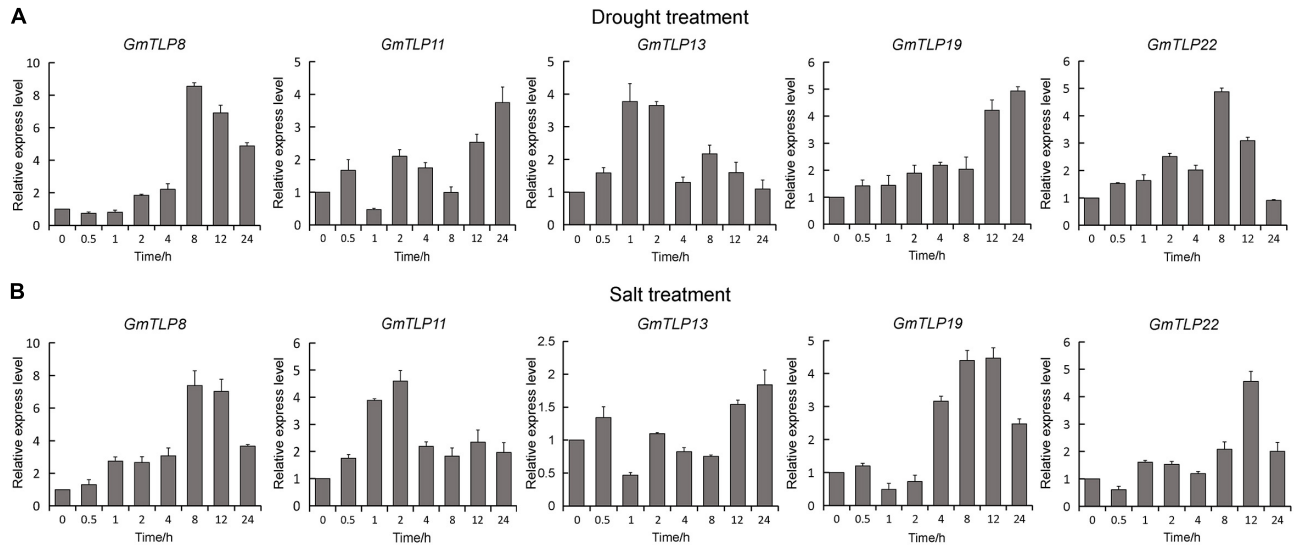


FIGURE 8 | Expression patterns of *GmTLPs* under drought and salt treatments. *GmTLP* gene expression was measured in response to drought **(A)** and NaCl **(B)** treatments using qRT-PCR. The *x*-axes show the duration of treatment and *y*-axes depict relative expression level. The data are shown as mean ± standard deviation (SD) of three technical and three biological replicates.

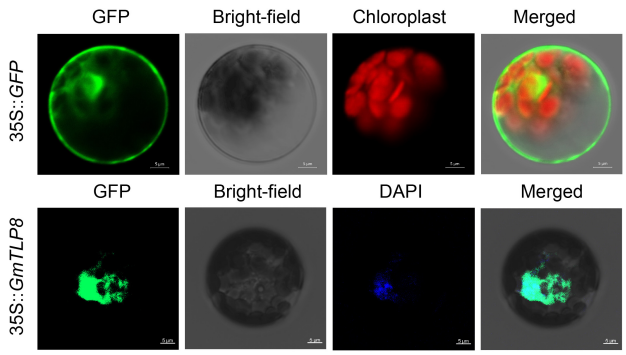


FIGURE 9 | Subcellular localization of *GmTLP8*-hGFP fusion protein. 35S:GFP was used as the control. Scale bar = 5 μm.

than that in EV-Control, and the expression level of *GmTLP8*-RNAi transgenic hairy roots was lower than that in EV-Control (**Supplementary Figure 2G**). Under normal growth conditions, no significant differences were observed between *GmTLP8*-OE, the EV-Control, and *GmTLP8*-RNAi lines (**Figure 10A**). However, after exposure to drought (**Figure 10B**) and salt (**Figure 10C**) treatments, there were significant phenotypic differences between *GmTLP8*-OE, EV-Control, and *GmTLP8*-RNAi plants. Compared with EV-Control, *GmTLP8*-RNAi plants showed more severe leaf dehydration and wilting stress phenotype, whereas *GmTLP8*-OE showed fewer rolled leaves and a delayed leaf wilting phenotype. The survival rates of *GmTLP8*-OE, EV-Control, and *GmTLP8*-RNAi lines under drought stress were 93, 67, and 40%, respectively; these survival rates were comparable to those of salt-stressed plants (**Figure 10H**).

Proline (Pro), malondialdehyde (MDA), hydrogen peroxide (H_2O_2), and superoxide anion (O_2^-) levels are important indicators of the effects of abiotic stresses on plant growth (Leng et al., 2021). Proline is a protective agent against osmotic stress; MDA reflects the degree of lipid oxidative damage; H_2O_2 and O_2^- play immune and signal transduction roles, although excessive accumulation may lead to cell membrane damage (Du et al., 2018; Zhang et al., 2019). Chlorophyll levels are an important indicator of plant photosynthetic capacity (Tanaka and Tanaka, 2006). To further analyze the potential physiological mechanism of *GmTLP8* in plant stress tolerance, we measured the levels of Pro, MDA, H_2O_2, O_2^-, and chlorophyll in the leaves of *GmTLP8*-OE, EV-Control, and *GmTLP8*-RNAi plants under normal growth conditions and under drought or salt stress (**Figures 10F,G,I–K**). Levels of Pro and chlorophyll were higher in *GmTLP8*-OE compared with EV-Control, whereas levels of MDA, H_2O_2, and O_2^- were lower. In contrast, the *GmTLP8*-RNAi lines had lower Pro and chlorophyll levels but higher MDA, H_2O_2, and O_2^- levels than EV-Control.

H_2O_2 and O_2^-, produced by the reactive oxygen species (ROS) pathway in leaf cells under abiotic stress, were measured to assess the degree of damage in leaf cells (Cui et al., 2019). This was done using DAB and NBT to stain the leaves of *GmTLP8*-OE, EV-Control, and *GmTLP8*-RNAi plants (**Figures 10D,E**). Under normal conditions, leaves from the *GmTLP8*-OE, EV-Control, and *GmTLP8*-RNAi lines showed minimal staining, with no significant difference between lines. Under drought and salt stresses, compared with EV-Control, *GmTLP8*-OE leaves showed shallow staining, whereas *GmTLP8*-RNAi showed deeper staining. These results demonstrated that the *GmTLP8*-OE line had lower levels of leaf damage and the *GmTLP8*-RNAi line had more severe leaf damage compared to EV-Control in response to exogenous abiotic stresses. The results of staining leaves with

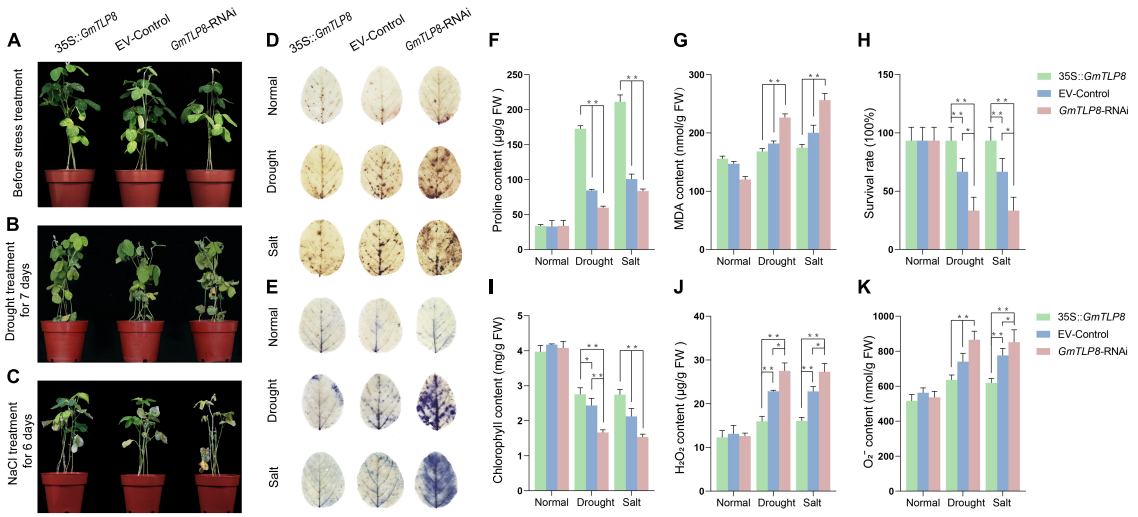


FIGURE 10 | Analysis of the function of soybean *GmTLP8*. **(A–C)** Phenotypes of transgenic soybean hairy root composites *GmTLP8*-OE (35S:*GmTLP8*), EV-Control (empty plasmid), and *GmTLP8*-RNAi plants under **(A)** normal conditions, **(B)** drought stress, or **(C)** salt stress. **(D)** DAB and **(E)** NBT leaf staining of the *GmTLP8*-OE, EV-Control, and *GmTLP8*-RNAi lines under normal conditions and drought or salt stress. The depth of color corresponds to the concentrations of H_2O_2 and O_2^- in the leaves. **(F)** Proline (Pro) content, **(G)** malondialdehyde (MDA) content, **(H)** survival rate, **(I)** chlorophyll content, **(J)** H_2O_2 content, and **(K)** O_2^- content in transgenic soybean hairy root composite plants and EV-control plants under normal conditions and drought or salt stress. Vertical bars indicate ± SD of three technical and three biological replicates. $^*p < 0.05$, $^{**}p < 0.01$ (Student' s *t*-test).

DAB and NBT were consistent with the physiological indexes of H_2O_2 and O_2^- contents.

GmTLP8 Activated Stress-Responsive Genes in Soybean

To analyze the potential stress tolerance mechanism of *GmTLP8*, genes known to be involved in drought and salt stress responses were selected, namely *GmDREB1* (Kasuga et al., 1999), *GmDREB2* (Chen et al., 2007), *GmNAC11* (An et al., 2018), *GmNCED3* (Pandey and Gautam, 2020), *GmSOS1* (Ma et al., 2020), and *GmWRKY27* (Wang et al., 2015) (**Figures 11A–F,H–M**). Expression of these genes in the hairy roots of *GmTLP8*-OE, EV-Control, and *GmTLP8*-RNAi transgenic soybean lines were measured via qRT-PCR. Plants were drought-treated by withholding water for 7 days or salt-treated with 150 mM NaCl for 3 days. Under normal growth conditions, the selected stress responsive genes were expressed at lower levels in all three plant lines compared to plants that had been exposed to drought or salt stress (**Supplementary Figures 2A–F** and **Figures 11A–F,H–M**). In plants that had been stressed, compared with EV-Control, the six stress-related genes were significantly up-regulated in *GmTLP8*-OE plants and down-regulated in *GmTLP8*-RNAi plants. These results suggest that overexpression of *GmTLP8* may activate expression of downstream drought- and salt-response genes.

DISCUSSION

Previous reports have proven that TLP family members participate in plant growth and development, response to abiotic stress, and can also be involved in the ABA signaling pathway (North et al., 1997; Lai et al., 2004; Bao et al., 2014; Chen et al., 2020). Also, reports have confirmed the resistance of TLPs members in *Arabidopsis* (Lai et al., 2004), maize (Chen et al., 2016), wheat (Hong et al., 2015), tomato (Zhang et al., 2020), apple (Xu et al., 2016), and cotton (Li Z. et al., 2021) to abiotic stress, but no report has been found in soybean. We used the NJ method to construct the phylogenetic tree of multiple species. According to the homology of protein sequences, they were divided into five groups, which were similar to the phylogenetic tree group in cotton previously reported (Li Z. et al., 2021). In group I, there is only one GmTLPs member, GmTLP14, which is the same as AtTLP8 in *Arabidopsis* previously reported (Lai et al., 2004). N-terminal of GmTLP14 and AtTLP8 do not contain F-Box domain, indicating that they may come from the same ancestor, so they are classified as the same group (**Figure 2**).

Gene structure analysis showed that each member of the *GmTLPs* had introns and exons, and their numbers were similar to those previously reported in *Arabidopsis* (Lai et al., 2004), indicating that the soybean *TLPs* was evolutionary conserved (**Figure 4A**). Motif analysis showed that except GmTLP14 containing two motifs, the number of motifs contained by other members was not less than five (**Figure 4C**). Analysis of protein conserved domains showed that except GmTLP14 had only one domain, other members contained two/three conserved domains (**Figure 4D**). Multiple sequence alignment marks the protein sites of two key conserved domains (**Figure 3**), consistent with the domain distribution shown in **Figure 4D**. Above results showed that the protein structures of other members of GmTLPs were similar except GmTLP14. Further analysis of transcriptome data showed that the up-regulated gene was *GmTLP8* under drought and salt stresses, while the down-regulated gene was *GmTLP14* (**Figure 7**). The down-regulated expression of *GmTLP14* under

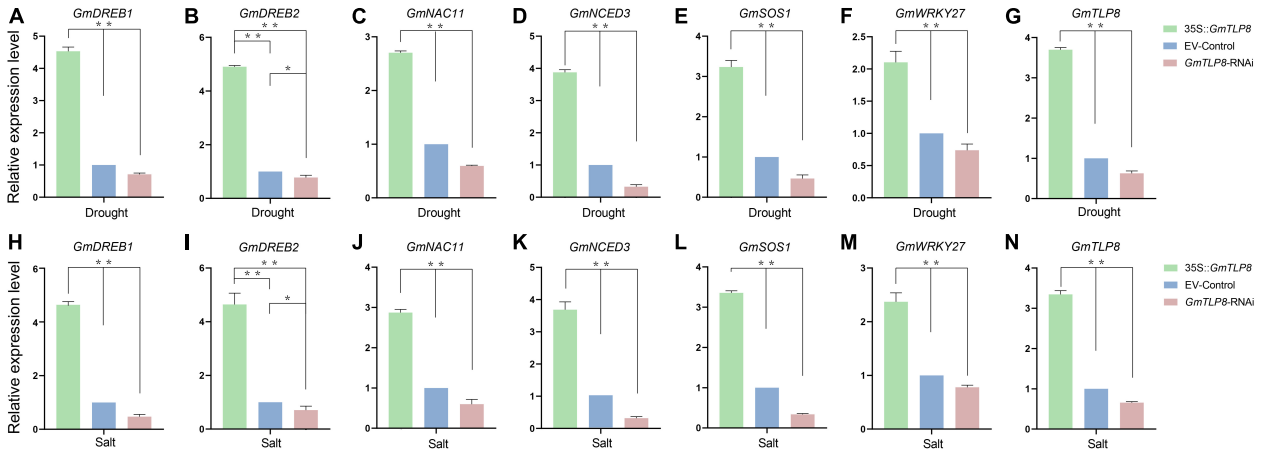


FIGURE 11 | *GmTLP8* regulates stress-responsive gene expression in transgenic soybean plants. **(A–F)** Expression levels of selected stress-related genes in transgenic soybean plants under drought stress. **(H–M)** Expression levels of selected stress-related genes in transgenic soybean plants under salt stress. **(G)** Expression levels of *GmTLP8* in transgenic soybean plants under drought stress. **(N)** Expression levels of *GmTLP8* in transgenic soybean plants under salt stress. Vertical bars indicate ± SD of three technical and three biological replicates. $*p < 0.05$, $**p < 0.01$ (Student' s *t*-test).

drought and salt stresses might be due to the lack of N-terminal F-box domain (**Figure 4D**).

Two key conserved domains are in plant TLPs, the F-box at the N-terminal and the Tub domain at the C-terminal, and these differ from the conserved domains in mammalian TLPs. In mammals, TLPs are binary transcription factors; the N-terminal induces transcriptional activation, and the Tub domain binds to double-stranded DNA (Boggon et al., 1999; Li S. et al., 2021). In plants, the N-terminal F-box can participate in the formation of the Skp1-Cullin1-F-box (SCF) complex, which is an important part of E3 ubiquitin ligase and can participate in protein ubiquitination process (Gagne et al., 2002). It has been reported that GhTULP34 interacts with the subunit GhSKP1A of the SCF complex to form a functional SCF-type E3 ligase, which may be involved in the response of plants to abiotic stresses (Li Z. et al., 2021). *Arabidopsis* AtTLPs and wheat TaTULPs have been shown to interact with specific S-phase kinase-associated protein 1 (SKP1)-like proteins (Lai et al., 2004; Bao et al., 2014; Hong et al., 2015). These findings suggest that TLPs may play a role as subunits of the SCF complex in plants. Yeast two-hybrid assays showed that AtTLP7 and AtTLP11 interacted with NDR1/HIN1-like protein NHL6 (Song et al., 2019). Because both AtTLP11 and AtTLP7 are functional E3 ligases (Bao et al., 2014), it is possible that AtTLP11 and AtTLP7 redundantly manipulate the function of NHL6 by regulating its protein turnover (Bao et al., 2016). The above reports confirmed that the TLPs family, as F-box proteins, played a key role in protein ubiquitination, and may play a key role in plant response to various adverse environmental conditions. Based on these findings, it is speculated that GmTLP14 may be due to the lack of F-Box domain that affects the protein ubiquitination process and then down-regulates its expression under drought and salt stresses. However, the detailed functions of *GmTLP14* gene need to be verified by related experiments. In this study, through RNA-Seq transcriptome data analysis and qRT-PCR verification, we determined the up-regulated expression of *GmTLP8* under drought and salt stresses for subsequent studies (**Figures 7**, **8**). Conserved domain analysis showed that GmTLP8 had two key conserved domains, namely, F-Box and Tub domains (**Figure 4D**), 3D modeling showed the integrity of GmTLP8 C-terminal tubby structure, which might play a role in the response of its to abiotic stress (**Figure 5**).

In this study, *Agrobacterium rhizogenes*-mediated transformation of soybean hairy roots was used to induce transgenic roots in soybean to study the function of *GmTLP8* gene (Kereszt et al., 2007). Through phenotypic observation, leaf staining and physiological index analysis of soybean, it was confirmed that the overexpression of *GmTLP8* enhanced the tolerance of soybean to drought and salt stresses (**Figure 10**). However, this genetic transformation mode is transient expression and cannot be stably inherited to the next generation through sexual reproduction. Therefore, further exploration of the application of *GmTLP8* gene in transgenic drought-resistant and salt-resistant soybean needs further research on transformation.

Previous studies identified genes that play important roles in response to drought and salt stresses. To further analyze the molecular mechanism of *GmTLP8* in regulating stress tolerance, we chose several confirmed stress-related genes (**Figure 11** and **Supplementary Figure 2**). *GmDREB1*, *GmDREB2*, *GmNAC11*, and *GmWRKY27* can specifically recognize and bind to *cis*-acting elements to up-regulate the expression of downstream stress-responsive genes, improving stress tolerance (Marè et al., 2004; Tran et al., 2004; Xu et al., 2011; Bouaziz et al., 2013, 2015; Sarkar et al., 2019). *GmNCED3* is considered to be an important contributor to ABA synthesis and its overexpression enhances drought tolerance in seedlings (Li et al., 2019). *GmSOS1* improves the salt tolerance of plants, potentially playing a role in Na^+ extrusion out of the roots and regulation of Na^+ transport from roots to shoots (Nie et al., 2015; Cao et al., 2018). These

selected stress-related genes were up-regulated in *GmTLP8*-OE plants under drought and salt treatments. Taken together, these indicated that *GmTLP8* responds to drought and salt stresses by activating stress-related transcription factors and the SOS pathway, which provides a scientific basis for further analysis of the function of *GmTLP8* gene under drought and salt stresses. However, further studies were needed to fully elucidate its internal mechanism in abiotic stress response.

CONCLUSION

In the present study, we identified 22 *TLP* genes in the soybean genome. Based on expression patterns in response to abiotic stresses, we found that GmTLP14 showed different structural characteristics and expression patterns from most other members, but the function of *GmTLP14* still needs further experimental verification. In this study, we selected *GmTLP8* with complete structure and up-regulated expression under drought and salt stresses, and verified its expression level under abiotic stress by qRT-PCR. *GmTLP8* was responsive to drought and salt stresses. Overexpression of *GmTLP8* enhanced the tolerance of soybean to drought and salt stresses by activating downstream stress-responsive genes. These results improve understanding of the GmTLP family and provide a basis for further study of the molecular mechanism of *GmTLP8* in soybean abiotic stress responses.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, PRJNA694374.

AUTHOR CONTRIBUTIONS

Z-SX coordinated the project, conceived and designed the experiments, and edited the manuscript. H-RX performed the experiments and wrote the first draft. W-LW and Z-SX revised the manuscript. YL, Z-HH, JC, Y-BZ, and MC contributed to data analysis and managed reagents. T-FY, J-DF, J-CZ, and Y-ZM contributed with valuable discussions. All authors reviewed and approved the final manuscript.

FUNDING

This work was financially supported by the National Natural Science Foundation of China (32071967 and 31871624), the Agricultural Science and Technology Innovation Program (CAAS-ZDRW202109 and CAAS-ZDRW202002), and the Central Public-interest Scientific Institution Basal Research Fund (S2022ZD02).

ACKNOWLEDGMENTS

We are grateful to Lijuan Qiu and Shi Sun of the Institute of Crop Science, Chinese Academy of Agricultural Sciences (CAAS) for kindly providing soybean seeds. Also, we are grateful to Wensheng Hou and Hui Zhang of the Institute of Crop Science, CAAS for kindly providing vectors and the protocol for high-efficiency *A. rhizogenes*-mediated transformation, respectively.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.844545/full#supplementary-material

REFERENCES

An, J. P., Li, R., Qu, F. J., You, C. X., Wang, X. F., and Hao, Y. J. (2018). An apple NAC transcription factor negatively regulates cold tolerance via CBF-dependent pathway. *J. Plant Physiol.* 221, 74–80. doi: 10.1016/j.jplph.2017.12.009

Artimo, P., Jonnalagedda, M., Arnold, K., Baratin, D., Csardi, G., de Castro, E., et al. (2012). ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res.* 40, W597–W603. doi: 10.1093/nar/gks400

Badgandi, H. B., Hwang, S. H., Shimada, I. S., Loriot, E., and Mukhopadhyay, S. (2017). Tubby family proteins are adapters for ciliary trafficking of integral membrane proteins. *J. Cell Biol.* 216, 743–760. doi: 10.1083/jcb.201607095

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., et al. (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 37, W202–W208. doi: 10.1093/nar/gkp335

Bao, Y., Song, W. M., Jin, Y. L., Jiang, C. M., Yang, Y., Li, B., et al. (2014). Characterization of *Arabidopsis* tubby-like proteins and redundant function of AtTLP3 and AtTLP9 in plant response to ABA and osmotic stress. *Plant Mol. Biol.* 86, 471–483. doi: 10.1007/s11103-014-0241-6

Bao, Y., Song, W. M., Pan, J., Jiang, C. M., Srivastava, R., Li, B., et al. (2016). Overexpression of the NDR1/HIN1-like gene NHL6 modifies seed germination in response to abscisic acid and abiotic stresses in *Arabidopsis*. *PLoS One* 11:e0148572. doi: 10.1371/journal.pone.0148572

Bhushan, D., Pandey, A., Choudhary, M. K., Datta, A., Chakraborty, S., and Chakraborty, N. (2007). Comparative proteomics analysis of differentially expressed proteins in chickpea extracellular matrix during dehydration stress. *Mol. Cell Proteomics* 6, 1868–1884. doi: 10.1074/mcp.M700015-MCP200

Boggon, T. J., Shan, W. S., Santagata, S., Myers, S. C., and Shapiro, L. (1999). Implication of tubby proteins as transcription factors by structure-based functional analysis. *Science* 286, 2119–2125. doi: 10.1126/science.286.5447.2119

Borman, A. D., Pearce, L. R., Mackay, D. S., Nagel-Wolfrum, K., Davidson, A. E., Henderson, R., et al. (2014). A homozygous mutation in the TUB gene associated with retinal dystrophy and obesity. *Hum. Mutat.* 35, 289–293. doi: 10.1002/humu.22482

Bouaziz, D., Jbir, R., Charfeddine, S., Saidi, M. N., and Gargouri-Bouzid, R. (2015). The StDREB1 transcription factor is involved in oxidative stress response and enhances tolerance to salt stress. *Plant Cell* 121, 237–248. doi: 10.1007/s11240-014-0698-7

Bouaziz, D., Pirrello, J., Charfeddine, M., Hammami, A., Jbir, R., Dhieb, A., et al. (2013). Overexpression of StDREB1 transcription factor increases tolerance to salt in transgenic potato plants. *Mol. Biotechnol.* 54, 803–817. doi: 10.1007/s12033-012-9628-2

Cai, M., Qiu, D., Yuan, T., Ding, X., Li, H., Duan, L., et al. (2008). Identification of novel pathogen-responsive cis-elements and their binding proteins in the promoter of OsWRKY13, a gene regulating rice disease resistance. *Plant Cell Environ.* 31, 86–96. doi: 10.1111/j.1365-3040.2007.01739.x

Cao, D., Li, Y. Y., Liu, B. H., Kong, F. J., and Tran, L. P. (2018). Adaptive mechanisms of soybean grown on salt-affected soils. *Land Degrad. Dev.* 29, 1054–1064. doi: 10.1002/ldr.2754

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 13, 1194–1202. doi: 10.1016/j.molp.2020.06.009

Chen, M., Wang, Q. Y., Cheng, X. G., Xu, Z. S., Li, L. C., Ye, X. G., et al. (2007). GmDREB2, a soybean DRE-binding transcription factor, conferred drought and high-salt tolerance in transgenic plants. *Biochem. Biophys. Res. Commun.* 353, 299–305. doi: 10.1016/j.bbrc.2006.12.027

Chen, Y., Dai, W., Sun, B., Zhao, Y., and Ma, Q. (2016). Genome-wide identification and comparative analysis of the TUBBY-like protein gene family in maize. *Genes Genomics* 38, 25–36. doi: 10.1007/s13258-015-0338-6

Chen, Z. F., Ru, J. N., Sun, G. Z., Du, Y., Chen, J., Zhou, Y. B., et al. (2021). Genomic-wide analysis of the PLC family and detection of GmPI-PLC7 responses to drought and salt stresses in soybean. *Front. Plant Sci.* 12:631470. doi: 10.3389/fpls.2021.631470

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G., et al. (2003). Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.* 31, 3497–3500. doi: 10.1093/nar/gkg500

Coleman, D. L., and Eicher, E. M. (1990). Fat (fat) and tubby (tub): two autosomal recessive mutations causing obesity syndromes in the mouse. *J. Hered.* 81, 424–427. doi: 10.1093/oxfordjournals.jhered.a111019

Cui, X. Y., Gao, Y., Guo, J., Yu, T. F., Zheng, W. J., Liu, Y. W., et al. (2019). BES/BZR transcription factor TaBZR2 positively regulates drought responses by activation of TaGST1. *Plant Physiol.* 180, 605–620. doi: 10.1104/pp.19.00100

Dong, M. Y., Fan, X. W., Pang, X. Y., and Li, Y. Z. (2019). Decrypting tubby-like protein gene family of multiple functions in starch root crop cassava. *AoB Plants* 11:lz075. doi: 10.1093/aobpla/plz075

Du, Y. T., Zhao, M. J., Wang, C. T., Gao, Y., Wang, Y. X., Liu, Y. W., et al. (2018). Identification and characterization of GmMYB118 responses to drought and salt stress. *BMC Plant Biol.* 18:320. doi: 10.1186/s12870-018-1551-7

Eddy, S. R. (1998). Profile hidden Markov models. *Bioinformatics* 14, 755–763. doi: 10.1093/bioinformatics/14.9.755

Fernandez-Pozo, N., Menda, N., Edwards, J. D., Saha, S., Tecle, I. Y., Strickler, S. R., et al. (2015). The sol genomics network (SGN)–from genotype to phenotype to breeding. *Nucleic Acids Res.* 43, D1036–D1041. doi: 10.1093/nar/gku1195

Finn, R. D., Clements, J., and Eddy, S. R. (2011). HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* 39, W29–W37. doi: 10.1093/nar/gkr367

Gagne, J. M., Downes, B. P., Shiu, S. H., Durski, A. M., and Vierstra, R. D. (2002). The F-box subunit of the SCF E3 complex is encoded by a diverse superfamily of genes in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 99, 11519–11524. doi: 10.1073/pnas.162339999

He, G. H., Xu, J. Y., Wang, Y. X., Liu, J. M., Li, P. S., Chen, M., et al. (2016). Drought-responsive WRKY transcription factor genes TaWRKY1 and TaWRKY33 from wheat confer drought and/or heat resistance in *Arabidopsis*. *BMC Plant Biol.* 16:116. doi: 10.1186/s12870-016-0806-4

Hong, M. J., Kim, D. Y., and Seo, Y. W. (2015). Interactions between wheat Tubby-like and SKP1 like proteins. *Genes Genet. Syst.* 90, 293–304. doi: 10.1266/ggs.14-00084

Hu, B., Jin, J., Guo, A. Y., Zhang, H., Luo, J., and Gao, G. (2015). GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics* 31, 1296–1297. doi: 10.1093/bioinformatics/btu817

Kapeller, R., Moriarty, A., Strauss, A., Stubdal, H., Theriault, K., Siebert, E., et al. (1999). Tyrosine phosphorylation of tub and its association with Src homology 2 domain-containing proteins implicate tub in intracellular signaling by insulin. *J. Biol. Chem.* 274, 24980–24986. doi: 10.1074/jbc.274.35.24980

Kasuga, M., Liu, Q., Miura, S., Yamaguchi-Shinozaki, K., and Shinozaki, K. (1999). Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. *Nat. Biotechnol.* 17, 287–291. doi: 10.1038/7036

Kereszt, A., Li, D., Indrasumunar, A., Nguyen, C. D., Nontachaiyapoom, S., Kinkema, M., et al. (2007). Agrobacterium rhizogenes-mediated transformation of soybean to study root biology. *Nat. Protoc.* 2, 948–952. doi: 10.1038/nprot.2007.141

Kleyn, P. W., Fan, W., Kovats, S. G., Lee, J. J., Pulido, J. C., Wu, Y., et al. (1996). Identification and characterization of the mouse obesity gene tubby: a member of a novel gene family. *Cell* 85, 281–290. doi: 10.1016/s0092-8674(00)81104-6

Kou, Y., Qiu, D., Wang, L., Li, X., and Wang, S. (2009). Molecular analyses of the rice tubby-like protein gene family and their response to bacterial infection. *Plant Cell Rep.* 28, 113–121. doi: 10.1104/pp.103.037820

Lai, C. P., Lee, C. L., Chen, P. H., Wu, S. H., Yang, C. C., and Shaw, J. F. (2004). Molecular analyses of the *Arabidopsis* TUBBY-like protein gene family. *Plant Physiol.* 134, 1586–1597.

Le, D. T., Aldrich, D. L., Valliyodan, B., Watanabe, Y., Ha, C. V., Nishiyama, R., et al. (2012). Evaluation of candidate reference genes for normalization of quantitative RT-PCR in soybean tissues under various abiotic stress conditions. *PLoS One* 7:e46487. doi: 10.1371/journal.pone.0046487

Le, D. T., Nishiyama, R., Watanabe, Y., Mochida, K., Yamaguchi-Shinozaki, K., Shinozaki, K., et al. (2011). Genome-wide expression profiling of soybean two-component system genes in soybean root and shoot tissues under dehydration stress. *DNA Res.* 18, 17–29. doi: 10.1093/dnares/dsq032

Leng, Z. X., Liu, Y., Chen, Z. Y., Guo, J., Chen, J., Zhou, Y. B., et al. (2021). Genome-wide analysis of the DUF4228 family in soybean and functional identification of GmDUF4228-70 in response to drought and salt stresses. *Front. Plant Sci.* 12:628299. doi: 10.3389/fpls.2021.628299

Letunic, I., Doerks, T., and Bork, P. (2012). SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res.* 40, D302–D305. doi: 10.1093/nar/gkr931

Li, S., Wang, N., Ji, D., Zhang, W., Wang, Y., Yu, Y., et al. (2019). A GmSIN1/GmNCED3s/GmRbohBs feed-forward loop acts as a signal amplifier that regulates root growth in soybean exposed to salt stress. *Plant Cell* 31, 2107–2130. doi: 10.1105/tpc.18.00662

Li, S., Wang, Z., Wang, F., Lv, H., Cao, M., Zhang, N., et al. (2021). A tubby-like protein CsTLP8 acts in the ABA signaling pathway and negatively regulates osmotic stresses tolerance during seed germination. *BMC Plant Biol.* 21:340. doi: 10.1186/s12870-021-03126-y

Li, S., Zhang, J., Liu, L., Wang, Z., Li, Y., Guo, L., et al. (2020). SlTLFP8 reduces water loss to improve water-use efficiency by modulating cell size and stomatal density via endoreduplication. *Plant Cell Environ.* 43, 2666–2679. doi: 10.1111/pce.13867

Li, Z., Wang, X., Cao, X., Chen, B., Ma, C., Lv, J., et al. (2021). GhTULP34, a member of tubby-like proteins, interacts with GhSKP1A to negatively regulate plant osmotic stress. *Genomics* 113(1 Pt 2), 462–474. doi: 10.1016/j.ygeno.2020.09.024

Liu, Q. (2008). Identifification of rice TUBBY-like genes and their evolution. *FEBS J.* 275, 163–171. doi: 10.1111/j.1742-4658.2007.06186.x

Ma, X. J., Fu, J. D., Tang, Y. M., Yu, T. F., Yin, Z. G., Chen, J., et al. (2020). GmNFYA13 improves salt and drought tolerance in transgenic soybean plants. *Front. Plant Sci.* 11:587244. doi: 10.3389/fpls.2020.587244

Marè, C., Mazzucotelli, E., Crosatti, C., Francia, E., Stanca, A. M., and Cattivelli, L. (2004). Hv-WRKY38: a new transcription factor involved in cold- and drought-response in barley. *Plant Mol. Biol.* 55, 399–416. doi: 10.1007/s11103-004-0906-7

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E., et al. (2021). Pfam: the protein families database in 2021. *Nucleic Acids Res.* 49, D412–D419. doi: 10.1093/nar/gkaa913

Mukhopadhyay, S., and Jackson, P. K. (2011). The tubby family proteins. *Genome Biol.* 12:225. doi: 10.1186/gb-2011-12-6-225

Nie, W. X., Xu, L., and Yu, B. J. (2015). A putative soybean GmsSOS1 confers enhanced salt tolerance to transgenic *Arabidopsis* sos1-1 mutant. *Protoplasma* 252, 127–134. doi: 10.1007/s00709-014-0663-7

Noben-Trauth, K., Naggert, J. K., North, M. A., and Nishina, P. M. (1996). A candidate gene for the mouse mutation tubby. *Nature* 380, 534–538. doi: 10.1038/380534a0

North, M. A., Naggert, J. K., Yan, Y., Noben-Trauth, K., and Nishina, P. M. (1997). Molecular characterization of TUB, TULP1, and TULP2, members of the novel tubby gene family and their possible relation to ocular diseases. *Proc. Natl. Acad. Sci. U.S.A.* 94, 3128–3133. doi: 10.1073/pnas.94.7.3128

Pandey, A. K., and Gautam, A. (2020). Stress responsive gene regulation in relation to hydrogen sulfide in plants under abiotic stress. *Physiol. Plant.* 168, 511–525. doi: 10.1111/ppl.13064

Papiernik, S. K., Grieve, C. M., Lesch, S. M., and Yates, S. R. (2005). Effects of salinity, imazethapyr, and chlorimuron application on soybean growth and yield. *Commun. Soil Sci. Plant Anal.* 36, 951–967. doi: 10.1081/css-200050280

Santagata, S., Boggon, T. J., Baird, C. L., Gomez, C. A., Zhao, J., Shan, W. S., et al. (2001). G-protein signaling through tubby proteins. *Science* 292, 2041–2050. doi: 10.1126/science.1061233

Sarkar, T., Thankappan, R., Mishra, G. P., and Nawade, B. D. (2019). Advances in the development and use of DREB for improved abiotic stress tolerance in transgenic crop plants. *Physiol. Mol. Biol. Plants.* 25, 1323–1334. doi: 10.1007/s12298-019-00711-2

Song, W. M., Cheng, Z. H., Guo, X. T., Yu, C. Y., Wang, H. H., Wang, J., et al. (2019). Overexpression of NHL6 affects seed production in transgenic *Arabidopsis* plants. *Plant Growth Regul.* 88, 41–47. doi: 10.1007/s10725-019-00486-2

Stretton, C., Litherland, G. J., Moynihan, A., Hajduch, E., and Hundal, H. S. (2009). Expression and modulation of TUB by insulin and thyroid hormone in primary rat and murine 3T3-L1 adipocytes. *Biochem. Biophys. Res. Commun.* 390, 1328–1333. doi: 10.1016/j.bbrc.2009.10.147

Tanaka, A., and Tanaka, R. (2006). Chlorophyll metabolism. *Curr. Opin. Plant Biol.* 9, 248–255. doi: 10.1016/j.pbi.2006.03.011

Tang, Y., Yan, J., Peng, Y., Weng, W., Yao, X., Gao, A., et al. (2021). First report of *Botryosphaeria dothidea* causing gray mold on tartary buckwheat in Southwest China. *Plant Dis.* 25, DIS–07–21–1403–DN. doi: 10.1094/PDIS-07-21-1403-PDN

Tran, L. S., Nakashima, K., Sakuma, Y., Simpson, S. D., Fujita, Y., Maruyama, K., et al. (2004). Isolation and functional analysis of *Arabidopsis* stress-inducible NAC transcription factors that bind to a drought-responsive cis-element in the early responsive to dehydration stress 1 promoter. *Plant Cell* 16, 2481–2498. doi: 10.1105/tpc.104.022699

Wang, F., Chen, H. W., Li, Q. T., Wei, W., Li, W., Zhang, W. K., et al. (2015). GmWRKY27 interacts with GmMYB174 to reduce expression of GmNAC29 for stress tolerance in soybean plants. *Plant J.* 83, 224–236. doi: 10.1111/tpj.12879

Wang, M., Xu, Z., Ahmed, R. I., Wang, Y., Hu, R., Zhou, G., et al. (2019). Tubby-like protein 2 regulates homogalacturonan biosynthesis in *Arabidopsis* seed coat mucilage. *Plant Mol. Biol.* 99, 421–436. doi: 10.1007/s11103-019-00827-9

Wang, T. T., Yu, T. F., Fu, J. D., Su, H. G., Chen, J., Zhou, Y. B., et al. (2020). Genome-wide analysis of the GRAS gene family and functional identification of GmGRAS37 in drought and salt tolerance. *Front. Plant Sci.* 11:604690. doi: 10.3389/fpls.2020.604690

Wang, X., Wang, Y., Tian, J., Lim, B. L., Yan, X., and Liao, H. (2009). Overexpressing AtPAP15 enhances phosphorus efficiency in soybean. *Plant Physiol.* 151, 233–240. doi: 10.1104/pp.109.138891

Wang, Z. Q., Yu, T. F., Sun, G. Z., Zheng, J. C., Chen, J., Zhou, Y. B., et al. (2021). Genome-wide analysis of the *Catharanthus roseus* RLK1-Like in soybean and GmCrRLK1L20 responds to drought and salt stresses. *Front. Plant Sci.* 12:614909. doi: 10.3389/fpls.2021.614909

Xu, J., Xing, S., Sun, Q., Zhan, C., Liu, X., Zhang, S., et al. (2019). The expression of a tubby-like protein from *Malus domestica* (MdTLP7) enhances abiotic stress tolerance in *Arabidopsis*. *BMC Plant Biol.* 19:60. doi: 10.1186/s12870-019-1662-9

Xu, J. N., Xing, S. S., Zhang, Z. R., Chen, X. S., and Wang, X. Y. (2016). Genome-wide identification and expression analysis of the tubby-like protein family in the *Malus domestica* genome. *Front. Plant Sci.* 7:1693. doi: 10.3389/fpls.2016.01693

Xu, Z. S., Chen, M., Li, L. C., and Ma, Y. Z. (2011). Functions and application of the AP2/ERF transcription factor family in crop improvement. *J. Integr. Plant Biol.* 53, 570–585. doi: 10.1111/j.1744-7909.2011.01062.x

Xu, Z. S., Ni, Z. Y., Liu, L., Nie, L. N., Li, L. C., Chen, M., et al. (2008). Characterization of the TaAIDFa gene encoding a CRT/DRE-binding factor responsive to drought, high-salt, and cold stress in wheat. *Mol. Genet. Genomics* 280, 497–508. doi: 10.1007/s00438-008-0382-x

Xu, Z. S., Xia, L. Q., Chen, M., Cheng, X. G., Zhang, R. Y., Li, L. C., et al. (2007). Isolation and molecular characterization of the *Triticum aestivum* L. ethylene-responsive factor 1 (TaERF1) that increases multiple stress tolerance. *Plant Mol. Biol.* 65, 719–732. doi: 10.1007/s11103-007-9237-9

Yang, Z., Zhou, Y., Wang, X., Gu, S., Yu, J., Liang, G., et al. (2008). Genomewide comparative phylogenetic and molecular evolutionary analysis of tubby-like protein family in *Arabidopsis*, rice, and poplar. *Genomics* 92, 246–253. doi: 10.1016/j.ygeno.2008.06.001

Yu, T. F., Liu, Y., Fu, J. D., Ma, J., Fang, Z. W., Chen, J., et al. (2021). The NF-Y-PYR module integrates the abscisic acid signal pathway to regulate plant stress tolerance. *Plant Biotechnol. J.* 19, 2589–2605. doi: 10.1111/pbi.13684

Zhang, X. Z., Zheng, W. J., Cao, X. Y., Cui, X. Y., Zhao, S. P., Yu, T. F., et al. (2019). Genomic analysis of stress associated proteins in soybean and the role of GmSAP16 in abiotic stress responses in *Arabidopsis* and soybean. *Front. Plant Sci.* 10:1453. doi: 10.3389/fpls.2019.01453

Zhang, Y., He, X., Su, D., Feng, Y., Zhao, H., Deng, H., et al. (2020). Comprehensive profiling of tubby-like protein expression uncovers ripening-related TLP genes in tomato (*Solanum lycopersicum*). *Int. J. Mol. Sci.* 21:1000. doi: 10.3390/ijms21031000

Zuo, Z. F., Kang, H. G., Hong, Q. C., Park, M. Y., Sun, H. J., Kim, J., et al. (2020). A novel basic helix-loop-helix transcription factor, ZjICE2 from *Zoysia japonica* confers abiotic stress tolerance to transgenic plants via activating the DREB/CBF regulon and enhancing ROS scavenging. *Plant Mol. Biol.* 102, 447–462. doi: 10.1007/s11103-019-00957-0

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 27 April 2022
doi: 10.3389/fpls.2022.857149

Genome-Wide Association Studies of Salt-Alkali Tolerance at Seedling and Mature Stages in *Brassica napus*

*Guofang Zhang[1,2], Yan Peng[1], Jinzhi Zhou[1], Zengdong Tan[1,2], Cheng Jin[1], Shuai Fang[1,2], Shengzhu Zhong[3], Cunwang Jin[4], Ruizhen Wang[3], Xiaoliang Wen[3], Binrui Li[3], Shaoping Lu[1,2], Guangsheng Zhou[1,2], Tingdong Fu[1,2], Liang Guo[1,2] and Xuan Yao[1,2]**

[1] National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan, China, [2] Hubei Hongshan Laboratory, Wuhan, China, [3] Agriculture and Animal Husbandry Technology Promotion Center, Inner Mongolia, China, [4] Green Industry Development Center, Inner Mongolia, China

Most plants are sensitive to salt-alkali stress, and the degree of tolerance to salt-alkali stress varies from different species and varieties. In order to explore the salt-alkali stress adaptability of *Brassica napus*, we collected the phenotypic data of 505 *B. napus* accessions at seedling and mature stages under control, low and high salt-alkali soil stress conditions in Inner Mongolia of China. Six resistant and 5 sensitive materials, respectively, have been identified both in Inner Mongolia and Xinjiang Uygur Autonomous Region of China. Genome-wide association studies (GWAS) for 15 absolute values and 10 tolerance coefficients (TCs) of growth and agronomic traits were applied to investigate the genetic basis of salt-alkali tolerance of *B. napus*. We finally mapped 9 significant QTLs related to salt-alkali stress response and predicted 20 candidate genes related to salt-alkali stress tolerance. Some important candidate genes, including *BnABA4*, *BnBBX14*, *BnVTI12*, *BnPYL8*, and *BnCRR1*, were identified by combining sequence variation annotation and expression differences. The identified valuable loci and germplasms could be useful for breeding salt-alkali-tolerant *B.napus* varieties. This study laid a foundation for understanding molecular mechanism of salt-alkali stress adaptation and provides rich genetic resources for the large-scale production of *B. napus* on salt-alkali land in the future.

Keywords: GWAS, salt-alkali stress, seedling stage, mature stage, *Brassica napus*

OPEN ACCESS

Edited by:
Prasanta Kumar Subudhi, Louisiana State University, United States

Reviewed by:
Fugui Zhang, Anhui Agricultural University, China
Tao Hu, Wuhan Botanical Garden (CAS), China

****Correspondence:***
Xuan Yao
xuanyao@mail.hzau.edu.cn

Specialty section:
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

Received: *18 January 2022*
Accepted: *28 March 2022*
Published: *27 April 2022*

Citation:
Zhang G, Peng Y, Zhou J, Tan Z, Jin C, Fang S, Zhong S, Jin C, Wang R, Wen X, Li B, Lu S, Zhou G, Fu T, Guo L and Yao X (2022) Genome-Wide Association Studies of Salt-Alkali Tolerance at Seedling and Mature Stages in Brassica napus. Front. Plant Sci. 13:857149. doi: 10.3389/fpls.2022.857149

INTRODUCTION

Salt-alkali stress has been considered as a major environmental threat to the entire terrestrial ecosystem, which is one of the main reasons to inhibit the growth, development, and even the yield of crops in the world (Lin et al., 2014, 2018a,b; Guo et al., 2015). Although there are many ways to improve crop tolerance to salt-alkali stress (Sanchez-Bel et al., 2012), there is still a challenge to maintain the food supply as the population of the world continues to increase (Hickey et al., 2019; Tyerman et al., 2019). It is an important step to increase crop agricultural productivity through in-depth understanding of the genetic and molecular mechanisms for improving crop salt-alkali tolerance.

Soil salinization is a universal environmental problem worldwide. According to the incomplete statistics of the Food and Agriculture Organization of the United Nations, the area of salt-affected

land in the world is as high as 9.54×10^8 hm^2 (Munns and Tester, 2008; Hussain et al., 2019; Singh et al., 2020). The distribution area of salt-alkali soil in China is very wide (Wang et al., 2011). According to the estimation from the Ministry of Agriculture and Rural Affairs of China, it is about 20-million-hectare salt-alkali lands (Song and Liu, 2017; Xu et al., 2020), accounting for one fourth of the total area of the cultivated land in China. Saline sodic soil is the most widely distributed soil type of semi-arid areas. Northwestern, northeastern, and central China, especially Xinjiang and Inner Mongolia Autonomous Region of China, are typical arid areas, where soil salinization is very serious and a large area of saline-alkali land has been formed (Lin, 2005; Jin et al., 2008; Lin et al., 2016). It not only becomes a major obstacle to the development of local agriculture but also an important factor affecting the ecological stability. Such a large area of salt-alkali soil has a great impact on agricultural production of China, especially for large-scale mechanized planting and industrial production (Pang et al., 2016).

Salt-alkali stress actually includes two abiotic stresses, including salt stress and high pH value (Zhou and Yang, 2003; Fang et al., 2021), which causes osmotic stress, ion toxicity, high pH value, and other hazards to plants (Huan, 2015). When the salt content reaches 0.2% or pH reaches 8., salt-alkali stress will result in a serious harm to plants (Chen et al., 2018). When the salt ions and pH value in the soil are too high, it will seriously affect the water potential balances in the plant roots, thereby affecting the absorption and transportation of water, and indirectly affecting the transportation and absorption of plant nutrients, which severely restricts the normal growth and development of plants and affects crop yield (Zhu, 2002; Julkowska and Testerink, 2015). When plants are subjected to short-term salt-alkali stress, a higher osmotic pressure will cause cell membrane damages, thereby leading to water loss and showing leaf wilting (Khan and Weber, 2006; Chikelu et al., 2007; Munns and Tester, 2008). Long-term salt-alkali stress causes the imbalance of ion absorption and ion transport, resulting in nutritional imbalance and metabolic disorders. This will inhibit the growth inhibition of new leaves, early senescence, and shedding of old leaves (Zhu, 2001). The plants themselves can resist salt-alkali stress to ensure their normal growth and development to some extent. However, when the salt-alkali stress level is greater than the resistance of the plant, the normal growth and the development of the plant are restricted (Munns and Tester, 2008; Julkowska and Testerink, 2015). The regulation mechanism of plants to salt-alkali stress mainly includes osmotic regulation, an active oxygen scavenging mechanism, ion regional regulation, and selective absorption of ions (Dwivedi, 2004; Zhu, 2016; Sanders, 2020).

Brassica napus has an important economic value of oil use, feed use, vegetable use, and tourism (Lu et al., 2019; Song et al., 2020). Previous studies have revealed that overexpression of *AtNHX1* and *BnaABF2* could significantly improve salt tolerance in transgenic *B. napus* and *Arabidopsis*, respectively (Zhang et al., 2001; Zhao et al., 2016). Using RNA-seq, 582 transcription factors and 438 transporter genes have been identified by analyzing differential expression genes (DEGs) under salt stress (Yong et al., 2014). Moreover, using the linkage analysis, 9 QTLs were mapped on Chromosomes A02, A04, and C03 with recombinant inbred lines (RIL) population from GH06 and P174 using the linkage analysis (Hou et al., 2017). Genome-wide association study (GWAS) is an effective way to identify QTLs related to complex traits for crops, which can improve efficiency and cost savings to parse the genetic basis based on abundant genetic variation (Pearson and Manolio, 2008; Huang et al., 2012; De et al., 2014; Patishtan et al., 2017). In recent years, some QTLs and candidate genes related to salt stress response have gradually been identified through GWAS. About 225 significant SNPs (He et al., 2017), 44 QTLs including 56 possible candidate genes (Wan et al., 2018), and 65 candidate genes (Zhang et al., 2017) related to salt stress response of *B. napus* were identified by GWAS using the 60K Brassica Illumina® Infinium SNP array at the germination stage, respectively. In addition, 25 QTLs, including 38 promising candidate genes associated with salt stress tolerance at the seedling stage, were also identified by GWAS using the 60k *Brassica* Infinium® SNP array in 368 *Brassica* accessions (Wan et al., 2017). Based on the next-generation sequencing technology and GWAS approach, a total of 62 QTLs were identified for salt tolerance index, growth indicators, and Na^+/K^+ ratio traits at the seedling stage in 85 rapeseed inbred lines (Yong et al., 2015). About 142 significant SNP markers and 117 possible candidate genes associated with salt tolerance of *B. napus* were also identified for SSI and STI values using 201,817 high-quality SNPs (Wassan et al., 2021). In addition, our previous study has revealed that a total of 177 and 228 candidate genes related to salt stress tolerance were identified at germination and seedling stages, respectively, and some important candidate genes, such as *BnCKX5* and *BnERF3*, which were found to play a positive role in salt and mannitol stress response at the germination stage (Zhang et al., 2022). Although a great number of QTLs and candidate genes related to salt stress have been identified, the molecular mechanism for salt-alkali tolerance of *B. napus* remains unclear (Chen et al., 2021). Moreover, the experiments conducted in the greenhouse cannot completely reflect the true response of plants to salt-alkali stress. Therefore, it is of great significance to carry out field trials on salt-alkali lands.

In this study, 11 extreme materials have been screened from 505 *B. napus* accessions grown on the salt-alkali lands of Inner Mongolia and Xinjiang Uygur Autonomous Region of China. We also performed GWAS for 15 absolute values and 10 tolerance coefficients (TCs) of some growth and agronomic traits in 505 *B. napus* accessions under salt-alkali stress to identify main-effect QTLs and candidate genes associated with salt-alkali stress response. This study will help us to reveal the genetic mechanism of salt-alkali stress adaptation and provide a valuable reference to genetic improvement in salt-alkali tolerance to *B. napus*.

MATERIALS AND METHODS

Materials

These materials consisting of 505 natural accessions of *B. napus* (**Supplementary Table 1**) (Tang et al., 2021) were used as research objects under control, low salt-alkali, and high salt-alkali stress conditions at seedling and mature stages.

Determination of Salt Content and pH Value of Salt-Alkali Land

Eight spots covering the entire salt-alkali land in Inner Mongolia of China were randomly selected, and 3 soil samples (10–20-cm depth) were taken in within the area (50–100 m^2) of each spot. After the dried samples were filtered with a 1-mm sieve, 100 g of the filtered samples were transferred to a 1-L plastic bottle, and 500-ml carbon dioxide distilled water was added into the bottle (Water: Soil = 5:1). Suction and filtration were immediately performed after shaking for 8 min. We put 50–100 ml of the test solution into an evaporating dish with known weight, and then evaporated the evaporating dish to dryness. We wiped the outside of the evaporating dish with a filter paper and then put it in an oven at 100–105°C for 4 h. After it was transferred to a desiccator to cool for 30 min, it was weighed with an analytical balance. The drying step was repeated until the difference of the weights between the two drying steps was <.0003 g. Finally, we obtained weight of the dried residue (W). After adding 15% hydrogen peroxide solution, the residue continued to be heated in a water bath to remove organic matter. Then, the total amount of soluble salt (S) was obtained after drying. Salt content of the salt-alkali land was calculated as follows: soluble salt (%) = S (g)/W (g) *100.

For determination of pH value of salt-alkali land, 10 g of the dry soil sample after being filtered by a 1-mm mesh was weighed and transferred into a 25-ml beaker. We added 10-ml.01 mol/L $CaCl_2$ solution into the sample and mixed it well (Water: Soil = 1:1). And then the sample was placed for 10 min. pH value was measured with a pH meter (PHS-25, Shanghai Leici Instrument, Inc., Shanghai, China).

Treatment Conditions and Experimental Design

Wuyuan County (105°12′-109°53′ E, 40°13′-42°28′ N) of Bayan Nur City is located in the western part of the Inner Mongolia Autonomous Region, which belongs to the hinterland of the Hetao Plain. The ingredients of saline-alkali land are dominated by sulfate, chloride, chloride-sulfate or sulfate-chloride in the Hetao Plain. We designed 3 treatment groups, including control (Salt:0.1–0.25%, pH: 7.5–8.), low salt-alkali (Salt:0.35–0.53%, pH: 8.—8.5), and high salt-alkali (Salt:0.64–1.05%, pH: 8.3–9.) conditions (**Supplementary Figure 1**). Each germplasm of 505 *B. napus* accessions was sown in 4 rows, and the space between each row was 20 cm. Each row contained 8 individual plants, and the space between each individual plant was 20 cm. At the same time, the second technical repetition was set up in the adjacent plots in different salt-alkali lands (**Figure 1A**). The sowing time is May 7, 2017. Before sowing, 225 kg/hm^2 of complex fertilizer (N:P_2O_5:K_2O, 25:10:16; total nutrients ≥ 51%; Shandong Stanley Fertilizer Co., Ltd.) was applied as base fertilizer to maintain normal growth and development of plants on May 7, 2017. In addition, 150 kg/hm^2 of complex fertilizer as supplemental fertilizer was also applied on July 1, 2017. The irrigation was conducted on May 7, 2017, May 17, 2017, and July 1, 2017, respectively (**Supplementary Table 30**). We measured its salt content and pH value in control, low salt-alkali, and high salt-alkali stress conditions at seedling and mature stages (**Supplementary Table 2**). Then, some agronomic or growth indicators were determined (**Supplementary Table 3**) at the seedling (July 9, 2017) and harvest stages (September 9, 2017), respectively.

Shihezi City (84°58′-86°24′ E, 43°26′-45°20′ N) is located in the northern part of Xinjiang Uygur Autonomous Region, which belongs to the Southern margin of Junggar Basin. The ingredients of saline-alkali land are dominated by soda, chloride, sulfate-chloride, chloride-sulfate, or sulfate in Shihezi City. In order to further verify and screen extreme materials under salt-alkali condition, we designed 2 treatment groups, including control soil (Salt:0.05–0.15%, pH: 6.8–7.3), and high salt-alkali soil (Salt:0.80–1.15%, pH: 8.5–10). Each germplasm of 505 *B. napus* accessions was sown in 4 rows, and the space between each row was 20 cm. Eight individual plants were grown in each row for one replicate, and the space between each individual plant was 20 cm. At the same time, another technical repetition was conducted in the adjacent plots. The sowing date is July 10, 2018. Before sowing, 225 kg/hm^2 of complex fertilizer (N:P_2O_5:K_2O, 25:10:16; total nutrients ≥ 51%; Shandong Stanley Fertilizer Co., Ltd.) was applied as base fertilizer to maintain normal growth and development of plants on July 10, 2017. In addition, 150 kg/hm^2 of complex fertilizer as supplemental fertilizer was also applied on August 4, 2017. The irrigation was conducted on July 10, 2017, August 4, 2017, and August 20, 2017, respectively (**Supplementary Table 31**). Then, we observed the response of the extreme materials at the seedling stage on September 19, 2018.

Determination of Growth and Agronomic Indexes

Some traits, such as plant height (PH), ground dry weight (GDW), and yield (Y), were measured (**Supplementary Table 4**). For GDW measurement, materials were harvested at seedling and mature stages, respectively, and dried at 85°C for at least 72 h before the samples were weighed. We collected the traits data from 5 biological replicates for each accession (Lv et al., 2016; Yu et al., 2017). The tolerance coefficient (TC) was calculated as the ratio between the agronomic and growth indexes under salt-alkali stress and control condition (**Supplementary Table 5**).

Screening Criteria of Extreme Materials

Combined with the response of these extreme materials to salt-alkali land in Xinjiang Uygur and Inner Mongolia Autonomous Region, the trait values of each variety were ranked, and then the top 10% most tolerant materials and the bottom 10% most sensitive materials were selected as extreme materials, respectively. Finally, a number of reliable extreme materials were screened out through Venn analysis between the two developmental stages.

Statistical Analysis and Correlation Analysis

The statistically significant differences between different data sets were evaluated by the Student's *T*-test ($P < 0.05$). We finally calculated the average of all traits with 5 replicates per accession. We used the scale package of R language (https://www.r-project.org/) to analyze data standardization and R

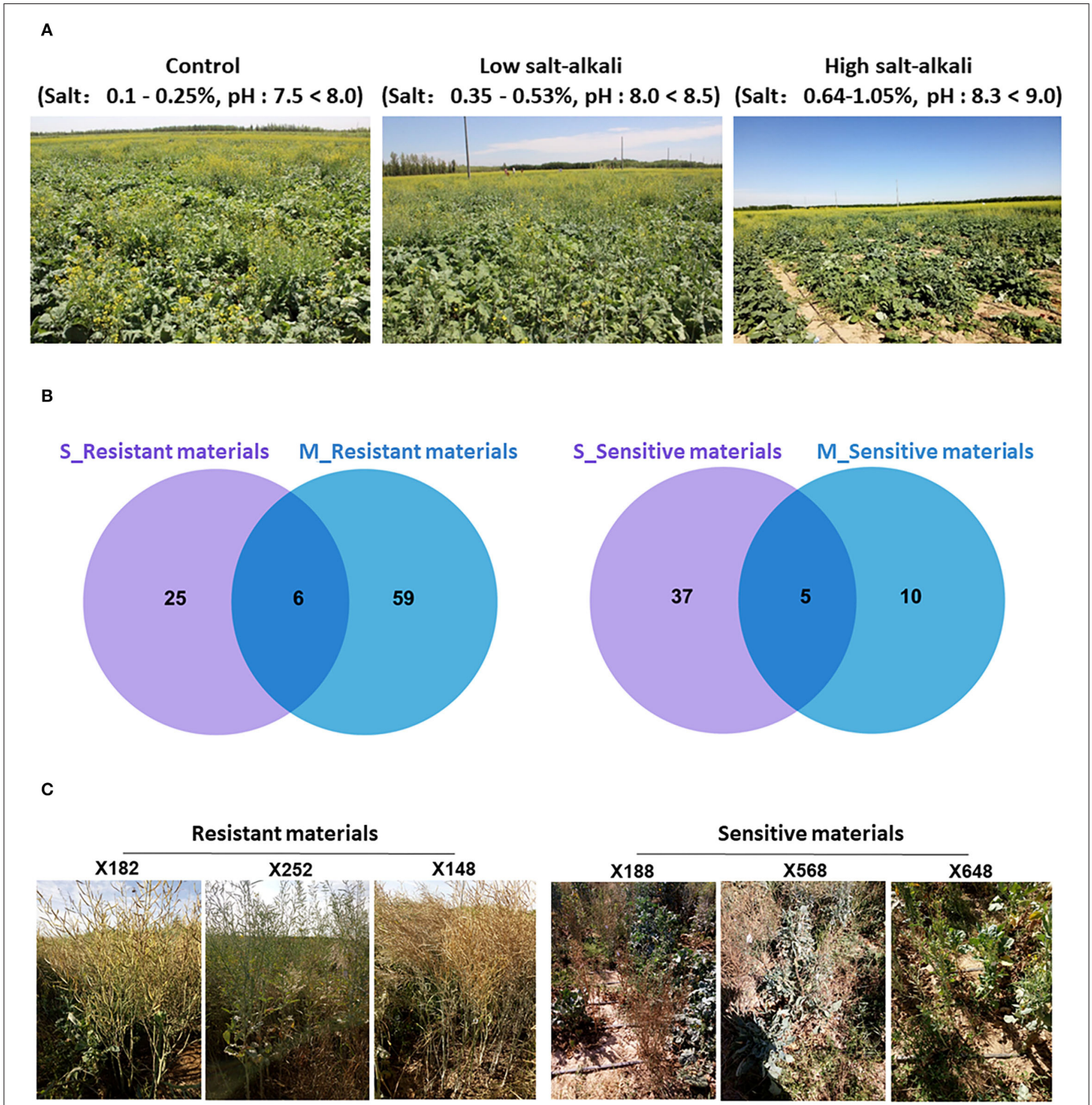


FIGURE 1 | The plant growth stage of 505 *B. napus* accessions and screening results of extreme materials under salt-alkali stress conditions at seedling and mature stages. **(A)** Growth state of 505 *B. napus* accessions under control, low salt-alkali, and high salt-alkali conditions. **(B)** Venn diagrams showing resistant and sensitive materials under the salt-alkali condition at seedling and mature stages. **(C)** The response of extreme materials to salt-alkali stress at the mature stage in Inner Mongolia Autonomous Region of China.

language stat.desc packages to analyze phenotypic variation. The correlation coefficient was calculated using the corr.test of psych package (https://www.r-project.org/), and figures were plotted using the corrplot package of R language (https://www.r-project.org/).

Analysis of Genetic Effect, Treatment Effect, and Heritability

In order to determine the variance components, we used the mixed-effects variance analysis of the avo package by R language (https://www.r-project.org/), including genetic effects (G_effect),

treatment effects (E_effect), and interaction effects between genetic and treatment effects (G&E effect). Heritability is a commonly used reference index in crop genetic breeding. The broad sense heritability (H_2b) of all traits is calculated as follows: $H_2b = \sigma_G^2/(\sigma_G^2 + \sigma_{GE}^2/n + \sigma_e^2/nr)$ where σ_G^2 was the genotype variance, $\sigma_e^2$was the error variance, and σ_{GE}^2 was interaction between genetic and environmental effects. The estimates of *H2b* were analyzed by the AVOVA for ANOVA using the lmer function of the lme4 package by R environment (https://www.r-project.org/) (Chen et al., 2014; Li et al., 2020; Zhang et al., 2022).

Genome-Wide Association Analysis (GWAS)

Our laboratory collected 505 *Brassica napus* germplasm resources to construct a re-sequencing association population (Tang et al., 2021). The high-quality clean reads data were obtained by the BWA (v0.75) software (https://github.com/lh3/bwa) (Li and Durbin, 2009). The reference genome is derived from "Brassica v4.1" ("Darmor-*bzh*") genome (http://www.Genoscope.cns.fr/brassicanapus/data/). We also performed GWAS with 7,671,951 SNPs (MAF >0.05) using GEMMA linear mixed models (Zhou and Stephens, 2012). It is worth noting that we divided the gene components into 300 kb bins. If there are multiple suggestive SNPs in the same bin, it is considered to be located at the same site, and the suggestive SNP with the smallest *P*-value is regarded as the lead SNP.

Identification of Candidate Genes and Polymorphisms

The *P*-value should be 1/marker number, 0.05/marker number or 0.01/marker number, which were named suggestive *P*-Value, significant *P*-value, and high-significant *P*-Value. High-density SNPs markers can improve the accuracy of candidate gene positioning, but there is often a certain linkage effect between adjacent SNP markers, so we often combined inter-linked SNP markers for one molecular marker. Therefore, we turned the final merged SNP markers, which were called effective SNP markers. The number of SNP markers used to calculate the *P*-value was far less than the total SNPs markers. Previous studies have shown that *B. napus* can be divided into two groups: spring and semi winter. In addition, semi winter was divided into Semi Winter 1 and Semi Winter 2 (Tang et al., 2021). The average number of SNP markers for different sub-species is about 7.38 million (Oberved_number) and 1.08 million (Effective_number), respectively. Effective SNPs account for about 14% of total SNPs. The average threshold is about 6 [∼1. e-06, -$\log_{10}$ (*P*-value)] between different subspecies. This value was previously used by our laboratory in GWAS analysis of the same 505 *B. napus* accessions (Tang et al., 2021). Therefore, this value is usually considered as a recognized threshold in the 505 *B. napus* accessions. The suggestive and significant *P*-value threshold of the entire population was set to be 1. e-06. All SNPs exceeded the significance criterion were assessed for location of candidate genes. If coding regions were present, the potential impact on protein of each SNP was subsequently determined using the "Allele finder" facility Gene Ontology analysis. The candidate genes with differential gene expression ratios between the control treatment and stress treatment (R ≥ 2 or ≤0.5) (https://bigd.big.ac.cn/) (Zhang et al., 2019) were identified within 200 kb upstream or downstream of the lead SNP. We finally determined the protein function corresponding to the candidate gene in *Brassica napus* based on the homologous gene information on *Arabidopsis* (Chalhoub et al., 2014).

RESULTS

Phenotypic Variation for Salt-Alkali Tolerance Traits to 505 *B. napus* Accessions at Seedling and Mature Stages

In order to evaluate the response of the 505 *B. napus* accessions to salt-alkali stress, we determined the related traits at seedling and mature stages under control (Salt:0.1–0.25%, pH: 7.5–8., CK), low salt-alkali (Salt:0.35–0.53%, pH: 8.−8.5, L), and high salt-alkali (Salt:0.64–1.05%, pH: 8.3–9., H) stress conditions, respectively (**Figure 1A**; **Supplementary Figure 1**). Plant height (S_PH_CK, S_PH_L, and S_PH_H) and aboveground dry weight (S_GDW_CK, S_GDW_L, and S_GDW_H) were measured at the seedling stage (**Supplementary Table 4**). In addition, the traits, such as plant height (M_PH_CK, M_PH_L, and M_PH_H), aboveground dry weight (M_GDW_CK, M_GDW_L, and M_GDW_H) and yield (M_Y_CK, M_Y_L, and M_Y_H) were obtained at the mature stage (**Supplementary Table 4**). We found that salt-alkali stress significantly inhibited the growth, development, and even yield of crops (**Supplementary Figure 2**).

Moreover, in order to better reflect the salt-alkali stress response, we also focused on the tolerance coefficients (TCs) of traits, which were calculated from the ratio of all traits between salt-alkali stress and control condition, denoted by the suffix type "trait_R1" under low salt-alkali stress conditions and "trait_R2" under high-salt-alkali stress conditions (**Supplementary Table 5**) (Guo et al., 2018). Descriptive statistics provided a good description of the phenotypic variation on 505 *B. napus* accessions under control condition and salt-alkali stress conditions. "Mean," 'Max," "Min," "SD," "SE," and "CV" were used to calculate the descriptive statistics of the average value of all traits (**Supplementary Table 6**). We found that the coefficient of variation (CV, %) of most traits reached more than 20%. Most traits of the 505 *B. napus* accessions show significant variations between control condition and salt-alkali stress conditions. Interestingly, large variations in plant height (M_PH), aboveground dry weight (M_GDW), and yield (M_Y) were observed among the accessions under salt-alkali stress conditions at the mature stage, whereas plant height (S_PH) and ground dry weight (S_GDW) showed small variations under the control condition at the seedling stage (**Supplementary Table 6**). These results suggest that the developmental stage has a big effect on the phenotypic variation under salt-alkali stress conditions.

After trait values of each strain were ranked, the top 10% most tolerant and bottom 10% most sensitive materials were selected as extreme materials, respectively. A number of reliable extreme materials have been screened out through Venn analysis between the two developmental stages (**Figure 1B**;

Supplementary Figure 3). Combining the response of these extreme materials on salt-alkali land in Xinjiang Uygur and Inner Mongolia Autonomous Region of China, we finally identified that "X182," "X252," "X148," "X200," "X278," and "X398" were as resistant varieties, and "X188," "X568," "X648," "X552," and "X756" were as sensitive varieties (**Figure 1C**; **Supplementary Figure 4**) under the two different salt-alkali stress environments. These extreme materials are valuable germplasm resources for the genetic improvement of salt-alkali tolerance in *B. napus*.

Frequency Distribution and Correlation Analysis Among all Traits at Seedling and Mature Stages

We also performed normal distribution detection on the original absolute value (**Supplementary Table 7**) and TCs of all traits for GWAS (**Supplementary Table 8**), respectively. The absolute value of all traits conformed to normal or lognormal distributions and the TCs of all traits for GWAS also belonged to normal distributions. The lognormal distribution of S_PH and S_GDW may be attributed to the significant phenotypic variations among 505 *B. napus* accessions. Correlation analysis for all traits at seedling and mature stages was calculated and presented in **Supplementary Table 9**. The two traits at the seedling stage and the three traits at the mature stage showed significant correlations with "r" under control, low, and high salt-alkali stress conditions, which ranged from 0.021 to 0.982 between each other (**Figure 2A**; **Supplementary Figures 5**, **6**; **Supplementary Table 9**). It is worth noting that the S_PH as strongly correlated with the S_GDW ($r \geq 0.5$) and the M_GDW was strongly correlated with the M_Y ($r \geq 0.6$), whereas the M_PH was weakly correlated with the M_GDW and M_Y ($r \geq 0.16$) under control, low, and high salt-alkali stress conditions. Moreover, the strong positive correlations among control, low, and high salt-alkali stress conditions were observed for almost all the traits, and their correlations were from 0.497 to 0.869 while the correlation of all traits between seedling and mature stages was found to be relatively low. These results indicated that the genetic mechanism of salt-alkali stress response at the two different developmental stages was likely to be different.

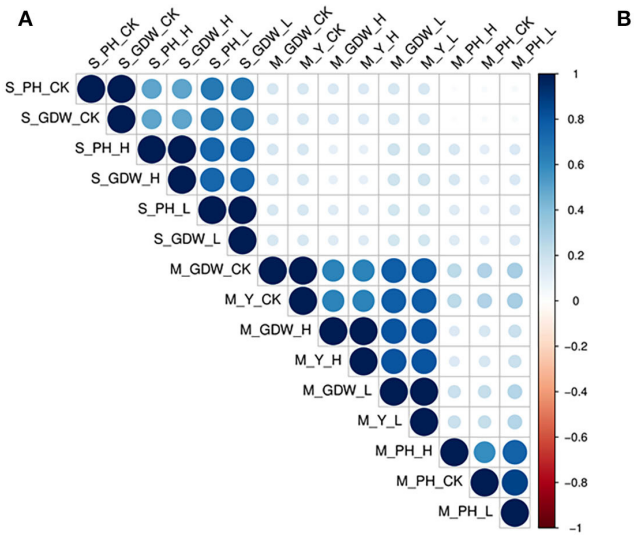
Heritability, Genetic, and Treatment Effect Analysis

Heritability is a key factor of decision-making in crop breeding, and it is essential for the effective design of breeding programs (Chen et al., 2014). Therefore, we also calculated the broad heritability (H_2b) of all traits at seedling and mature stages. As shown in **Supplementary Table 10**, the variations in 505 *B. napus* accessions were largely determined by genotype and heritability estimates, which ranged from 0.759 to 0.889 (**Figure 2B**; **Supplementary Table 10**). Moreover, the genetic effect and treatment effect (G_effect and E_effect, $P \leq 0.05$) could significantly represent the genotype effect and environment effect using the AVOVA by R environment (**Figure 2B**; **Supplementary Table 10**). Except for the S_GDW, the G&E_effects of other traits were found to be extremely significant. These results indicated that the response to salt-alkali stress is dynamic, which is largely determined by complex multiple genetic effects, and the mechanism of adaption to salt-alkali stress in *B. napus* is complicated.

GWAS Revealed the QTLs Related to Salt-Alkali Tolerance in *B. napus* and Prediction of Candidate Genes

The 15 absolute value and 10 TCs (**Supplementary Tables 4, 5**) were performed using the GEMMA linear mixed models for GWAS of all traits (**Figures 3A,B**; **Supplementary Figures 7A–C**) (Zhou and Stephens, 2012). About 3,536 (∼0.046% of in total) strongly associated SNPs ($-\log_{10}$ (P-value) ≥ 6.) were chosen to further map the QTLs with the significant effect SNPs (**Supplementary Table 11**). A total of 1,170 SNPs (CK_SNPs), 1,207 SNPs (L_SNPs), and 762 SNPs (H_SNPs) were identified for absolute value of all traits under control, low salt-alkali, and high salt-alkali conditions, respectively. Moreover, 367 SNPs (R1_SNPs) and 30 SNPs (R2_SNPs) were identified for TCs under low salt-alkali and high salt-alkali conditions, respectively (**Figure 3C**; **Supplementary Table 12**). Based on these results, we obtained different numbers of specific and co-localized SNPs for different traits under different salt-alkali conditions. Among these SNPs, 740 specialized SNPs were detected, including 169 CK_SNPs, 136 L_SNPs, 49 H_SNPs, 364 R1_SNPs, and 22 R2_SNPs, respectively. About 366 co-localized SNPs (CK_SNPs and L_SNPs) and 9 co-localized SNPs (H_SNPs and R2_SNPs) were located on Chromosome A03, respectively. We also identified 70 (L_SNPs and H_SNPs), 632 SNPs (CK_SNPs, L_SNPs, and H_SNPs) and 3 co-localized SNPs (CK_SNPs, L_SNPs, H_SNPs, and R1_SNPs) located on Chromosome A10, respectively. Additionally, 20 co-localized SNPs on Chromosome A10 were simultaneously identified by M_GDW_L, M_PH_CK, M_PH_H, M_PH_L, and M_Y_L; 8 co-localized SNPs on Chromosome A03 were simultaneously identified by M_GDW_H, M_GDW_R2, M_Y_H, and M_Y_R2 (**Figure 3C**; **Suppleentary Table 12**). We finally obtained 9 main effect and co-localized QTLs related to salt-alkali stress response (**Supplementary Table 13**).

In order to identify the candidate genes associated with different traits under different salt-alkali stress conditions, we finally focused on the SNPs identified by GWAS for TCs of all traits, such as R1_SNPs (**Figure 3C**; **Supplementary Tables 12**, **14**), R2_SNPs (**Figure 3C**; **Supplementary Tables 12**, **15**), H_SNPs & R2_SNPs (**Figure 3C**; **Supplementary Tables 12**, **16**), and CK_SNPs & L_SNPs & H_SNPs & R1_SNPs (**Figure 3C**; **Supplementary Tables 12**, **17**). By analyzing the differential gene expression between control and stress conditions ($R \geq 2$ or ≤ 0.5) (https://bigd.big.ac.cn/) (Zhang et al., 2019), 222 genes were identified within 200 kb upstream or downstream of the lead SNPs (**Supplementary Table 18**), such as 36 genes located on ChrA03 from 10,625,336 to 10,741,817 bp for M_Y_R1 and from 10,625,336 to 10,741,817 bp for M_GDW_R1 (**Supplementary Figures 8A–C**; **Supplementary Table 14**) and 15 candidate genes located on ChrA03 from 8,751,415 to 9,293,723 bp for M_Y_R2 and from 9,292,526 to 9,293,723



B

	G_effect (Pr)	E_effect (Pr)	G&E_effect (Pr)	*H2b*
S_GDW	1.82E-239***	0***	0.023*	0.828
S_PH	0***	0***	7.85E-48***	0.873
M_PH	0***	0***	0***	0.889
M_GDW	0***	0***	3.819E-270***	0.770
M_Y	1.38E-28***	0***	1	0.759

FIGURE 2 | Evaluation of correlation coefficient, heritability, and genetic and treatment effects of all traits under salt-alkali stress conditions at seedling and mature stages. **(A)** Heatmap showing Pearson's correlations (*r*) among all traits of the 505 *B. napus* accessions assessed under control (CK, control condition, 0.1–0.25% salt content and pH 7.5–8.), low salt-alkali (L, low salt-alkali condition, 0.35–0.53% salt content, and pH 8–8.5), and high salt-alkali conditions (H, high salt-alkali condition, 0.64–1.05% salt content and pH 8.3–9.). S_GDW, ground dry weight and S_PH, plant height at the seedling stage. M_PH, plant height; M_GDW, ground dry weight; and M_Y, yield at the mature stage. **(B)** Broad-sense heritability, genetic and treatment effect analyses. *H2b*, broad-sense heritability; E_effect, treatment effect; G_effect, genetic effect; and G&E_effect, interaction of genetic and treatment effects. Statistical significance was evaluated by Student *T*-test. $^{*}p < 0.05$ and $^{***}p < 0.001$.

bp for M_GDW_R2 (**Supplementary Figures 9A–C**). GO analysis revealed that most genes were associated with the cellular process, the metabolic process, biological regulation, and response to stimulus (**Supplementary Figure 10**). Moreover, we also found that the most of these candidate genes were significantly differentially expressed under salt stress according to the RNA-seq data (PRJCA008229, https://ngdc.cncb.ac.cn/omix/) (Zhang et al., 2022). The relative expression of some candidate genes under normal and salt conditions was taken as an example (**Figure 3D**; **Supplementary Table 19**), which was consistent with the previous study (Zhang et al., 2019), showing the accuracy of candidate gene prediction. These results suggest that the accuracy of candidate gene prediction and the tolerance of *B. napus* to salt-alkali stress are regulated by a complex genetic mechanism.

Prediction of Candidate Genes of Salt-Alkali Tolerance Within the Main Effect QTL *QSAT.A02.1* on ChrA02

A total of 36 genes within the QTL region on ChrA02 from 6,927,771 to 7,018,686 bp, which was named as to be *qSAT.A02.1* within 200 kb upstream and downstream of the candidate SNPs of M_PH_R1 (**Figures 4A,B**; **Supplementary Tables 14**, **18**), and the expression of some genes were found to be significantly induced by many abiotic stresses (https://bigd.big.ac.cn/) (Zhang et al., 2019). Among these genes, we finally focused on the BnaA02g12820D, whose functions related to salt-alkali stress response have not been reported. BnaA02g12820D named as *BnABA4* is a homologous gene of *Arabidopsis ABA4*, which is involved in the photoprotection of PSII and plays a vital role in neoxanthin production in abscisic acid biosynthesis. Based on our RNA-seq data, we found that *BnABA4* was significantly induced by salt stress (**Figure 3D**). Moreover, the results from the expression pattern showed that *BnABA4* was highly expressed in siliques and stems (**Figure 4C**).

BnABA4 consisting of five introns and six exons shows rich SNP variations, leading to synonymous and nonsense variation (**Supplementary Table 20**). The haplotypes of *BnABA4* were grouped into 6 haplotypes, including hap.A, hap.B, hap.C, hap.D, hap.E, and hap.F types (**Supplementary Table 21**). Both the absolute and relative values of plant height (M_PH) of hap.B type at the mature stage were significantly higher than other haplotypes under the control, low or high salt-alkali stress condition. We thus considered the accessions with hap.B types of *BnABA4* as extremely resistant haplotypes under salt-alkali stress conditions (**Figures 4D–H**). On the contrary, both the absolute and relative values of M_PH of hap.D type at the mature stage were significantly less than other haplotypes under control, low or high salt-alkali stress condition (**Figures 4D–H**). We thus considered the accessions with the hap.D types of *BnABA4* as extremely sensitive haplotypes under salt-alkali stress conditions. Therefore, *BnABA4* was considered as a candidate gene on the QTL *qSAT.A02.1* for salt-alkali stress response, showing a rich

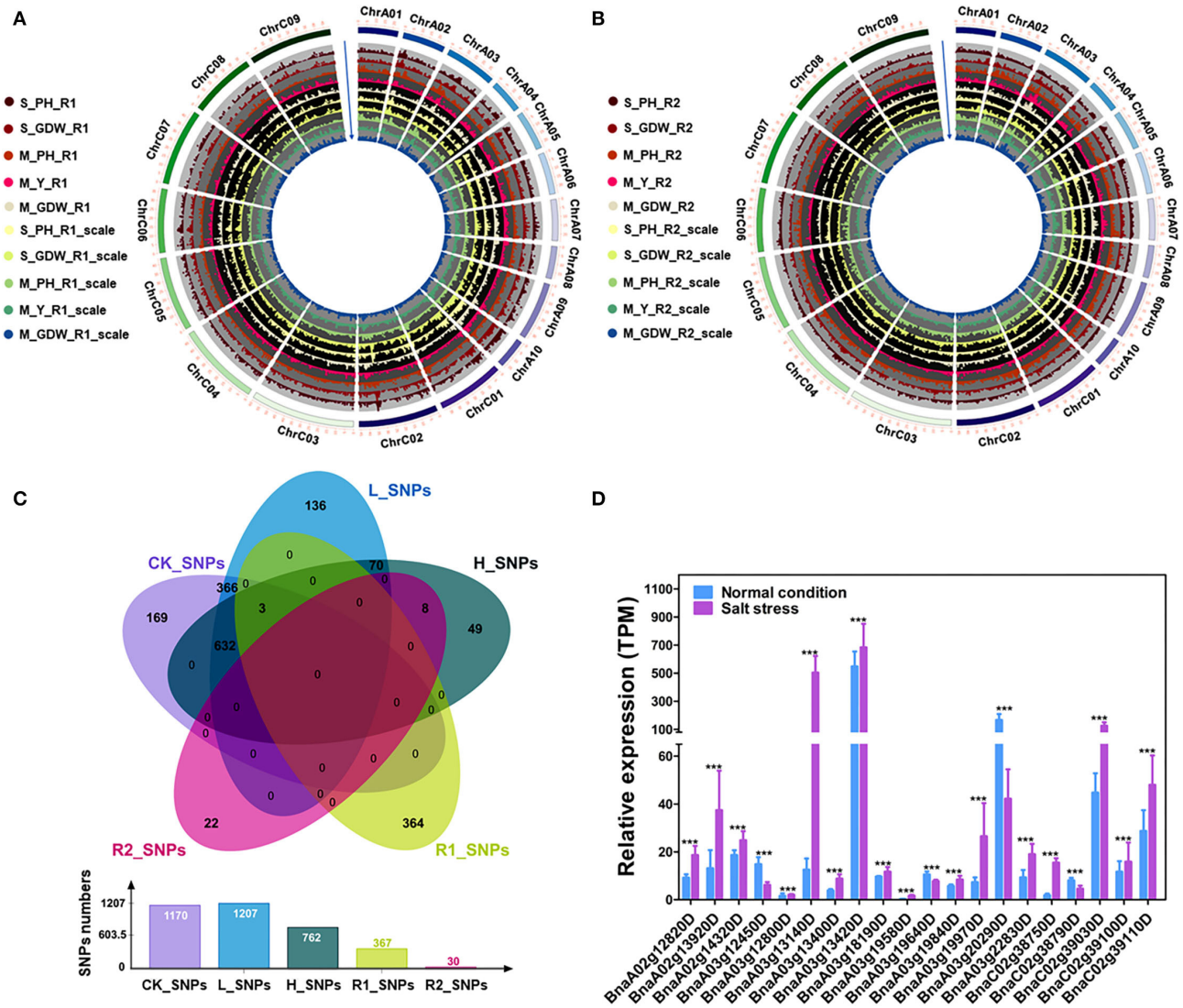


FIGURE 3 | GWAS for tolerance coefficients (TCs) and identification of candidate genes. **(A)** Circle Manhattan plots with the co-localized SNPs for TCs of multiple traits under low salt-alkali condition ($P \leq$ 1e-2). S_PH_R1, the ratio of plant height between low salt-alkali and control conditions at the seedling stage; S_GDW_R1, the ratio of ground dry weight between low salt-alkali and control conditions at the seedling stage; M_PH_R1, the ratio of plant height between low salt-alkali and control conditions at the mature stage; M_Y_R1, the ratio of yield between low salt-alkali and control conditions at the mature stage; M_GDW_R1, the ratio of ground dry weight between low salt-alkali and slight control conditions at the mature stage; S_PH_R1_scale, normalized value of the plant height ratio between low salt-alkali and control conditions in the seedling stage; S_GDW_R1_scale, normalized value of the ground dry weight ratio between low salt-alkali and control conditions; M_PH_R1_scale, normalized value of the plant height ratio between low salt-alkali and control conditions at the mature stage; M_Y_R1_scale, normalized value of the yield ratio between low salt-alkali and control conditions at the mature stage; M_GDW_R1_scale, normalized value of the ground dry weight ratio between low salt-alkali and control conditions at mature the stage. **(B)** Circle Manhattan plots with the co-localized SNPs for TCs of multiple traits under the high-salt-alkali condition ($P \leq$ 1e-2). S_PH_R2, the ratio of plant height between high salt-alkali and control conditions at the seedling stage; S_GDW_R2, the ratio of ground dry weight between high-salt-alkali and control conditions at the seedling stage; M_PH_R2, the ratio of plant height between high-salt-alkali and control conditions at the mature stage; M_Y_R2, the ratio of yield between high-salt-alkali and control conditions at the mature stage; M_GDW_R2, the ratio of ground dry weight between high-salt-alkali and control conditions at the mature stage; S_PH_R2_scale, the ratio of normalized value of plant height between high-salt-alkali and control conditions at the seedling stage; S_GDW_R2_scale, the ratio of normalized value of ground dry weight between high-salt-alkali and control conditions at the seedling stage; M_PH_R2_scale, the ratio of normalized value of plant height between high-salt-alkali and control conditions at the mature stage; M_Y_R2_scale, the ratio of normalized value of yield between high-salt-alkali and control conditions at the mature stage; M_GDW_R2_scale, the ratio of normalized value of ground dry weight between high-salt-alkali and control conditions at the mature stage. **(C)** Co-localized SNPs under different environmental conditions revealed by Venn analysis. CK_SNPs, candidate SNPs of all traits under the control condition; L_SNPs, candidate SNPs of all traits under the low-salt-alkali condition; H_SNPs, candidate SNPs of all traits under high-salt-alkali conditions; R1_SNPs, candidate SNPs of all TCs under the low-salt-alkali condition; R2_SNPs, candidate SNPs of all TCs under the high-salt-alkali condition. The bar graph represents the number of SNPs under each environmental condition. **(D)** Gene expression of candidate differential expressed genes (DEGs) under salt stress condition. Statistical significance was determined by Student's *T*-test. ***$p < 0.001$.

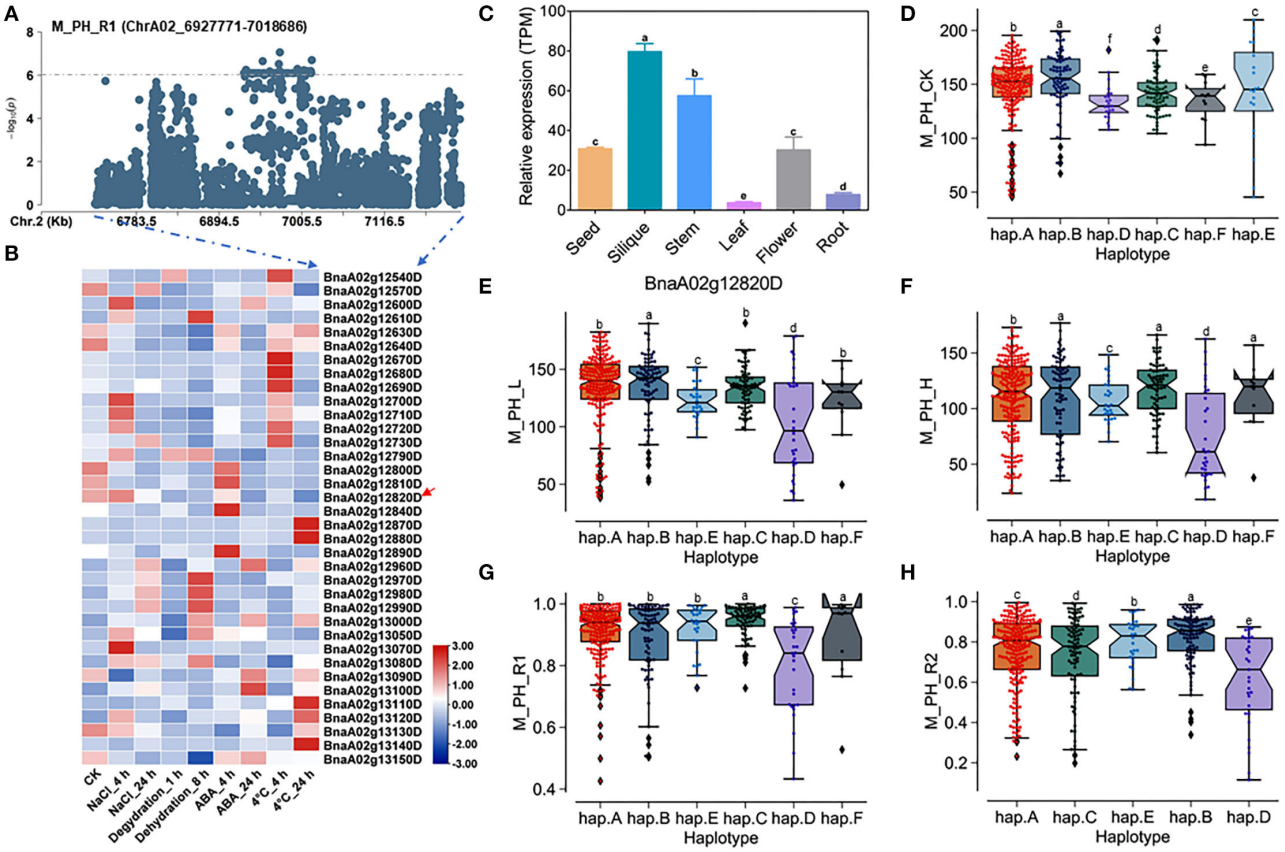


FIGURE 4 | Analysis of candidate genes within the QTL *qSAT.A02.1* on ChrA02 associated with M_PH_R1. **(A)** A zoom-in view of the Manhattan plot on ChrA02 from 6,727,771 to 7,218,686 bp in the QTL *qSAT.A02.1* region. These points represented the log-transformed *P*-values from GWAS for the ratio of plant height between low-salt-alkali and control conditions at the mature stage (M_PH_R1). **(B)** Gene expression of candidate genes identified within 200 kb upstream or downstream of the significantly associated SNPs under 200-mM NaCl, 25-μM ABA, 4°C low temperature conditions for 4 and 24 h and under the dehydration condition for 1 and 8 h in *B. napus* cultivar Zhong Shuang 11 ("*ZS11*"). **(C)** Expression levels of BnA02g12820D in different tissues of *B. napus* cultivar "*ZS11*." Values were means ± SD ($n = 3$ replicates), and different letters significantly indicate differences at $p < 0.05$ using two-way ANOVA. **(D–F)** Boxplots of haplotypes of BnA02g12820D for absolute value of the plant height (M_PH) at the mature stage under control conditions (M_PH_CK, **D**), the low-salt-alkali condition (M_PH_L, **E**), and under the high salt-alkali condition (M_PH_H, **F**). Values were means ± SD ($n = 5$ replicates), and different letters indicate significant difference at $p < 0.05$ using two-way ANOVA. **(G,H)** Boxplots of haplotypes of BnA02g12820D for relative value of the plant height (M_PH) at the mature stage under the low-salt-alkali condition (M_PH_R1, **G**) and under the high-salt-alkali condition (M_PH_R2, **H**). Values were means ± SD ($n = 5$ replicates), and different letters significantly indicate differences at $p < 0.05$ using two-way ANOVA.

SNP variation in 505 *B. napus* accessions. The identified extreme haplotypes related to salt-alkali tolerance would be useful for breeding salt-alkali-tolerant *B. napus* in the future.

Prediction of Candidate Genes of Salt-Alkali Tolerance Within the Main Effect QTL *QSAT.A02.2* on ChrA02

There were 22 genes within the QTL region on ChrA02 from 7,918,159 to 8,015,641 bp, which was named as to be *qSAT.A02.2* within 200 kb upstream and downstream of the candidate SNPs of M_PH_R1 (**Figures 5A,B**; **Supplementary Tables 14**, **18**), and the expressions of some genes were found to be significantly induced by many stresses (https://bigd.big.ac.cn/) (Zhang et al., 2019). Among these genes, we finally focused on the BnaA02g13920D and BnaA02g14320D, whose functions related to salt-alkali stress have not been reported. BnaA02g13920D and BnaA02g14320D were named as *BnBBX14* and *BnVTI12,* which is the area homologous gene of *Arabidopsis BBX14* and *VTI12*, respectively. *BnBBX14* encodes a transcription factor of B-box zinc finger type with a CCT domain, which normally localizes to the transgolgi network. However, the *BnVTI12* belongs to the *VTI* gene family, which normally localizes to plasma membrane. Based on our RNA-seq data, we also found that *BnBBX14* and *BnVTI12* were significantly induced by salt stress (**Figure 3D**; **Supplementary Table 19**). Additionally, *BnBBX14* and *BnVTI12* were found to be expressed in various tissues. The expression level of *BnBBX14* was relatively high (**Figure 5C**) in siliques and stems. However, the expression level of *BnVTI12* in flowers was higher than other tissues (**Supplementary Figure 11A**).

BnBBX14 consisting of one intron and two exons shows rich SNP sequence variations, leading to synonymous, missense, and

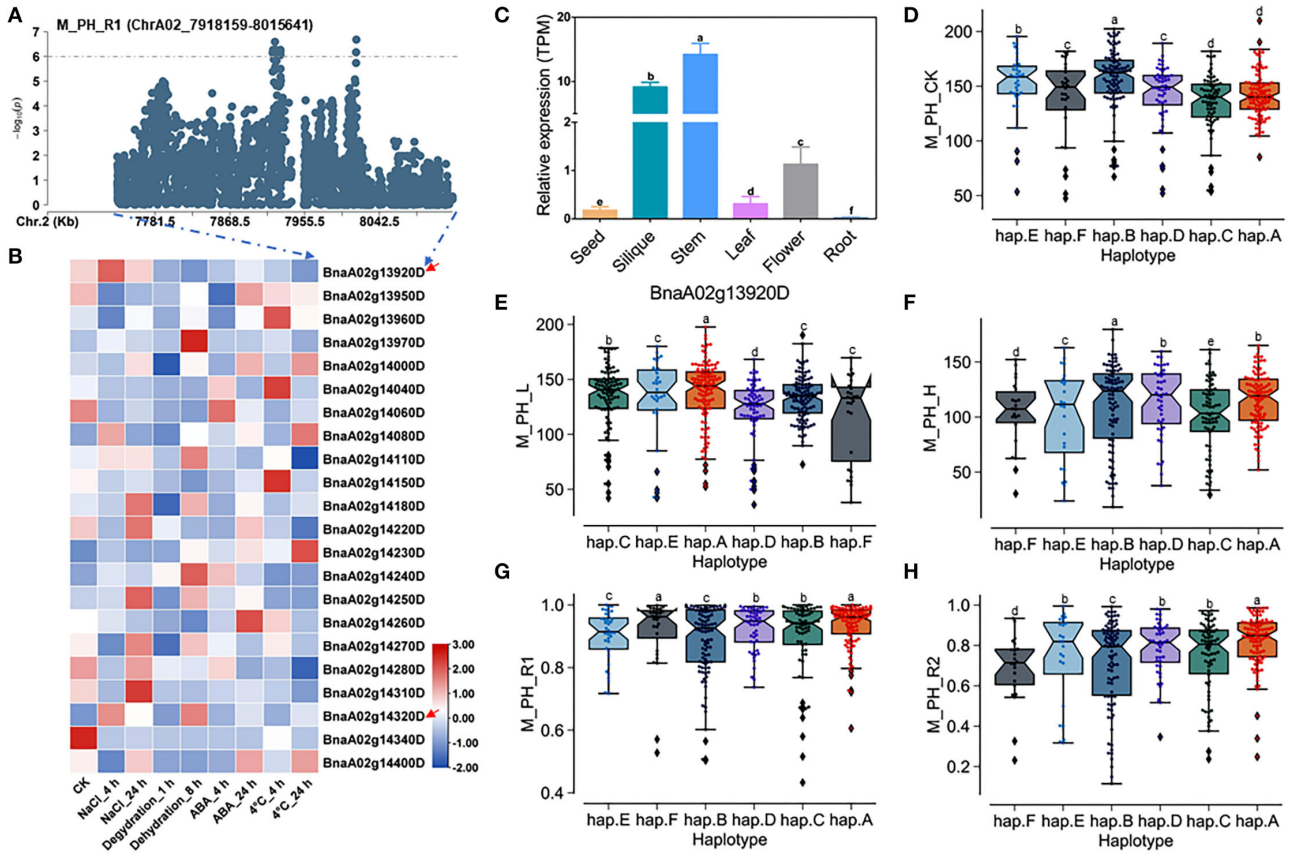


FIGURE 5 | Analysis of candidate genes within the QTL *qSAT.A02.2* on ChrA02 associated with M_PH_R1. **(A)** A zoom-in view of the Manhattan plot on ChrA02 from 7,718,159 to 8,215,641 bp in the QTL *qSAT.A02.2* region. These points represented the log-transformed *P*-values for variants from GWAS for the ratio of plant height between low-salt-alkali and control conditions at the mature stage (M_PH_R1). **(B)** Gene expression of candidate genes identified within 200 kb upstream or downstream of the significantly associated SNPs under 200-mM NaCl, 25-μM ABA, 4°C low temperature conditions for 4 and 24 h and the dehydration condition for 1 and 8 h in *B. naps* cultivar "*ZS11*." **(C)** Expression levels of BnA02g13920D in different tissues of *B. napus* cultivar "*ZS11*." Values were means ± SD ($n = 3$ replicates), and different letters significantly indicate differences at $p < 0.05$ using two-way ANOVA. **(D–F)** Boxplots of haplotypes of BnA02g13920D for absolute value of the plant height (M_PH) at the mature stage under control conditions (M_PH_CK, **D**), the low-salt-alkali condition (M_PH_L, **E**), and under the high salt-alkali condition (M_PH_H, **F**). Values were means ± SD ($n = 5$ replicates), and different letters significantly indicate difference at $p < 0.05$ using two-way ANOVA. **(G,H)** Boxplots of haplotypes of BnA02g13920D for relative value of the plant height (M_PH) at the mature stage under the low-salt-alkali condition (M_PH_R1, **G**) and under the high-salt-alkali condition (M_PH_R2, **H**). Values were means ± SD ($n = 5$ replicates), and different letters significantly indicate differences at $p < 0.05$ using two-way ANOVA.

disruptive inframe deletion variation (**Supplementary Table 22**). The haplotypes of *BnBBX14* were grouped into 6 haplotypes, including hap.A, hap.B, hap.C, hap.D, hap.E, and hap.F types (**Supplementary Table 23**). Except for hap.A types accessions of plant height (M_PH) under the control condition, both the absolute and relative values of M_PH of the hap.A type at the mature stage were significantly higher than other haplotypes under the control, low or high salt-alkali stress condition (**Figures 5D–H**). We thus considered the accessions with hap.A type of *BnBBX14* as extremely resistant materials under salt-alkali stress conditions. *BnVTI12* consisting of four introns and five exons exhibits rich SNP variations, leading to synonymous and splice region variants (**Supplementary Table 24**). The haplotypes of *BnVTI12* were grouped into 5 haplotypes, including hap.A, hap.B, hap.C, hap.D, and hap.E types (**Supplementary Table 25**). Both the absolute and relative values of plant height (M_PH) of the hap.E type at the mature stage were significantly higher than other haplotypes. We thus considered the accessions with hap.E type of *BnVTI12* as extremely resistant haplotypes under salt-alkali stress conditions. Except for M_PH trait absolute value of hap.A types accessions under control, low, and high salt-alkali stress condition, the relative values of hap.A type under low and high salt-alkali stress conditions were significantly higher than other haplotypes. We thus considered the accessions with hap.A type of *BnVTI12* as extremely resistant haplotypes under salt-alkali stress conditions. Therefore, *BnBBX14* and *BnVTI12* were considered as two candidate genes on *qSAT.A02.2* for salt-alkali stress response, showing a rich SNP variation among 505 *B. napus* accessions. The identified extreme haplotypes related to salt-alkali tolerance would be useful for breeding salt-alkali-tolerant *B. napus* in the future.

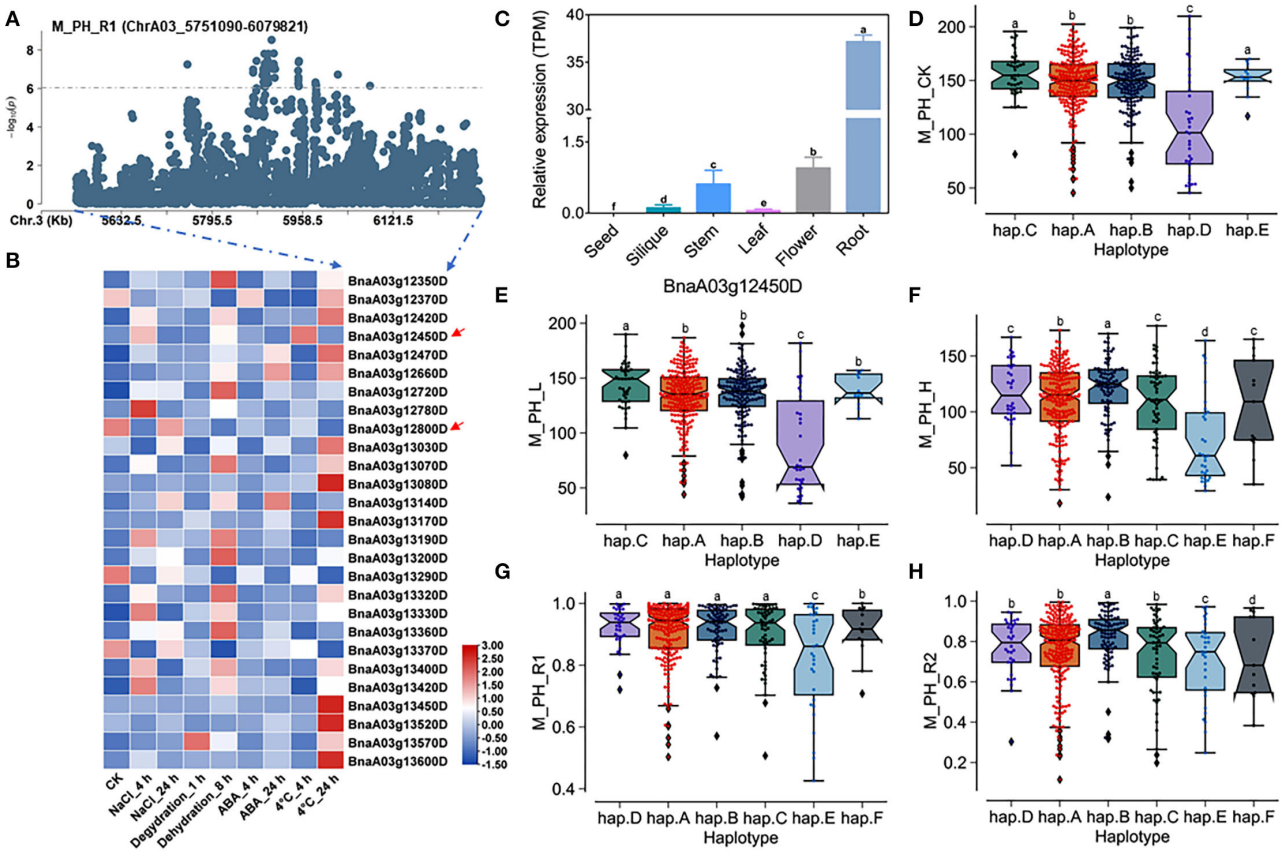


FIGURE 6 | Analysis of candidate genes within the QTL *qSAT.A03* on ChrA03 associated with M_PH_R1. **(A)** A zoom-in view of the Manhattan plot on ChrA03 from 5,751,090 to 6,079,821 bp in the QTL *qSAT.A03* region. These points represented the log-transformed *P*-values for variants from GWAS of the ratio of plant height between low-salt-alkali and control conditions at the mature stage (M_PH_R1). **(B)** Gene expression of candidate genes identified within 200 kb upstream or downstream of the significantly associated SNPs under 200-mM NaCl, 25-μM ABA, 4°C low temperature conditions for 4 and 24-h dehydration conditions for 1 and 8 h in *B. napus* cultivar "*ZS11*." **(C)** Expression levels of BnA03g12450D in different tissues of *B. napus* cultivar "*ZS11*." Values were means ± SD ($n = 3$ replicates), and different letters significantly indicate differences at $p < 0.05$ using two-way ANOVA. **(D–F)** Boxplots of haplotypes of BnA03g12450D for absolute value of the plant height (M_PH) at the mature stage under control conditions (M_PH_CK, **D**), the low-salt-alkali condition (M_PH_L, **E**), and under the high-salt-alkali condition (M_PH_H, **F**). Values were means ± SD ($n = 5$ replicates), and different letters significantly indicate difference at $p < 0.05$ using two-way ANOVA. **(G,H)** Boxplots of haplotypes of BnA03g12450D for relative value of the plant height (M_PH) at the mature stage under the low-salt-alkali condition (M_PH_R1, **G**) and under the high-salt-alkali condition (M_PH_R2, **H**). Values were means ± SD ($n = 5$ replicates), and different letters significantly indicate differences at $p < 0.05$ using two-way ANOVA.

Prediction of Candidate Genes of Salt-Alkali Tolerance Within the Main Effect of QTL *QSAT.A03.1* on ChrA03

A total of 27 genes within the QTL region were identified on ChrA03 from 5,751,090 to 6,079,821 bp, which was named as to be *qSAT.A03.1* within 200 kb upstream and downstream of the candidate SNPs of M_PH_R1 (**Figures 6A,B**; **Supplementary Table 14**), and the expressions of some genes was found to be significantly induced by many stresses (https://bigd.big.ac.cn/) (Zhang et al., 2019). Among these genes, we finally focused on the BnaA03g12450D and BnaA03g12800D, whose functions related to salt-alkali stress response have not been reported. BnaA03g12450D and BnaA03g12800D were named as *BnPYL8* and *BnCRR1,* which are the homologous gene of *Arabidopsis PYL8/RCAR3* and *CRR1*, respectively. *BnPYL8* encoding an ABA receptor interacts with protein phosphatase 2Cs *ABI1* and *ABI2* in ABA signaling, and *BnCRR1* is involved in electron transfer between PSI and PSII and plays a vital role in the biogenesis or stabilization of the NAD(P)H dehydrogenase. Based on our RNA-seq data, we also found that *BnPYL8* and *BnCRR1* were significantly induced by salt stress (**Figure 3D**; **Supplementary Table 19**). In addition, *BnPYL8* and *BnCRR1* were found to be expressed in various tissues. The expression level of *BnPYL8* was relatively high in roots (**Figure 6C**), whereas the expression level of *BnCRR1* in the silique was higher than that in other tissues (**Supplementary Figure 12A**).

BnPYL8 consisting of three introns and four exons exhibits rich SNP variations, leading to synonymous, missense, and splice region variation (**Supplementary Table 26**). The haplotypes of *BnPYL8* were grouped into 6 haplotypes, including hap.A, hap.B, hap.C, hap.D, hap.E, and hap.F types (**Supplementary Table 27**).

Both the absolute and relative values of plant height (M_PH) at the mature stage in the hap.A, hap.B, and hap.C type accessions were significantly higher than other haplotypes, while those in the hap.E type accessions were significantly less than other haplotypes under the control, low or high salt-alkali stress condition. Both the absolute and relative values of plant height (M_PH) at the mature stage in the hap.A, hap.B, and hap.C type accessions were significantly higher than other haplotypes, while those in the hap.E type accessions were significantly less than other haplotypes under the control, low or high salt-alkali stress condition (**Figures 6D–H**). We could consider the accessions with hap.A, hap.B, and hap.C types of *BnPYL8* as extremely resistant haplotypes and the accessions with hap.E type of *BnPYL8* as extremely sensitive haplotypes under salt-alkali stress conditions. In addition, *BnCRR1* consisting of nine introns and ten exons shows rich SNP variations, leading to synonymous, missense, and splice region variation (**Supplementary Table 28**). The haplotypes of *BnCRR1* were grouped into 6 haplotypes, including hap.A, hap.B, hap.C, hap.D, hap.E, and hap.F types (**Supplementary Table 29**). Except for M_PH trait absolute value of hap.F types accessions under high salt-alkali stress conditions, both the absolute and relative values in the hap.F type accessions were significantly less than other haplotypes under control, low, and high salt-alkali stress conditions (**Supplementary Figures 12B–F**). We thus considered the accessions with hap.F type of *BnCRR1* as extremely sensitive haplotypes under salt-alkali stress conditions. On the contrary, both the absolute and relative values of plant height in the hap.A, hap.D, and hap.E type accessions were significantly higher than other haplotypes under control, low, and high salt-alkali stress conditions. We thus considered the accessions with hap.A, hap.D, and hap.E types of *BnCRR1* as extremely resistant haplotypes under salt-alkali stress conditions. According to the above results, *BnPYL8* and *BnCRR1* were considered as two candidate genes on *qSAT.A03.1* for salt-alkali stress response, which showed a rich SNP variation among 505 *B. napus* accessions. The identified extreme haplotypes associated with salt-alkali tolerance would be helpful for breeding salt-alkali-tolerant *B. napus* in the future.

DISCUSSION

Salt-alkali stress has significant influences on seedling establishment and yield performance (Zhao et al., 2020; Wassan et al., 2021). In this study, we determined the response to salt-alkali stress at seedling and mature stages in 505 accessions of *B. napus*. We totally obtained 2 phenotypic indexes at the seedling stage, including plant height and aboveground dry weight, and 3 growth traits, including plant height, aboveground dry weight, and yield at the mature stage under control, low, and high salt-alkali stress conditions, respectively (**Supplementary Tables 2–4**) (Batool et al., 2015).

We finally have screened out 11 reliable extreme materials at different developmental stages under the two different salt-alkali environments (**Figures 1B,C**; **Supplementary Figures 3**, **4**). These materials may be useful for the genetic improvement of tolerance to salt-alkali stress in *B. napus*. In addition, it was found that the correlation coefficients between the various traits among different stress conditions at the same developmental stage were high, while the correlation coefficients between the same trait at seedling and mature stages were relatively low (**Figure 2A**; **Supplementary Table 9**). We also found that the plant height, aboveground dry weight, and yield are inhibited by salt-alkali stresses through the analysis of genetic effects (G_effect) and treatment effects (E_effect), and their broad-sense heritabilities (*H2b*) could better represent the genotypic effect of all traits, which were >0.75 (**Figure 2B**; **Supplementary Table 10**). These results suggest that the regulatory mechanism of salt-alkali stress responses at the seedling stage is likely to be different from that at the mature stage.

Previous studies have performed QTL mapping to explore the genetic basis of salt or salt-alkali stress responses at the germination or seedling stages (Yong et al., 2015; He et al., 2017; Hou et al., 2017; Wan et al., 2017, 2018; Zhang et al., 2017; Wassan et al., 2021). Although many QTLs and candidate genes related to salt or alkali stress response were identified in *B. napus*, no QTL or candidate gene has been reported to be associated with salt-alkali stress in *B. napus* until now. In this study, we finally chose 9 main-effect QTLs within 1,819 significant SNPs related to salt-alkali stress by GWAS of 15 absolute values and 10 TCs. We used the Venn analysis to identify co-localized QTLs and SNPs between different treatment conditions, developmental periods, and traits associated with salt-alkali stress response (**Supplementary Tables 12, 13**). In addition, we finally focused on the co-localized SNPs associated with the TCs of growth and agronomic traits (**Supplementary Tables 13–17**). Compared with the previous results, the two SNPs Bn-A03-p140441361 (ChrA03, site: 48,480 bp) and Bn-A06-p1015217 (ChrA06, site: 1,004,315 bp) previously identified by the GWAS for salt tolerance index of the germination rate (ST-GR) were co-localized with the QTL *qSAT.A03* on Chromosome A03 and *qSAT.A06* on Chromosome A06 according to our results, respectively. Moreover, the SNP Bn-A09-P1603989 (ChrA09, site: 2,323,836 bp) previously identified by the GWAS for salt tolerance index of fresh weight was co-localized with the QTL *qSAT.A09* on Chromosome A09 according to our results (Wan et al., 2018). Combining with expression analysis under various stress conditions, 222 candidate genes were identified (**Supplementary Table 18**) within 200 kb upstream and downstream of the candidate SNPs. Among these candidate genes, BnaA01g31420D (*BnBZIP28*), BnaA03g13820D (*BnWRKY25*), BnaA03g18900D (*BnCZF1*), BnaA03g18910D (*BnCZF1*), BnaA03g19270D (*BnCYSA*), BnaC02g39600D (*BnNHX1*), and BnaC02g39630D (*BnNHX1*) previously identified for root length, BnaA03g14410D (*BnSTH*) previously identified for the germination rate (He et al., 2017), BnaC03g22990D (*BnaUMAMIT11*) previously identified for hypocotyl length at seedling (HLW) and germination stages (HLN), and BnaA02g14680D (*BnRPK1*), BnaA02g14490D (*BnHVA22*), BnaC03g64030D (*BnCPL1*), and BnaC03g62830D (*BnMYB98*) previously identified for seedling dry weight and fresh weight by QTL Mapping (Hou et al., 2017) in previous studies on salt tolerance were also mapped

in this study (**Supplementary Table 18**). Moreover, some identified candidate genes, such as BnaA02g12820D (*BnABA4*), BnaA02g14040D (*ABI3*), BnaA03g18230D (*BnCPK20*), BnaA03g12450D (*BnABA3*), BnaA03g12470D (*BnPP2C*), BnaA03g13820D (*BnWRKY25*), and BnaC02g38750D (*BnRAB*), have been reported to be related to salt stress responses. Interestingly, some previously unreported salt-stress-related genes were also identified, such as BnaA03g13140D (*BnLTP3*), BnaA03g19840D (*BnCBK3*), BnaA03g19970D (*BnPIF4*), BnaC02g38790D (*BnSAD1*), and BnaC02g39030D (*BnBXL1*) (**Supplementary Tables 18**, **19**). Therefore, the GWAS approach was considered as a reliable and effective method to identify candidate genes related to salt-alkali stress.

Different from previous GWAS results, a larger population and high-density molecular markers were used in this study, which can improve the abundance of genetic variation within the population and the accuracy of candidate gene positioning. Moreover, compared with previous pieces of research on the salt stress response of plants grown in greenhouse at the germination and seedling stages, our study can truly reflect the response of *B. napus* to salt-alkali stress to *B. napus* under low and high salt-alkali land conditions. It was an important reference for the response of other crops to salt-alkali stress or other abiotic stresses through large-scale field trials on salt-alkali land. However, some shortcomings also existed in this study, such as uneven salt content of saline-alkali soil, a lack of more replicates in different years at different saline-alkali lands, and functional verification of candidate genes. Further studies are called for dissecting the complex genetic mechanism of salt-alkali tolerance in *B. napus.*

CONCLUSION

Our results revealed significant natural variation in some growth and agronomic traits under salt-alkali stress. Using the GWAS approach, we not only confirmed the variation in alleles of some known salt tolerance-related genes but also identified some unreported genes or QTLs related to salt-alkali tolerance. An effective method was applied to identify some tolerant and sensitive accessions as valuable breeding materials for genetic improvement of salt-alkali tolerance of *B. napus*. This study has laid the foundation for understanding the complex genetic mechanism of salt-alkali tolerance and large-scale production of *B. napus* on salt-alkali land.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: National Genomics Data Center (http://bigd.big.ac.cn), PRJCA008229.

AUTHOR CONTRIBUTIONS

XY, LG, GZha, and TF designed the research. GZha, YP, JZ, ZT, ChJ, SF, SZ, CuJ, RW, XW, BL, SL, and GZho performed the experiments or analyzed the data. GZha, LG, and XY analyzed the data and wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by Hubei Hongshan Laboratory Research Funding (2021HSZD004) and Higher Education Discipline Innovation Project (B20051).

ACKNOWLEDGMENTS

We appreciate the suggestions and technical support from Professor Lijun Li from Inner Mongolia Agricultural University and Professor Fenghua Zhang from Shihezi University.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.857149/full#supplementary-material

REFERENCES

Batool, N., Noor, T., Ilyas, N., and Shahzad, A. (2015). Salt stress impacts on seed germination and seedling growth of *Brassica napus L. Pure Appl. Biol.* 4, 398–406. doi: 10.19045/bspab.2015.43016

Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A. P., Tang, H., Wang, X., et al. (2014). Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science* 345, 950–953. doi: 10.1126/science.1253435

Chen, D., Neumann, K., Friedel, S., Kilian, B., Chen, M., Altmann, T., et al. (2014). Dissecting the phenotypic components of crop plant growth and drought responses based on high-throughput image analysis. *Plant Cell* 26, 4636. doi: 10.1105/tpc.114.129601

Chen, J., Zhang, H., Tong, J., Liu, C., Ran, J., Tang, J., et al. (2021). Genome-wide association analysis of root length traits in *Brassica napus* at germination stage under sodium carbonate stress. *Euphytica* 217, 197. doi: 10.1007/s10681-021-02928-3

Chen, T., Johnson, R., Chen, S., Lv, H., Zhou, J., and Li, C. (2018). Infection by the fungal endophyte Epichloë bromicola enhances the tolerance of wild barley (*Hordeum brevisubulatum*) to salt and alkali stresses. *Plant Soil.* 428, 353–370. doi: 10.1007/s11104-018-3643-4

Chikelu, A. R., Jain, S. M., Gregorio, G. B., and Zapata, F. J. (2007). Induced mutations for enhancing salinity tolerance in rice. *Biology* 413–454. doi: 10.1007/978-1-4020-5578-2_17

De, R., Bush, W. S., and Moore, J. H. (2014). Bioinformatics challenges in genome-wide association studies (GWAS). *Methods Mol. Biol.* 1168, 63–81. doi: 10.1007/978-1-4939-0847-9_5

Dwivedi, N. (2004). Genotypic difference in salinity tolerance of green gram cultivars. *Plant Sci.* 166, 1135–1142. doi: 10.1016/j.plantsci.2003.11.028

Fang, S., Hou, X., and andLiang, X. (2021). Response mechanisms of plants under saline-alkali stress. *Front. Plant Sci.* 12, 667458. doi: 10.3389/fpls.2021.667458

Guo, R., Yang, Z., Li, F., Yan, C., Zhong, X., Liu, Q., et al. (2015). Comparative metabolic responses and adaptive strategies of wheat (*Triticum aestivum*) to salt and alkali stress. *BMC Plant Biol.* 15, 170. doi: 10.1186/s12870-015-0546-x

Guo, Z., Yang, W., Chang, Y., Ma, X., Tu, H., Xiong, F., et al. (2018). Genome-wide association studies of image traits reveal genetic architecture of drought resistance in rice. *Mol. Plant*. 11, 789–805. doi: 10.1016/j.molp.2018.03.018

He, Y., Wu, D., You, J., and Qian, W. (2017). Genome-wide association analysis of salt tolerance related traits in *Brassica napus* and candidate gene prediction. *Sci. Agricul. Sin.* 50, 1189–1201. doi: 10.3864/j.issn.0578-1752.2017.07.002

Hickey, L. T., N., Hafeez, A., Robinson, H., Jackson, S. A., Leal Bertioli, S. C. M., et al. (2019). Breeding crops to feed 10 billion. *Nat. Biotechnol*. 37, 744–754. doi: 10.1038/s41587-019-0152-9

Hou, L. Y., Wang, T., Jian, H., Wang, J., Jiana, L., and Liu, L. (2017). QTL mapping for seedling dry weight and fresh weight under salt stress and candidate genes analysis in *Brassica napus L. Acta Agr. Sin*. 43, 179–189. doi: 10.3724/SP.J.1006.2017.00179

Huan, Z. (2015). Study on physiological and biochemical responses to salt and alkali stress in two varieties of Eggplant. *Heilongjiang Agricul. Sci*.

Huang, X., Zhao, Y., Wei, X., Li, C., Wang, A., Qiang, Z., et al. (2012). Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat. Genet*. 44, 32–39. doi: 10.1038/ng.1018

Hussain, S., Shaukat, M., Ashraf, M. A. B., Zhu, C. Q., Jin, Q. Y., and Zhang, J. (2019). Salinity stress in arid and semi-arid climates: effects and management in field crops. *Climate Change Agricul*. 1–26. doi: 10.5772/intechopen.87982

Jin, H., Kim, H., Plaha, P., Liu, S., Park, J. I., Piao, Y., et al. (2008). Expression profiling of the genes induced by Na_2CO_3 and NaCl stresses in leaves and roots of *Leymus chinensis*. *Plant Sci*. 175, 784–792. doi: 10.1016/j.plantsci.2008.07.016

Julkowska, M. M., and Testerink, C. (2015). Tuning plant signaling and growth to survive salt. *Trends Plant Sci*. 20, 586–594. doi: 10.1016/j.tplants.2015.06.008

Khan, M. A., and Weber, D. J. (2006). Ecophysiology of high salinity tolerant plants. *Tasks Veg Sci*. 40, 1–24. doi: 10.1007/1-4020-4018-0

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with *Burrows-Wheeler* transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

Li, H., Feng, H., Guo, C., Yang, S., Huang, W., Xiong, X., et al. (2020). High-throughput phenotyping accelerates the dissection of the dynamic genetic architecture of plant growth and yield improvement in rapeseed. *Plant Biotechnol. J*. 18, 2345–2353. doi: 10.1111/pbi.13396

Lin, J., Hua, X., Peng, X., Dong, B., and Yan, X. (2018a). Germination responses of ryegrass (*Annual vs. Perennial*) seed to the Interactive effects of temperature and salt-alkali stress. *Front. Plant Sci*. 9, 1458. doi: 10.3389/fpls.2018.01458

Lin, J., Peng, X., Hua, X., Sun, S., Wang, Y., and Yan, X. (2018b). Effects of arbuscular mycorrhizal fungi on *Leymus chinensis* seedlings under salt alkali stress and nitrogen deposition conditions: from osmotic adjustment and ion balance. *RSC Adv*. 8, 14500–14509. doi: 10.1039/C8RA00721G

Lin, J., Yu, D., Shi, Y., Sheng, H., Li, C., Wang, Y., et al. (2016). Salt-alkali tolerance during germination and establishment of *Leymus chinensis* in the *Songnen Grassland* of China. *Ecol. Eng*. 95, 763–769. doi: 10.1016/j.ecoleng.2016.07.011

Lin, J. S., Mu, C., Wang, Y., Li, Z., and Li, X. (2014). Physiological adaptive mechanisms of *Leymus chinensis* during germination and early seedling stages under saline and alkaline conditions. *J. Animal Plant Sci*. 24, 904–912.

Lin, N. (2005). Study on the environment evolution and the analysis of causes to Land salization and desertification in songnen plain. *Quaternary Sci*. 152, 55–66. doi: 10.1007/BF02042141

Lu, K., Wei, L., Li, X., Wang, Y., Wu, J., Liu, M., et al. (2019). Whole-genome resequencing reveals *Brassica napus* origin and genetic loci involved in its improvement. *Nat. Commun*. 10, 1154. doi: 10.1038/s41467-019-09134-9

Lv, Y., Guo, Z., Li, X., Ye, H., Li, X., and Xiong, L. (2016). New insights into the genetic basis of natural chilling and cold shock tolerance in rice by genome-wide association analysis. *Plant Cell Environ*. 39, 556–570. doi: 10.1111/pce.12635

Munns, R., and Tester, M. A. (2008). Mechanisms of salinity tolerance. *Annu. Rev. Plant Biol*. 59, 651–681. doi: 10.1146/annurev.arplant.59.032607.092911

Pang, Q., Zhang, A., Zang, W., Wei, L., and Yan, X. (2016). Integrated proteomics and metabolomics for dissecting the mechanism of global responses to salt and alkali stress in *Suaeda corniculata*. *Plant Soil*. 402, 379–394. doi: 10.1007/s11104-015-2774-0

Patishtan, J., Hartley, T. N., Raquel, F., and Maathuis, F. J. (2017). Genome-wide association studies to identify rice salt-tolerance markers. *Plant Cell Environ*. 41, 970–982. doi: 10.1111/pce.12975

Pearson, T. A., and Manolio, T. A. (2008). How to interpret a genome-wide association study. *JAMA* 299, 1335–1344. doi: 10.1001/jama.299.11.1335

Sanchez-Bel, P., Egea, I., Flores, F. B., Bolarin, M. C., Sanchez-Bel, P., Egea, I., et al. (2012). Tomato: grafting to improve salt tolerance. *Improv. Crop. Resist. Abiotic Stress* 10, 263. doi: 10.1002/9783527632930.ch42

Sanders, D. (2020). The salinitychallenge. *New Phytol*. 225, 1047–1048. doi: 10.1111/nph.16357

Singh, H., Kumar, P., Kumar, A., Kyriacou, M. C., Colla, G., and Rouphael, Y. (2020). Grafting tomato as a tool to improve salt tolerance. *Agron* 10, 263. doi: 10.3390/agronomy10020263

Song, J., Guan, Z., Hu, J., Guo, C., Yang, Z., Wang, S., et al. (2020). Eight high-quality genomes reveal pan-genome architecture and ecotype differentiation of *Brassica napus*. *Nat Plants* 6, 34–45. doi: 10.1038/s41477-019-0577-7

Song, W., and Liu, M. (2017). Farmland conversion decreases regional and national land quality in China. *Land Degrad Dev*. 28, 459–471. doi: 10.1002/ldr.2518

Tang, S., Zhao, H., Lu, S., Yu, L., Zhang, G., Zhang, Y., et al. (2021). Genome- and transcriptome-wide association studies provide insights into the genetic basis of natural variation of seed oil content in *Brassica napus*. *Mol. Plant*. 14, 470–487. doi: 10.1016/j.molp.2020.12.003

Tyerman, S. D., Munns, R., Fricke, W., Arsova, B., Barkla, B. J., Bose, J., et al. (2019). Energy costs of salinity tolerance in crop plants. *New Phytol*. 221, 25–29. doi: 10.1111/nph.15555

Wan, H., Chen, L., Guo, J., Li, Q., Wen, J., Yi, B., et al. (2017). Genome-wide association study reveals the genetic architecture underlying salt tolerance-related traits in rapeseed (*Brassica napus L.*). *Front. Plant Sci.* 8, 593. doi: 10.3389/fpls.2017.00593

Wan, H., Wei, Y., Qian, J., Gao, Y., Wen, J., Yi, B., et al. (2018). Association Mapping of salt tolerance traits at germination stage of rapeseed (*Brassica napus* L.). *Euphytica* 214. doi: 10.1007/s10681-018-2272-6

Wang, H., Wu, Z., Chen, Y., Yang, C., and Shi, D. (2011). Effects of salt and alkali stresses on growth and ion balance in rice (*Oryza sativa L.*). *Plant Soil Environ*. 57, 286–294. doi: 10.17221/36/2011-PSE

Wassan, G. M., Khanzada, H., Zhou, Q., Mason, A. S., Keerio, A. A., Khanzada, S., et al. (2021). Identification of genetic variation for salt tolerance in *Brassica napus* using genome-wide association mapping. *Mol. Genet. Genomics*. 296, 391–408. doi: 10.1007/s00438-020-01749-8

Xu, Y., Magwanga, R. O., Yang, X., Jin, D., Cai, X., Hou, Y., et al. (2020). Genetic regulatory networks for salt-alkali stress in *Gossypium hirsutum* with differing morphological characteristics. *BMC Genomics* 21, 15. doi: 10.1186/s12864-019-6375-9

Yong, H. Y., Wang, C., Bancroft, I., Li, F., Wu, X., Kitashiba, H., et al. (2015). Identification of a gene controlling variation in the salt tolerance of rapeseed (*Brassica napus L.*). *Planta* 242, 313–326. doi: 10.1007/s00425-015-2310-8

Yong, H. Y., Zou, Z., Kok, E. P., Kwan, B. H., and Nishio, T. (2014). Comparative transcriptome analysis of leaves and roots in response to sudden increase in salinity in *Brassica napus* by RNA-seq. *J. Biomed. Biotechnol*. 2014, 467395. doi: 10.1155/2014/467395

Yu, J., Zao, W., He, Q., Kim, T. S., and Park, Y. J. (2017). Genome-wide association study and gene set analysis for understanding candidate genes involved in salt tolerance at the rice seedling stage. *Mol. Genet. Genomics* 292, 1391–1403. doi: 10.1007/s00438-017-1354-9

Zhang, G., Zhou, J., Peng, Y., Tan, Z., Li, L., Yu, L., et al. (2022). Genome-wide association studies of salt tolerance at seed germination and seedling stages in *Brassica napus*. *Front. Plant Sci.* 12, 772708. doi: 10.3389/fpls.2021.772708

Zhang, H. X., Hodson, J.oanna, N., Williams, John, P., and Blumwald, E. (2001). Engineering salt-tolerant Brassica plants: characterization of yield and seed oil quality in transgenic plants with increased vacuolar sodium accumulation. *Proc. Natl. Acad. Sci. U.S.A.* 98, 12832–12836. doi: 10.1073/pnas.231476498

Zhang, R., Deng, W., Yang, L., Wang, Y., and Lu, K. (2017). Genome-wide association study of root length and hypocotyl length at germination stage under saline conditions in *Brassica napus*. *Sci. Agricul. Sin.* 50, 15–27. doi: 10.3864/j.issn.0578-1752.2017.01.002

Zhang, Y., Ali, U., Zhang, G., Yu, L., Fang, S., Iqbal, S., et al. (2019). Transcriptome analysis reveals genes commonly responding to multiple abiotic stresses in rapeseed. *Mol. Breed*. 39, 158. doi: 10.1007/s11032-019-1052-x

Zhao, B. Y., Hu, Y. F., Li, J., Yao, X., and Liu, K. (2016). *BnaABF2,* a bZIP transcription factor from rapeseed (*Brassica napus L.*), enhances drought and salt tolerance in transgenic *Arabidopsis*. *Bot. Stud.* 57, 12. doi: 10.1186/s40529-016-0127-9

Zhao, C., Zhang, H., Song, C., Zhu, J. K., and Shabala, S. (2020). Mechanisms of plant responses and adaptation to soil salinity. *Innovation* 1, 100017. doi: 10.1016/j.xinn.2020.100017

Zhou, C., and Yang, Y. (2003). Physiological response to salt-alkali stress in experimental populations in two ecotypes of *Leymus chinensis* in the Songnen Plains of China. *Ying Yong Sheng Tai Xue Bao.* 14, 1842–1846.

Zhou, X., and Stephens, M. (2012). Genome-wide efficient mixed model analysis for association studies. *Nat. Genet*. 44, 821–824. doi: 10.1038/ng.2310

Zhu, J. K. (2001). Cell signaling under salt, water and cold stresses. *Curr. Opin. Plant Biol*. 4, 401–406. doi: 10.1016/S1369-5266(00)00192-8

Zhu, J. K. (2002). Salt and drought stress signal transduction in plants. *Annu. Rev. Plant Biol*. 53, 247–273. doi: 10.1146/annurev.arplant.53.091401.143329

Zhu, J. K. (2016). Abiotic stress signaling and responses in plants. *Cell* 167, 313–324. doi: 10.1016/j.cell.2016.08.029

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 29 April 2022
doi: 10.3389/fpls.2022.863283

CBL-Interacting Protein Kinase OsCIPK18 Regulates the Response of Ammonium Toxicity in Rice Roots

Tong Sun[1], Ting Wang[2], Yalin Qiang[1], Gangqing Zhao[1], Jian Yang[1], Hua Zhong[1], Xiaojue Peng[3], Jing Yang[1,3]* and Yangsheng Li[1]**

[1] State Key Laboratory of Hybrid Rice, College of Life Sciences, Wuhan University, Wuhan, China, [2] Department of Chemistry, University of Kentucky, Lexington, KY, United States, [3] College of Life Sciences, Nanchang University, Nanchang, China

Ammonium (NH_4^+) is one of the major nitrogen sources for plants. However, excessive ammonium can cause serious harm to the growth and development of plants, i.e., ammonium toxicity. The primary regulatory mechanisms behind ammonium toxicity are still poorly characterized. In this study, we showed that OsCIPK18, a CBL-interacting protein kinase, plays an important role in response to ammonium toxicity by comparative analysis of the physiological and whole transcriptome of the T-DNA insertion mutant (*cipk18*) and the wild-type (WT). Root biomass and length of *cipk18* are less inhibited by excess NH_4^+ compared with WT, indicating increased resistance to ammonium toxicity. Transcriptome analysis reveals that OsCIPK18 affects the NH_4^+ uptake by regulating the expression of OsAMT1;2 and other NH_4^+ transporters, but does not affect ammonium assimilation. Differentially expressed genes induced by excess NH_4^+ in WT and *cipk18* were associated with functions, such as ion transport, metabolism, cell wall formation, and phytohormones signaling, suggesting a fundamental role for OsCIPK18 in ammonium toxicity. We further identified a transcriptional regulatory network downstream of OsCIPK18 under NH_4^+ stress that is centered on several core transcription factors. Moreover, OsCIPK18 might function as a transmitter in the auxin and abscisic acid (ABA) signaling pathways affected by excess ammonium. These data allowed us to define an OsCIPK18-regulated/dependent transcriptomic network for the response of ammonium toxicity and provide new insights into the mechanisms underlying ammonium toxicity.

Keywords: ammonium toxicity, NH_4^+, OsCIPK18, rice (*Oryza sativa* L.), RNA-Seq

OPEN ACCESS

Edited by:
Rajeev K. Varshney, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), India

Reviewed by:
Vikranth Kumar Chandrasekaran, University of Missouri, United States
Sibaji Kumar Sanyal, Jawaharlal Nehru University, India
Yunus Effendi, Leibniz University Hannover, Germany

****Correspondence:***
Yangsheng Li
lysh2001@whu.edu.cn
Jing Yang
yangjing@ncu.edu.cn
Xiaojue Peng
xiaojuepeng@ncu.edu.cn

Specialty section:
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

Received: *27 January 2022*
Accepted: *29 March 2022*
Published: *29 April 2022*

Citation:
Sun T, Wang T, Qiang Y, Zhao G, Yang J, Zhong H, Peng X, Yang J and Li Y (2022) CBL-Interacting Protein Kinase OsCIPK18 Regulates the Response of Ammonium Toxicity in Rice Roots. Front. Plant Sci. 13:863283. doi: 10.3389/fpls.2022.863283

INTRODUCTION

Nitrogen is one of the essential elements for plant growth and development, and ammonium (NH_4^+) is the predominant nitrogen source for plants (Forde, 2002; Hirano et al., 2008; Li et al., 2014). When ammonium is supplied at an appropriate concentration, it promotes plant growth and development. However, the elevated concentration of ammonium can give rise to symptoms of ammonium toxicity in many plants (Britto and Kronzucker, 2002; Liu and Wirén, 2017). Ammonium toxicity-related phenotypes include stunted growth, short and thick roots, lack of root gravitropism, yellowing of leaves, and even plant death in severe cases (Esteban et al., 2016). In agricultural production, overaccumulation of NH_4^+ in the soil normally happened at the excessive application of nitrogen fertilizers and unreasonable fertilization methods, which in turn affects plant growth and reproduction, and seriously harms the yield of crops (Dave and Nilsson, 2005; Romano and Zeng, 2007). Ammonium toxicity is of great ecological and economic importance. It is meaningful to study the response mechanism of ammonium toxicity.

To regulate the stress from ammonium toxicity, plants need to balance the activities of uptake, production, and consumption of NH_4^+. Thus, the mechanisms of ammonium toxicity have been found to be related to several cellular phenomena or activities, including ion transport, rhizosphere acidification, photosynthesis, phytohormones, and NH_4^+ futile cycle (Zhu et al., 2009; Esteban et al., 2016; Alencar et al., 2019; Meier et al., 2020). For example, members of the ammonium transporter 1 (AMT1) subfamily play a major role in NH_4^+ uptake, including OsAMT1;1, OsAMT1;2, and OsAMT1;3, and simultaneous knockout of these three genes resulted in a 95% reduction in NH_4^+ uptake (Yutaka et al., 2003; Miller and Cramer, 2005; Konishi and Ma, 2021). Besides, other proteins, such as the potassium transporters and plasma membrane H^+-ATPase, have also been reported to be involved in NH_4^+ transport (Kronzucker et al., 2001; Zhang M. et al., 2021). Meanwhile, NH_4^+ could be converted to organic nitrogen by assimilation through a metabolic cycle consisting of glutamine synthetase (GS) and glutamine-2-oxoglutarate aminotransferase (GOGAT) (Miflin and Habash, 2002; Li et al., 2014). Additionally, as the major players in the establishment and interconnection of plant signaling networks, phytohormones are also directly involved in ammonium toxicity (Krouk et al., 2010; Meier et al., 2020). For instance, the distribution of auxin is associated with NH_4^+-induced loss of root gravitropism (Zou et al., 2012). It has been proved that in rice the endogenous abscisic acid (ABA) could reduce reactive oxygen species (ROS) and free NH_4^+ of ammonium toxicity by regulating the SAPK9-bZIP20 pathway (Li et al., 2012; Sun et al., 2020). In spite of numerous studies on the mechanism of ammonium toxicity, the regulatory network related to ammonium toxicity in rice is still underdeveloped because of its complexity.

The calcineurin B-like protein (CBL), CBL-interacting protein kinase (CIPK) network, has been repetitively reported regulating several abiotic stress-induced signaling pathways, such as aluminum stress, K^+, NH_4^+, NO_3^- status, pH, salt stress, and oxidative stress (Hu et al., 2009; Yong et al., 2010; Hashimoto and Kudla, 2011). For example, in *Arabidopsis*, the AtCBL4-AtCIPK24 complex regulates the expression of the downstream functional gene AtSOS1 (Na^+/H^+ reverse transporter protein) to improve salt tolerance in roots (Sánchez-Barrena et al., 2005). AtCBL1/AtCBL9 interacts with AtCIPK23 to activate AtAKT1, a K^+ channel protein localized at the plasma membrane, thereby regulating K^+ uptake under low-K^+ conditions (Xu et al., 2006). An increasing number of studies have reported that several CIPKs altered their transcript levels and phosphorylation status during the ammonium response to its toxicity. When exposed to excess NH_4^+, the AtCBL1-AtCIPK23 complex phosphorylates AtAMT1s to inhibit NH_4^+ transport in *Arabidopsis* (Straub et al., 2017). In rice, the expression of OsCIPK23, OsCIPK8, OsCIPK9, and OsCIPK14/15 was sensitive to exogenous NH_4^+ (Xuan et al., 2019). Among them, OsCIPK9 regulates NH_4^+-dependent root growth downstream of OsIDD10 (Xuan et al., 2019). However, research in this area is still poorly understood and needs to be further explored.

In this study, we used the T-DNA insertion mutant *cipk18* to investigate the role of OsCIPK18 in ammonium toxicity. By observing the physiological and biochemical difference between WT and the mutant *cipk18* in the absence and presence of excess NH_4^+, we found that *cipk18* exhibited the decreased toxicity of ammonium and NH_4^+ accumulation, whereas GS/GOGAT enzyme activity for NH_4^+ assimilation remained the same. It was further demonstrated that OsCIPK18 regulates free NH_4^+ in roots by affecting the expression of NH_4^+ transporters, including OsAMT1;2. Finally, RNA-seq was used to analyze the transcriptome data to further characterize the molecular mechanisms and identify promising candidates of transcription factors (TFs) affecting ammonium toxicity that may depend on the OsCIPK18 regulatory pathway.

MATERIALS AND METHODS

Plant Materials and Growth Conditions

Wild-type lines of rice (*Japonica*, Hwayoung), T-DNA insertion mutant *cipk18,* and three complementary strains of *cipk18* (com1, com2, com3) were used in this study. The OsCIPK18 T-DNA insertion line (1C-05857) in the Hwayoung background was obtained from Kyung Hee University, Korea (Jeon et al., 2000). To generate the complemented lines, the 35S::OsCIPK18 fragment was constructed and transformed into *Agrobacterium tumefaciens* strain EHA105. Calli from *cipk18* were transformed as described by Cheng et al. (1998). Seeds used in this study were surface-sterilized with 1% sodium hypochlorite for 10 min, washed extensively with distilled water, and then incubated at 37°C in the dark for 2 days to break dormancy. Consistently growing germinating seeds were, respectively, transferred to two groups of hydroponic media for growth, namely, control (pure water) and ammonium toxicity treatment (4 mM NH_4Cl solution) with no other nutrients present in the hydroponic media. All seedlings subjected to hydroponic treatments were grown in a temperature-controlled incubator set at 28°C, 14 h light/22°C, 10 h dark. All mature plants were grown and harvested in Wuhan, Hubei, and Lingshui, Hainan, China.

Phenotyping of Rice Seedlings

The roots of seedlings of WT, *cipk18*, com1, com2, and com3 cultured for 3 days under ddH_2O and 4 mM NH_4Cl, respectively, were photographed at high resolution with a Nikon D7100 digital SLR camera, and root length data were obtained using smartRoot in ImageJ (Lobet and Draye, 2011). Seedlings of WT and *cipk18* that had been cultured for 7 days under control and NH_4Cl treatments were selected and divided into groups of five plants to measure biomass, and the average value of individual plants was calculated and repeated three times. The roots of seedlings were drained and weighed directly to obtain fresh weight data, dried in an oven at 70°C for 3 days until their constant weight was obtained, and then weighed again to obtain dry weight data. The roots of 7-day-old rice seedlings were spread as far as possible, and high-resolution photographs were taken with a Nikon D7100 digital SLR camera to obtain root length, diameter, and number data using smartRoot in ImageJ.

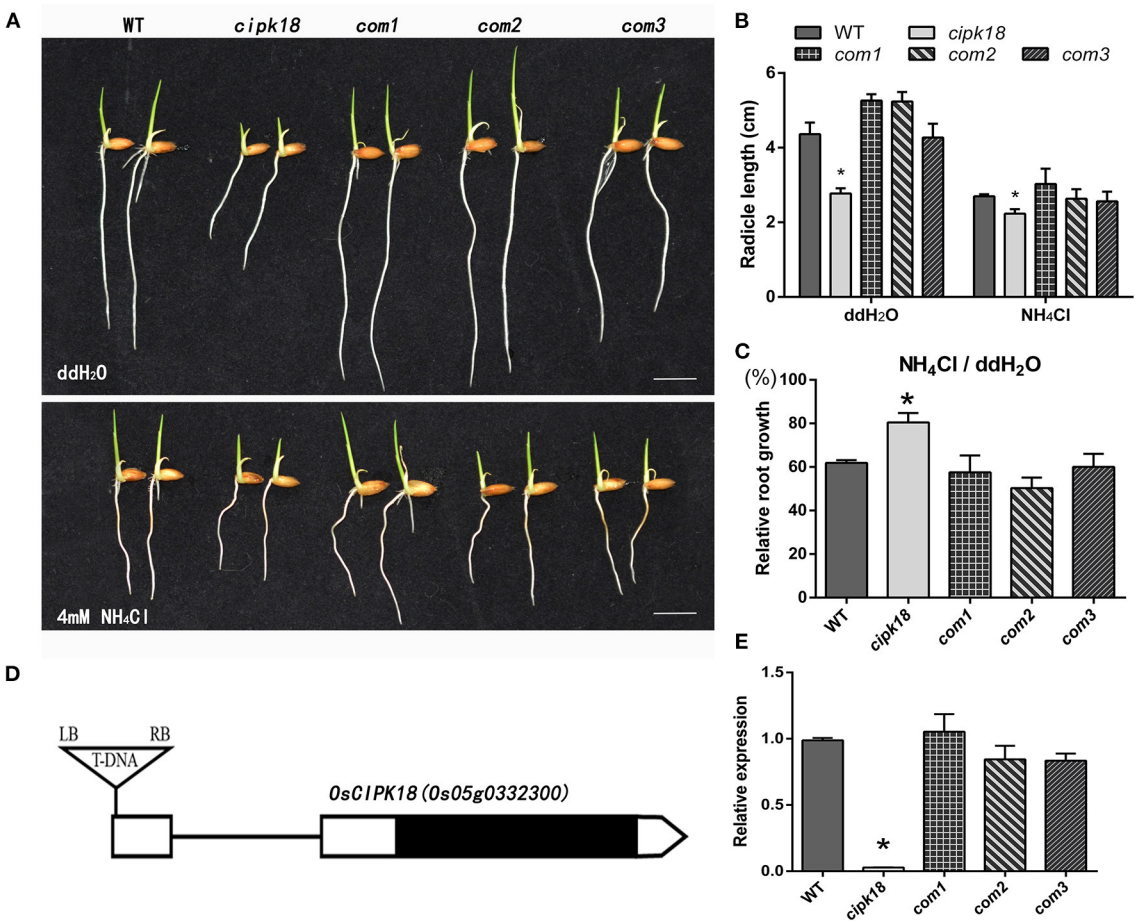


FIGURE 1 | The mutant *cipk18* shows low levels of ammonium toxicity. **(A,B)** Rice seedlings phenotypes **(A)**, radicle length **(B)** of 3-day-old WT, *cipk18,* and three complementary materials of *cipk18* (com1, com2, com3) under treatments with ddH_2O and 4 mM NH_4Cl, respectively. Scale bar = 1 cm. **(C)** Compared with the control, relative root growth of the five lines under 4 mM NH_4Cl. **(D)** Gene structures of the OsCIPK18 insertion mutant. UTRs are shown as white boxes and exons as black box. Triangle indicates T-DNA insertion sites. **(E)** Relative expression of OsCIPK18 in WT, *cipk18,* and three complementary materials of *cipk18*. Data are means ± SDs, significant differences using Student's *t*-test: $^*p < 0.05$. RB, right border of T-DNA; LB, left border of T-DNA.

Tissue Ammonium Concentration Determination

The roots of 7-day-old rice seedlings were removed from the culture medium, rinsed three times with water, dried with paper towels to absorb the water, ground to powder with liquid nitrogen, weighed 0.1 g of the powder, and treated with 1 ml of 10 mM formic acid solution. After mixing thoroughly, they were centrifuged at 10,000 × g at 4°C for 10 min. Then, 1 ml OPA (100 mM KH_2PO_4, 100 mM K_2HPO_4, 3.75 mM *o*-phthalaldehyde, 2 mM β-ME) was added to 250 μl supernatants and reacted in the dark for 30 min, the absorbance was measured at 410 nm using infinite M200 Enzyme Scale. A standard curve was plotted with different concentrations of NH_4Cl solution, and then it was used to calculate the concentration of tissue NH_4^+.

Assays of GS and GOGAT Activity

The roots of WT and *cipk18* seedlings cultured for 7 days under two treatments described above were ground to powder with liquid nitrogen. Then, 0.1 g of tissue was used to measure GS and GOGAT activity according to the method described on the kit (cominbio). Absorbance value measurement was done using infinite M200 Enzyme Scale.

RNA Extraction and Quantitative Real-Time Reverse Transcriptase-Polymerase Chain Reaction (RT-PCR)

To verify the expression changes of genes in WT and *cipk18* under ammonium toxicity, we extracted total RNA from roots of rice seedlings under ddH_2O and 4 mM NH_4Cl treatment and performed qRT-PCR using specific primers. RNA was extracted as one biological replicate from the roots of approximately 5–8 seedlings grown for 7 days in the same growth state under both treatments. We rapidly ground the treated roots in liquid nitrogen, extracted total RNA from rice samples using TRIzol (Invitrogen), removed genomic DNA from total RNA using DNase I (Thermo Scientific), and performed cDNA first-strand

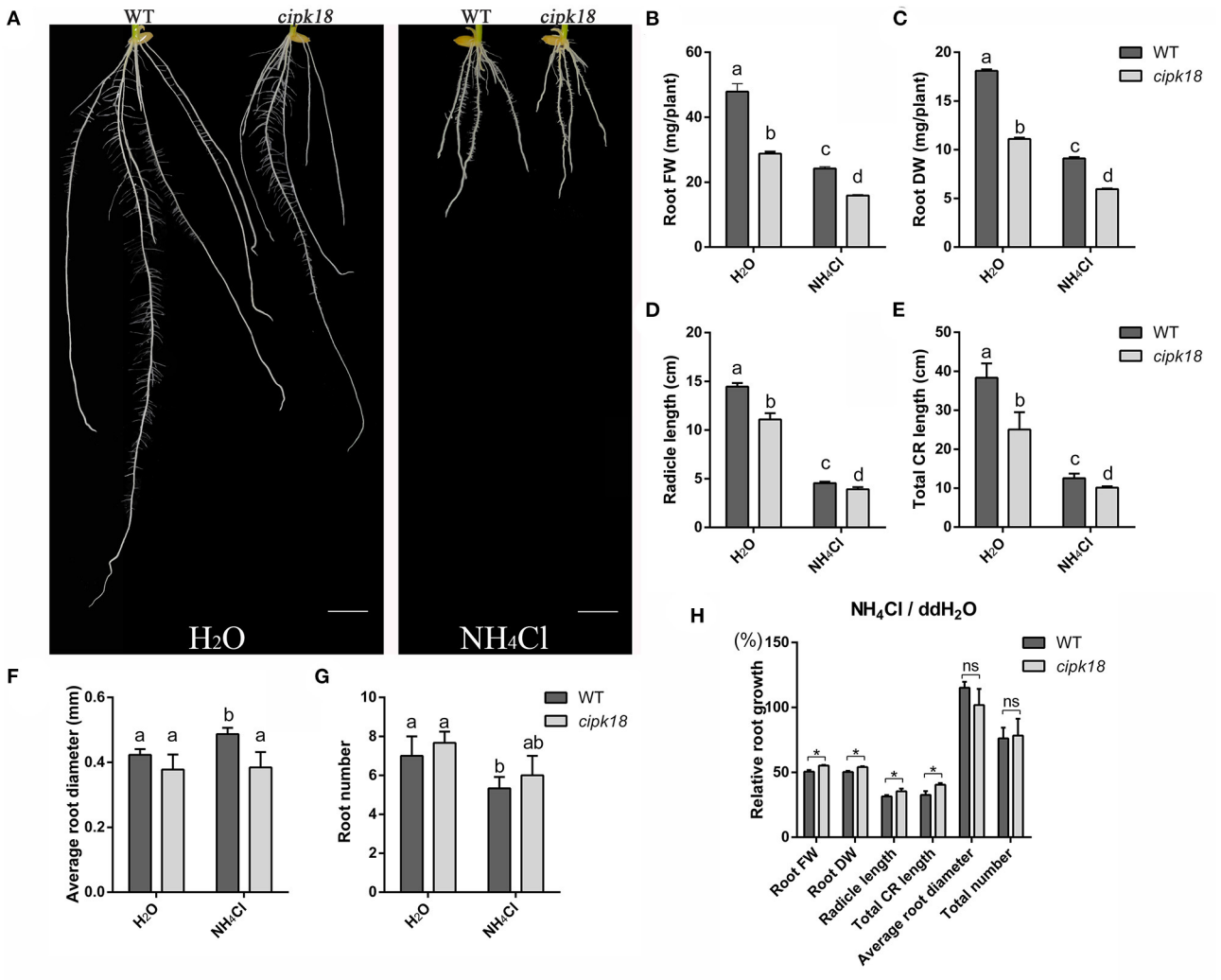


FIGURE 2 | OsCIPK18 affects the root growth of rice seedlings under ammonium toxicity. Rice seedlings phenotypes **(A)**, root FW **(B)**, root DW **(C)**, radicle length **(D)**, total CR length **(E)**, average root diameter **(F)**, and root number **(G)** of 7-day-old WT and *cipk18* plant cultured in ddH_2O and 4 mM NH_4Cl, respectively. Scale bar = 1 cm. Different letters represent the significant difference, $p < 0.05$. **(H)** Relative root growth, i.e., the ratio of root growth data in the presence of ammonium toxicity to that of the control. Data are means ± SDs, significant differences using Student's *t*-test: $^*p < 0.05$. FW, fresh weight; DW, dry weight; CR, crown root.

synthesis using M-MLV reverse transcriptase (Promega). qRT-PCR was done using 2× ChamQ SYBR qPCR Master Mix (UE), specific primers P+/P–, and cDNA using the Bio-Rad CFX96TM Real-Time System. The relative expressions were calculated using the $2^{-\Delta\Delta CT}$ method. Three technical replicates were used to calculate the mean of the expression levels for each biological replicate, and three biological replicates were used to generate the mean of the expression levels for each sample. Graphing and significance analysis were performed using GraphPad Prism 6.

RNA Sequencing and Data Analysis

The roots of WT and *cipk18* seedlings treated for 7 days under two treatments as described previously were divided into three replicates to prepare specific RNA-seq libraries, respectively, for a total of 12 libraries. The libraries were submitted to BGI for sequencing using the Illumina HiSeq 2000 platform. The raw data were filtered using trimmomatic (Bolger et al., 2014) to remove the low-quality reads, and the resulting high-quality reads were aligned to the NIP reference genome (ftp://ftp.ensemblgenomes.org/pub/plants/release-44/fasta/oryza_sativa/dna/) using STAR software (Dobin et al., 2013). After alignment, the raw counts were normalized to trimmed mean of M value (TMM) using RSEM (Li and Dewey, 2011). Difference analysis between the two samples was performed using DESeq2, where genes with $p_{adj} < 0.05$ and $|\log2FC| > 1$ were identified as differential genes. Gene Ontology (GO) enrichment analysis was performed using TBtools, and the GO background was provided on the AgriGO V2.0 website (Tian et al., 2017; Chen et al., 2020). Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis was done using the R package from clusterprofiler (Yu et al., 2012).

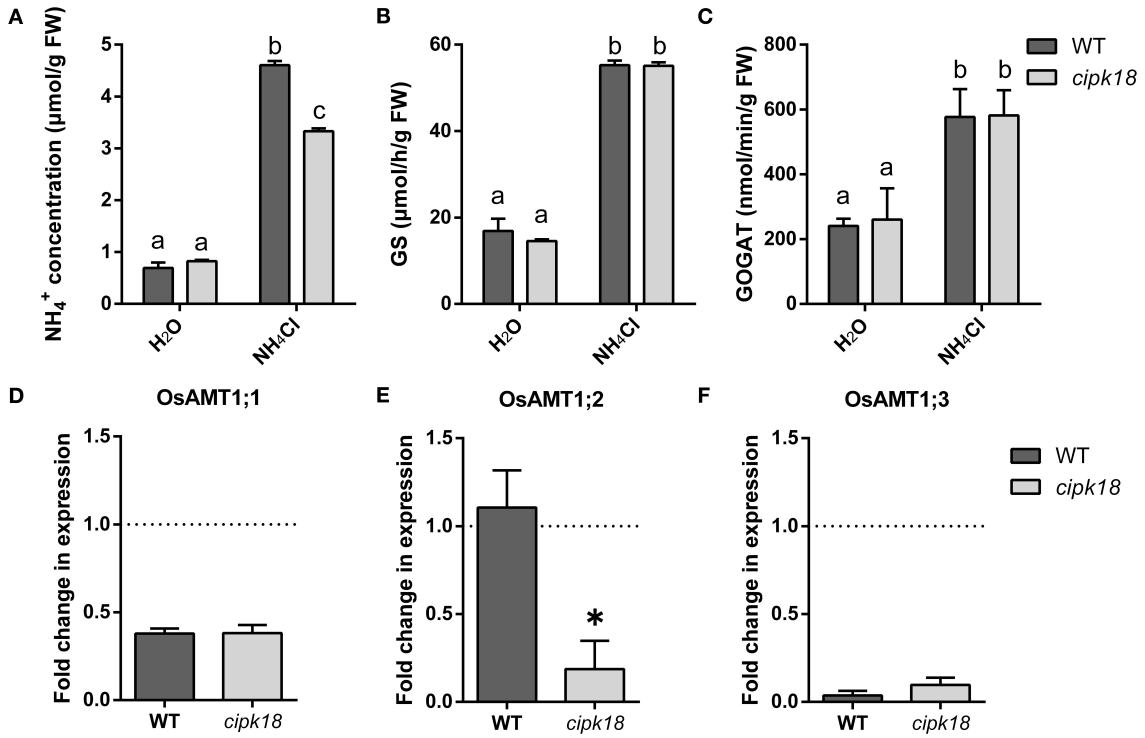


FIGURE 3 | OsCIPK18 affects NH_4^+ accumulation in rice seedlings root under ammonium toxicity. NH_4^+ concentration **(A)**, GS activity **(B)**, and GOGAT activity **(C)** of 7-day-old rice seedlings root. Fold changes in the expression of OsAMT1;1 **(D)**, OsAMT1;2 **(E)**, and OsAMT1;3 **(F)** under ammonium toxicity compared with control, as determined by quantitative real-time PCR. Data are means ± SDs, different letters represent the significant difference, $p < 0.05$. GS, glutamine synthetase; GOGAT, glutamate synthase.

TF Prediction

Transcription factor prediction was performed on the plantregmap website (http://plantregmap.gao-lab.org/), and the predicted results were filtered. TF network visualization was done by Cytoscape.

Statistical Analysis

All statistical analyses were conducted using GraphPad Prism 6. The error bars in all the charts represent the standard deviation of the mean. The different letters above the error bar represent significant differences between groups. The comparison method uses one-way ANOVA combined with Duncan's *post-hoc* multiple test method. $p < 0.05$ was set as the significance cutoff.

RESULTS

Knockdown of OsCIPK18 Enhances Ammonium Resistance in Rice Roots

To screen the specific CIPK associated with ammonium toxicity in rice, we compared the ammonium resistance of WT with the mutant of OsCIPK T-DNA insertion. Ammonium resistance was quantified based on the phenotype of the relative root growth (i.e., root growth in the presence of high NH_4^+ compared with root growth in water as control). Growth inhibition but relatively enhanced ammonium resistance was identified in a mutant (1C-05857), in which ammonium toxicity only shortened its root length by about 20%, but by about 40% in WT (**Figures 1A–C**). This mutant carried a homozygous T-DNA insertion in the 5′ untranslated region (UTR) of OsCIPK18 (*Os05g0332300*), leading to a knockdown of the OsCIPK18 transcript (**Figures 1D,E**, **Supplementary Figure 1**). We obtained three OsCIPK18 complementation lines (com1, com2, and com3) and confirmed that the increased ammonium resistance was attributed by the mutation of OsCIPK18 (**Figures 1A–C,E**). With increasing incubation time, *cipk18* consistently showed resistance to excess NH_4^+ in terms of fresh weight and root length compared to WT and complementary lines (**Supplementary Figure 2**).

Furthermore, we investigated the development of root system between WT and the mutant *cipk18* under control and ammonium stress conditions. Exposing to excess NH_4^+, WT exhibited a significant decrease in biomass, root length and root number, and an increase in root diameter, which indicated that external high NH_4^+ severely affects rice root development of WT (**Figures 2A–H**). However, ammonium toxicity did not cause the same changes in root diameter and root number in the mutant *cipk18*. The effects of ammonium toxicity on root length of *cipk18* were diminished, and the fresh

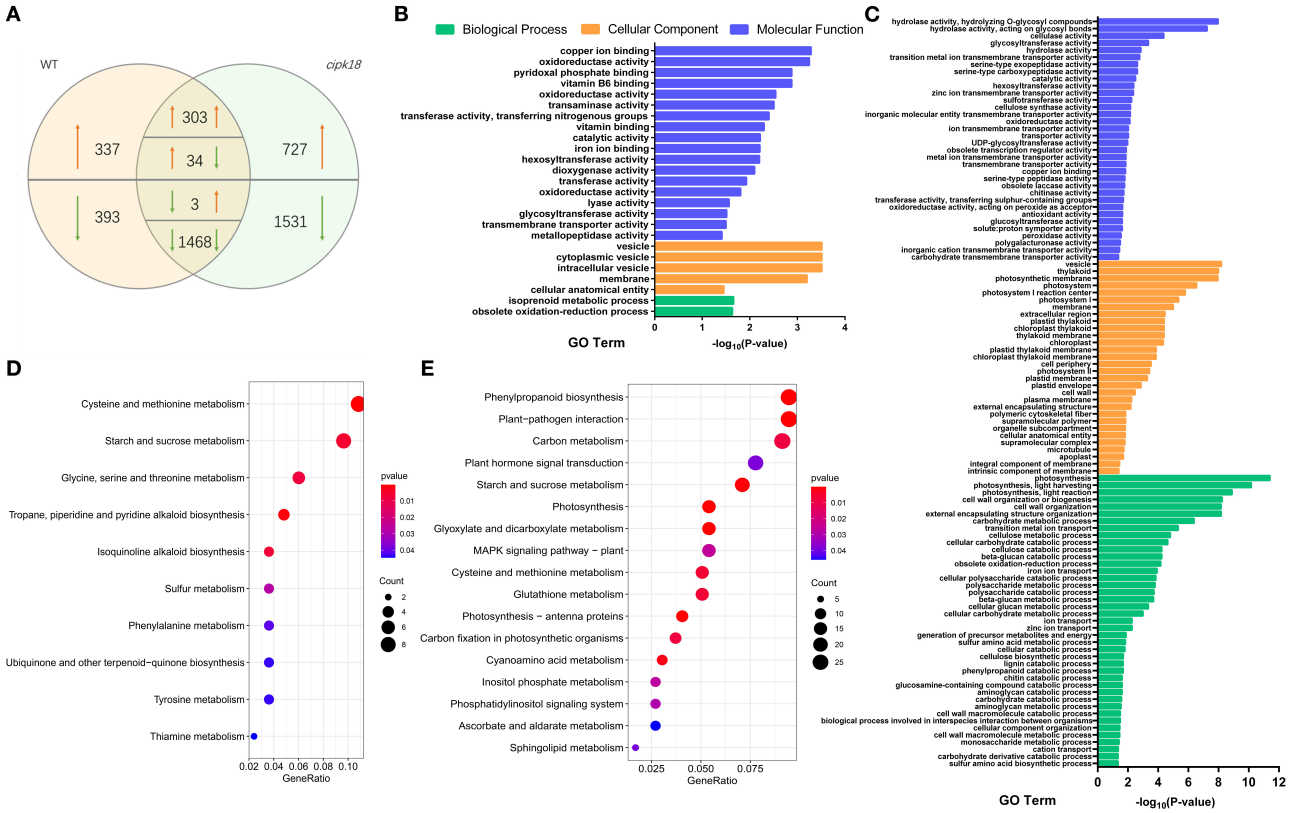


FIGURE 4 | Transcriptomic analysis and enrich analysis of WT and *cipk18* in response to ammonium toxicity. **(A)** Venn diagram illustrates the number of DEGs in WT and *cipk18* under ammonium toxicity. **(B)** GO analysis of WT-specific DEGs. **(C)** GO analysis of *cipk18*-specific DEGs. **(D)** KEGG analysis of WT-specific DEGs. **(E)** KEGG analysis of *cipk18*-specific DEGs. DEGs, differentially expressed genes; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

and dry weight also showed stronger ammonium resistance compared with WT (**Figures 2A–H**). Together, these results suggest that *cipk18* was more resilient to ammonium toxicity than WT, and protein kinase OsCIPK18 might be involved in the adjustment of the root structure of rice seedlings by ammonium toxicity.

OsCIPK18 Affects the Accumulation of NH_4^+ Under Ammonium Toxicity Without Affecting Assimilation

The excessive accumulation and assimilation of NH_4^+ in plants are considered as the important causes accounting for ammonium toxicity (Chen et al., 2013). To estimate the accumulation and assimilation of NH_4^+ in the roots of both WT and *cipk18*, we examined NH_4^+ content and enzymatic activity of GS/GOGAT under the control and ammonium stress conditions (**Figures 3A–C**). When only water was provided, the roots of both WT and *cipk18* had low and comparably basal amounts of NH_4^+ (**Figure 3A**). When excessive NH_4^+ was provided, a significant accumulation of NH_4^+ was observed in both lines (WT and *cipk18*). But the amount of NH_4^+ in the *cipk18* roots was significantly lower than that in the WT (**Figure 3A**). The enzymatic activities of GS and GOGAT for NH_4^+ assimilation in roots of both WT and *cipk18* were significantly increased due to external high NH_4^+. But there was no significant difference in the enzymatic activity between WT and *cipk18* (**Figures 3B,C**). These data indicated that the observed differential ammonium resistance between WT and *cipk18* lines could be a result of differences in root NH_4^+ accumulation, which was not associated with GS/GOGAT-dependent assimilation.

The difference in NH_4^+ accumulation in roots prompted us to investigate the changes in transcript abundance of OsAMT1s that encoded ammonium transporters. The expression of OsAMT1;1 and OsAMT1;3 was significantly suppressed to the same extent in both WT and *cipk18* by external high NH_4^+ (**Figures 3D,F**). The expression of OsAMT1;2 stayed unchanged in WT but was significantly downregulated in *cipk18* upon the external high NH_4^+ (**Figure 3E**). These findings suggest that the differences in NH_4^+ accumulation between WT and *cipk18* lines may be the result of direct or indirect involvement of OsCIPK18-mediated signaling in the transcriptional regulation of OsAMT1;2. Furthermore, this correlation prompted us to focus on the transcript changes in WT and *cipk18* during high NH_4^+ stress and identify the key genetic elements involved in differential ammonium resistance between the two lines.

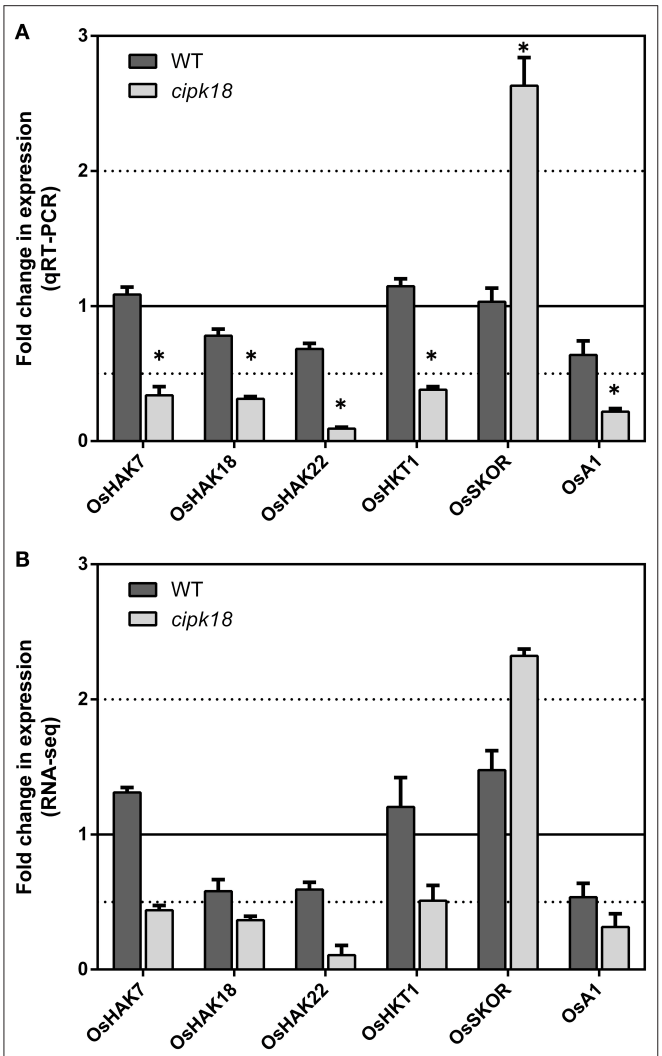


FIGURE 5 | The expression of NH_4^+ transporters regulated by OsCIPK18 under ammonium toxicity. Fold changes in the expression of NH_4^+ transporters under ammonium toxicity, as determined by qRT-PCR **(A)** and RNA-seq analysis **(B)**. Data are means ± SDs. Significant differences using Student's *t*-test: $^*p < 0.05$.

Genes Are Differentially Regulated Between WT and *cipk18*

Transcriptome comparisons were performed between the control and ammonium stress treatments. In the WT line, a total of 2,538 excess NH_4^+-induced differentially expressed genes (DEGs) were identified, of which 674 DEGs were upregulated and 1,864 DEGs were downregulated in the presence of excess NH_4^+ (**Figure 4A**). In mutant *cipk18*, 4,066 excess NH_4^+-induced DEGs were identified, of which 43.6% (303 + 1,468 out of 4,066) had the same expression change trend in WT (**Figure 4A**). These overlapping DEGs most likely reflect regulatory networks induced by excess NH_4^+ independent of OsCIPK18. Notably, there were 730 (337 + 393) genes differentially expressed only in WT but not in *cipk18*, named WT-specific DEGs, reflecting ammonium toxicity-induced changes in gene expression dependent on OsCIPK18 (**Figure 4A**, **Supplementary Table 1**). These transcript changes may be one of the reasons for the phenotypic differences between WT and *cipk18*. In addition, 2,258 (727 + 1,531) DEGs were uniquely induced by high NH_4^+ in *cipk18* but not in WT, i.e., *cipk18*-specific DEGs, possibly reflecting an additional mechanism of ammonium resistance caused by knockdown of OsCIPK18 (**Figure 4A**, **Supplementary Table 2**). It is necessary to be concerned that there were 270 DEGs between WT and *cipk18*, even in the control conditions (**Supplementary Figure 3A**). These DEGs demonstrate important regulatory roles for OsCIPK18 in various signaling and physiological pathways in rice roots (**Supplementary Figures 3C,D**), which could explain the phenotypic differences observed in *cipk18* and WT lines (**Figures 1A**, **2A**). There were 4.9% WT-specific DEGs (36 out of 730) and 2.1% *cipk18*-specific DEGs (48 out of 2258) overlapped with DEGs between WT and *cipk18* (**Supplementary Figure 3B**), and these overlapped DEGs may constitute factors that influence the response of *cipk18* to ammonium toxicity.

Multiple Channels of NH_4^+ Transport Regulated by OsCIPK18

Ammonium could be transported through simple osmotic diffusion, non-selective cation channels, and potassium transport channels (Britto et al., 2014; Bittsánszky et al., 2015). Our previous expression analysis of AMT1 subfamily members had demonstrated that OsCIPK18 regulated NH_4^+ influx by altering the expression of OsAMT1;2 (**Figure 3E**). Using RNA-seq, we further identified four K^+ transporters, OsHAK7 (*Os07g0669700*), OsHAK18 (*Os09g0563200*), OsHAK22 (*Os07g0102100*) (Banuelos et al., 2002), and OsHKT1 (*Os06g0701700*) (Yao et al., 2010), a plasma membrane H^+-ATPase OsA1 (*Os03g0689300*), as well as an outward-rectifying shaker-like potassium channel OsSKOR (*Os04g0445000*) (Kim et al., 2015) in *cipk18*-specific DEGs (**Figure 5**). In *cipk18*, four K^+ transporters were uniquely downregulated by excess NH_4^+ to reduce NH_4^+ uptake; meanwhile, OsSKOR was uniquely upregulated to increase NH_4^+ efflux (**Figure 5**). OsA1 could cooperatively improve N and C utilization and facilitates ammonium absorption in rice (Zhang M. et al., 2021). Its expression was significantly downregulated by excess NH_4^+ in *cipk18*, which may be one of the pathways through which OsCIPK18 regulates NH_4^+ uptake (**Figure 5**). In summary, we considered that the lower amount of NH_4^+ in *cipk18* could be the result of the simultaneous downregulation of OsAMT1;2, OsA1, OsHAK7/18/22, OsHKT1, and upregulation of OsSKOR. To further investigate how OsCIPK18 functions in the inhibition of root growth by NH_4^+, we investigated the functional distribution of WT-specific and *cipk18*-specific DEGs under ammonium toxicity, respectively, and tried to find the key genes.

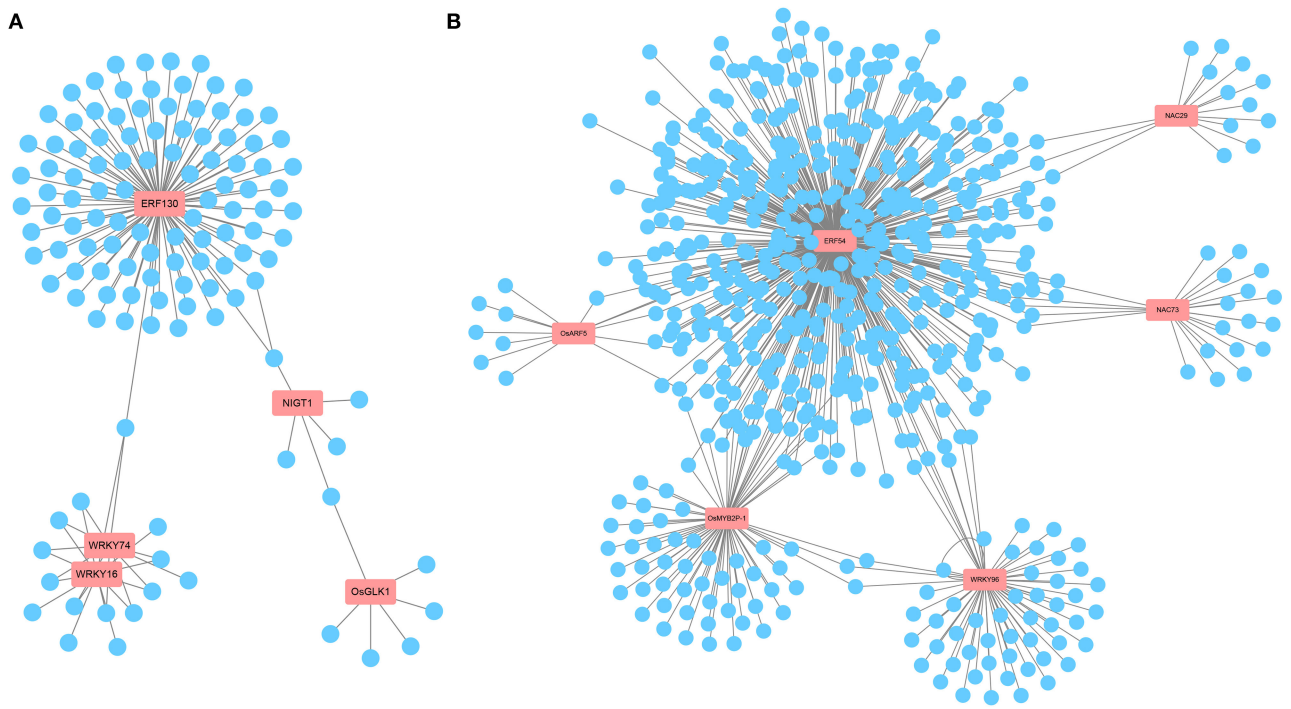


FIGURE 6 | The key TFs and prediction network composed of DEGs. **(A)** The prediction network composed of WT-specific DEGs. **(B)** The prediction network composed of *cipk18*-specific DEGs. The genes in the red boxes are key TFs; blue circles represent predicted genes that may be activated or repressed in expression by TFs; gray lines represent possible transcriptional activation or repression. TFs, transcription factors.

Ammonium Toxicity Response Process Dependent on OsCIPK18

WT-specific DEGs induced by excessive NH_4^+ differentially expressed only in WT represent an OsCIPK18-dependent ammonium toxicity response process. We explored the function of WT-specific DEGs based on GO and KEGG classification (**Figures 4B,D**). A total of 25 GO pathways and 10 KEGG pathways were enriched, mainly including transmembrane transporter activity, ion binding, vesicle, oxidoreductase activity, starch and sucrose metabolism, amino acid metabolism, and biosynthesis of small molecules, such as isoquinoline alkaloid and ubiquinone (**Figures 4B,D**).

Transcription factors, as key factors coordinating the expression of downstream genes involved in metabolic and developmental pathways, are important players in the response to ammonium toxicity (Kikuchi et al., 2000; Huang et al., 2015; Gu et al., 2017). Consistent with this, we found that multiple TF family members, such as AP2/ERF, MYB, WRKY, and NAC, were induced to be upregulated or downregulated upon exposure to excess NH_4^+ (**Supplementary Table 1**). In WT-specific DEGs, according to the predicted TF network using PlantRegMap, we found that five core TFs, OsERF130 (*Os05g0497200*), OsWRKY74 (*Os09g0334500*), OsWRKY16 (*Os01g0665500*), OsNIGT1 (*Os02g0325600*), and OsGLK1 (*Os06g0348800*), had possible interactions with multiple genes among the WT-specific DEGs (**Figure 6A**, **Supplementary Figure 4**, **Supplementary Table 3**). Among them, OsWRKY74 (Dai et al., 2016), whose expression was downregulated by NH_4^+ in WT, affected the elongation of roots and the increase of crown root number. This led us to speculate that ammonium toxicity signals might be transmitted through OsCIPK18 to OsWRKY74 to influence root development. Overall, this network revealed some key TFs that may be involved in root growth regulated by OsCIPK18 under NH_4^+ stress and regulatory relationships in WT-specific DEGs.

In addition, genes associated with ABA signaling in WT-specific DEGs attracted our attention. OsRePRP2.1 (*Os07g0418700*) and OsRePRP2.2 (*Os07g0418600*) were reported to be repressors of root cell expansion and were induced by ABA to be expressed in the elongation zone of roots (Tseng et al., 2013). In this study, their expression was significantly upregulated in WT (**Figures 7A,B**). Meanwhile, OsSAPK6 (*Os02g0551100*), encoded stress-activated protein kinase and involved in ABA signaling (Kobayashi et al., 2004), was uniquely downregulated in *cipk18* (**Figures 7A,B**). The expression changes of these genes implied that excess NH_4^+ activated ABA signaling and increased the expression of OsRePRP2.1/2.2 to inhibit root elongation. In contrast, due to the deletion of OsCIPK18, the expression of the ABA signaling factor OsSAPK6 was downregulated and the expression of OsRePRP2.1/2.2 was no longer upregulated, allowing the mutant *cipk18* to show greater resistance to ammonium.

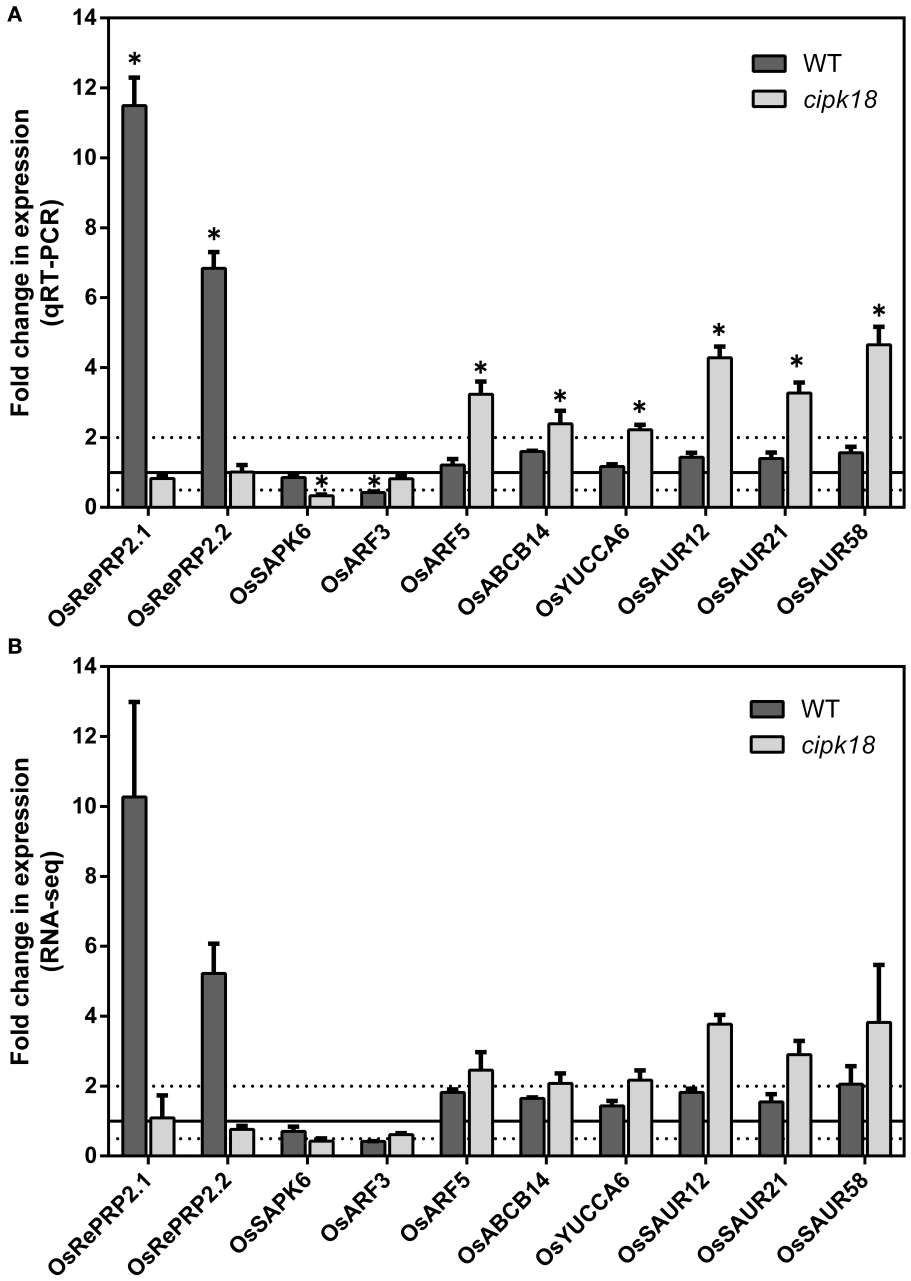


FIGURE 7 | The expression of genes related to ABA and auxin signaling regulated by OsCIPK18 under ammonium toxicity. Fold changes in the expression of genes under ammonium toxicity, as determined by qRT-PCR **(A)** and RNA-seq analysis **(B)**. Data are means ± SDs. Significant differences using Student's *t*-test: $^*p < 0.05$.

Mechanisms of Ammonium Resistance in the Mutant *cipk18*

In this study, *cipk18*-specific DEGs induced by excess NH_4^+ in *cipk18* were considered as components of the ammonium resistance mechanisms. Based on the GO and KEGG classification, the function of *cipk18*-specific DEGs is mainly related to ion transport, transmembrane transporter activity, cell wall macromolecule metabolic process, plant hormone signal transduction, cellular catabolic process, glutathione metabolism, sugar catabolic process, and photosystem (**Figures 4C,E**).

In *cipk18*-specific DEGs, we identified six key TFs that are likely to be the regulatory center of this part network, including OsERF54 (*Os01g0657400*), OsMYB2P-1 (*Os05g0140100*), OsWRKY96 (*Os12g0507300*), OsNAC29 (*Os08g0115800*), OsNAC73 (*Os01g0672100*), and ARF5 (*Os04g0664400*) (**Figure 6B**, **Supplementary Figure 4**, **Supplementary Table 4**). Among these genes, OsMYB2P-1, an R2R3 MYB TF engaged

in phosphorus starvation response and regulation of root architecture in rice, was uniquely downregulated by NH_4^+ in *cipk18* (Dai et al., 2012). OsNAC29, a top-layer TF for secondary wall formation (Huang et al., 2015), was uniquely downregulated by NH_4^+ in *cipk18,* which was consistent with the cellulose synthesis-related pathway was enriched by GO classification. We, therefore, hypothesized that OsCIPK18 is associated with cell wall formation under ammonium toxicity.

Notably, OsARF5 is an auxin response factor whose expression is upregulated by excess NH_4^+ in the mutant *cipk18* (**Figures 7A,B**). This reminds us to be concerned about the role of auxin signaling in ammonium toxicity response. As we all know, ammonium feeding has been shown to suppress root auxin content (Kudoyarova et al., 1997; Britto and Kronzucker, 2002). In *cipk18*, we revealed that exposure to high NH_4^+ resulted in the upregulation of auxin synthesis and response gene expression, including the auxin influx transporter OsABCB14 (*Os04g0459000*) (Xu et al., 2014), the IAA synthetic pathway gene OsYUCCA6 (*Os07g0437000*) (Yamamoto et al., 2007), and three auxin-responsive SAUR gene family members [OsSAUR12 (*Os02g0769100*), OsSAUR21 (*Os04g0617050*), OsSAUR58 (*Os12g0626200*)] (**Figures 7A,B**) (Jain et al., 2006; Zhang T. et al., 2021). These results implied that knockdown of OsCIPK18 may have prevented the inhibition of auxin synthesis in roots by excess NH_4^+, allowing upregulation of auxin synthesis and response genes, thereby enhancing the resistance to ammonium toxicity in *cipk18*. Meanwhile, we found that OsARF3 (*Os01g0753500*) (Wang et al., 2007) was uniquely downregulated by excess NH_4^+ in WT-specific DEGs (**Figures 7A,B**), further supporting this hypothesis. Therefore, we hypothesized that seedling roots inhibit auxin signaling *via* OsCIPK18 in response to ammonium toxicity signals, resulting in inhibition of root growth in rice.

DISCUSSION

The Mutant *cipk18* Shows NH_4^+-Resistance Phenotype

As a result of human intervention in the nitrogen cycle, including increasing soil nitrogen input and irrational fertilization practices in the biosphere, plants have to deal with unprecedented NH_4^+ stress (Gerendás et al., 1997; Britto and Kronzucker, 2002). Even though rice is a recognized ammonium-tolerant species and highly adapted to NH_4^+ as a nitrogen source, it is still threatened by ammonium toxicity (Balkos et al., 2010). Our study showed strong evidence that WT exhibited significant growth inhibition when it was grown under excessive NH_4^+ treatment, with few and short thick roots and reduced biomass (**Figures 1**, **2**). However, a T-DNA insertion mutant *cipk18* exhibits enhanced resistance in response to ammonium toxicity as both biomass and root architecture showed that the inhibition of high NH_4^+ on roots growth in *cipk18* was less than that in WT (**Figures 1**, **2**).

The CIPK family interacts with members of the CBL family in response to the Ca_2^+-mediated signaling pathway (Straub et al., 2017). OsCIPK18, as a member of the CIPK family, must also play an important role in rice growth and development, but this has not been reported yet. In this study, the mutant *cipk18* exhibited significant growth inhibition compared with WT (**Figures 1**, **2**). Using GO and KEGG analysis, the 270 DEGs that were identified between WT and *cipk18* were associated with biological functions, such as signaling, response to stimuli, ion binding, and nitrogen metabolism (**Supplementary Figures 3A,C,D**). This discovery facilitates the study of the effects of OsCIPK18 in rice, but this study focuses on the role of OsCIPK18 in response to ammonium toxicity. The mutant *cipk18* shows NH_4^+-resistance phenotype, but mutants of two other members of the CIPK family, OsCIPK9 and OsCIPK23, were sensitive to NH_4^+ (Xuan et al., 2019), which suggested that the function of OsCIPK18 in response to ammonium toxicity might be opposite to that of OsCIPK9 and OsCIPK23.

OsCIPK18 Affects NH_4^+ Transport

Traditional theories suggest that high intracellular ammonium concentration is one of the major causes of ammonium toxicity in higher plants, and there is also emerging evidence that acidic stress caused by excessive NH_4^+ assimilation is the primary cause (Hachiya et al., 2012, 2021; Esteban et al., 2016). In this study, intracellular NH_4^+ accumulation and GS/GOGAT-dependent NH_4^+ assimilation in roots were significantly increased by excess NH_4^+ in both WT and the mutant, but relatively low NH_4^+ accumulation was observed in *cipk18* (**Figures 3A–C**), suggesting that OsCIPK18 is involved in the NH_4^+ uptake without affecting NH_4^+ assimilation. In the presence of elevated external NH_4^+, NH_4^+ influx dependent on the high-affinity transport system is downregulated to prevent ammonium toxicity (Kronzucker et al., 2001; Beier et al., 2018; Kumar et al., 2020). For example, CBL1-CIPK23 phosphorylates and inactivates AMT1;1/1;2 in response to ammonium stress, thereby reducing NH_4^+ transport (Straub et al., 2017). In this study, expression of OsAMT1;1 and OsAMT1;3 was significantly downregulated by excess NH_4^+ in both WT and mutant (**Figures 3D,E**). However, in *cipk18*, the expression of OsAMT1;2, OsHAK7, OsHAK18, OsHAK22, OsHKT1, and OsA1, which are reported to transport NH_4^+, were all inhibited by excess NH_4^+ (**Figures 3E**, **5A,B**). Changes in the expression of these genes coincided with reduced ammonium accumulation in the mutant, suggesting that OsCIPK18 plays an important role in maintaining the expression of ammonium transporters at the transcriptional level during the response to ammonium toxicity.

Ammonium Toxicity-Induced Transcriptional Regulatory Network Dependent on OsCIPK18

Apart from the regulation of NH_4^+ uptake, it is not known how OsCIPK18 regulates root growth under ammonium stress. Transcriptome-wide analysis facilitated our observation of high NH_4^+-induced changes in gene expression associated with root growth at the genome-wide level. Sun et al. used RNA-seq to reveal the spatiotemporal specificity of gene expression in rice after high ammonium treatment for 4 and 12 h, highlighting the role of TFs and phytohormones in

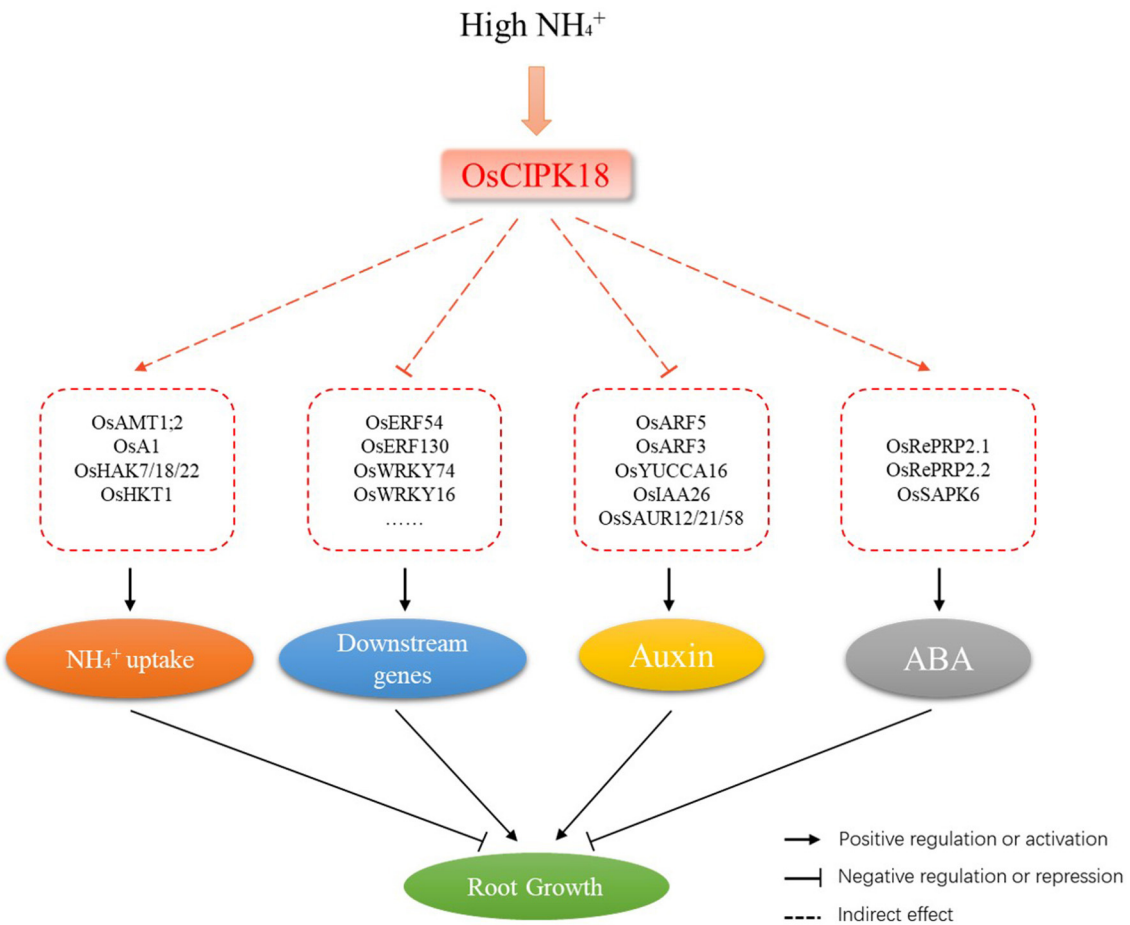


FIGURE 8 | A schematic model of the involvement of OsCIPK18 in ammonium toxicity in rice seedling roots. osCIPK18 directly or indirectly regulates the expression of several genes that affect NH_4^+ accumulation, the ABA and auxin signaling pathways, and the transcription factor regulatory network, ultimately regulating root growth.

ammonium resistance (Sun et al., 2017). In our data, WT-specific DEGs and *cipk18*-specific DEGs were identified as differential genes involved in response to ammonium toxicity and located downstream of OsCIPK18 (**Figure 4A**). Strategies for coping with salt stress in plants include adjusting ATP formation and enhancing energy metabolism (Zhao et al., 2013). Genes related to amino acid metabolism, starch, and sucrose metabolism, as well as oxidoreductase, lysozyme, and many other enzymatic activities are regulated by OsCIPK18 under ammonium toxicity (**Figures 4B–E**). Recent evidence suggests that the biosynthesis of vitamin B6 significantly improves root tolerance to ammonium (Liu et al., 2022). Our data show that genes associated with vitamin B6 binding are also regulated by OsCIPK18 (**Figure 4B**). The above evidence highlights the fundamental function of OsCIPK18 in the ammonium toxicity response.

Several core TFs were identified in WT-specific and *cipk18*-specific DEGs that could be directly or indirectly regulated by OsCIPK18 under ammonium stress (**Figures 6A,B**). Members of the AP2/ERF family are associated with plant defense programs against abiotic stresses and share a conserved DNA-binding domain that activates downstream gene expression by binding specifically to *cis*-acting elements in the promoters of abiotic stress-responsive genes (Mizoi et al., 2012). In this study, two ERF family genes (OsERF54, OsERF130) were found to be at the center of the transcriptional regulatory network involved in ammonium toxicity by OsCIPK18 (**Figures 6A,B**), suggesting that OsCIPK18 might regulate the expression of ERF genes to further regulate downstream genes in response to ammonium toxicity (**Figure 8**). WRKY TFs in rice are rapidly induced in response to abiotic stresses, such as salinity, aluminum, drought, and osmotic stress, to regulate developmental processes, such as seed development, root growth, and leaf senescence (Ross et al., 2007). Three key genes (OsWRKY16, OsWRKY74, and OsWRKY96) belonging to the WRKY family were identified in WT-specific DEGs or *cipk18*-specific DEGs (**Figures 6A,B**). Among them, OsWRKY74 was reported to be involved in the response of Pi starvation, Fe starvation, as well as cold stress, and promote the elongation of primary and adventitious roots (Dai et al., 2016). In this study, inhibition of OsWRKY74

by excess NH_4^+ was lost in the mutant, suggesting that OsWRKY74 lies downstream of OsCIPK18 in regulating root growth under ammonium toxicity. In addition, members of the NAC, MYB, and other families were identified to be regulated by OsCIPK18 and involved in the inhibition of root growth in response to ammonium toxicity. These TFs and their interacting functional genes together form a response network of excess NH_4^+ regulated by OsCIPK18, which will contribute to the understanding of the process of ammonium toxicity production.

Role of OsCIPK18 in the Regulation of Auxin and ABA Signaling by Ammonium Toxicity

Phytohormones play an integral role in the regulatory network of response to ammonium toxicity in plants (Zou et al., 2013; Lei et al., 2016; Di et al., 2018; Sun et al., 2020). Excessive ammonium supply reduced free IAA content in the roots and significantly accelerated tissue ABA accumulation (Di et al., 2018; Sun et al., 2020). In the mutant *cipk18*, an auxin biosynthetic gene OsYUCCA6, an auxin influx transporter OsABCB14, and auxin-related response genes (OsARF5, OsSAUR12, OsSAUR21, and OsSAUR58) were specifically upregulated by excess NH_4^+ (**Figures 7A,B**). This demonstrated that both auxin synthesis and response pathway were activated in *cipk18*, suggesting that excessive ammonium suppressed root growth through direct or indirect regulation of the auxin synthesis gene by OsCIPK18, reducing IAA content and inhibiting the auxin response pathway (**Figure 8**). In the ABA signaling pathway, OsRePRP2.1 and OsRePRP2.2 were able to be activated by ABA and inhibited cell expansion (Tseng et al., 2013), and OsSAPK6 expression was significantly increased by ABA treatment (Chae et al., 2007). Under ammonium stress, OsSAPK6 was uniquely downregulated, and OsRePRP2.1/2.2 were no longer upregulated in *cipk18*, implying that knockdown of OsCIPK18 results in blocked ABA signaling, thereby reducing the inhibitory effect of ammonium toxicity on root growth (**Figures 7A,B**). These data reveal that OsCIPK18 is a key node in hormone signaling under ammonium stress, which provides great insight into the involvement of key phytohormones in rice in response to NH_4^+ stress.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: BioProject ID:PRJNA776549, https://www.ncbi.nlm.nih.gov/bioproject/PRJNA776549.

AUTHOR CONTRIBUTIONS

JinY designed the experiments. YL and XP directed the experiments. JinY and TS performed most of the experiments and analyses. YQ, GZ, and JiaY helped with the quantification of phenotypes. TS wrote the manuscript. JinY, TW, and HZ revised the manuscript. All authors discussed the results and contributed to the finalization of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the National Natural Science Foundation of China (Nos. 31760377, 31960124, and 32172074), the National Special Key Project for Transgenic Breeding (Grant No. 2016ZX08001001), and the National Key Research and Development Program of China (2016YFD0100400).

ACKNOWLEDGMENTS

We thank the reviewers and editor for their careful reading and helpful comments on this manuscript. We also thank the lab members for their assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.863283/full#supplementary-material

Supplementary Figure 1 | Identification of the T-DNA insertion site of *cipk18*. **(A)** A schematic map of the T-DNA insertion sites on the genomic regions of OsCIPK18. **(B)** Detection of the flanking sequence of the T-DNA insertion site in the mutant *via* PCR/RE assay.

Supplementary Figure 2 | The mutant *cipk18* shows resistance to excess NH_4^+. **(A–D)** Rice seedlings phenotypes **(A,B)**, root FW **(C)**, and radicle length **(D)** of 5-day-old WT, *cipk18,* and three complementary materials of *cipk18* (com1, com2, and com3) under treatments with ddH_2O and 4 mM NH_4Cl, respectively. Scale bar = 1 cm. **(E)** Compared with the control, relative root growth of the five lines under 4 mM NH_4Cl. Data are means ± SDs, significant differences using Student's *t*-test: $*p < 0.05$.

Supplementary Figure 3 | Differentially expressed genes between WT and *cipk18* cultured in ddH_2O. **(A)** Volcano plot of DEGs between WT and *cipk18*. **(B)** Venn diagram of DEGs between WT and *cipk18*, WT-specific DEGs, and *cipk18*-specific DEGs. **(C)** GO analysis of DEGs between WT and *cipk18*. **(D)** KEGG analysis of DEGs between WT and *cipk18*.

Supplementary Figure 4 | The expression of key TF genes regulated by OsCIPK18 under ammonium toxicity. Fold changes in the expression of genes under ammonium toxicity, as determined by qRT-PCR **(A)** and RNA-seq analysis **(B)**. Data are means ± SDs. Significant differences using Student's *t*-test: $*p < 0.05$.

Supplementary Table 1 | Expression of WT-specific differentially expressed genes.

Supplementary Table 2 | Expression of *cipk18*-specific differentially expressed genes.

Supplementary Table 3 | Predicted transcriptional regulatory networks in WT-specific DEGs under ammonium toxicity.

Supplementary Table 4 | Predicted transcriptional regulatory networks in *cipk18*-specific DEGs under ammonium toxicity.

Supplementary Table 5 | Primers used in this study.

REFERENCES

Alencar, V., Lobo, A. K. M., Carvalho, F. E. L., and Silveira, J. A. G. (2019). High ammonium supply impairs photosynthetic efficiency in rice exposed to excess light. *Photosyn. Res.* 140, 321–335. doi: 10.1007/s11120-019-00614-z

Balkos, K. D., Britto, D. T., and Kronzucker, H. J. (2010). Optimization of ammonium acquisition and metabolism by potassium in rice (Oryza sativa L. cv. IR-72). *Plant Cell Environ.* 33, 23–34. doi: 10.1111/j.1365-3040.2009.02046.x

Banuelos, M. A., Garciadeblas, B., Cubero, B., and Rodríguez-Navarro, A. (2002). Inventory and functional characterization of the HAK potassium transporters of rice. *Plant Physiol.* 130, 784–795. doi: 10.1104/pp.007781

Beier, M. P., Obara, M., Taniai, A., Sawa, Y., Ishizawa, J., Yoshida, H., et al. (2018). Lack of ACTPK1, an STY kinase, enhances ammonium uptake and use, and promotes growth of rice seedlings under sufficient external ammonium. *Plant J.* 93, 992–1006. doi: 10.1111/tpj.13824

Bittsánszky, A., Pilinszky, K., Gyulai, G., and Komives, T. (2015). Overcoming ammonium toxicity. *Plant Sci.* 231, 184–190. doi: 10.1016/j.plantsci.2014.12.005

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Britto, D. T., Balkos, K. D., Becker, A., Coskun, D., Huynh, W. Q., and Kronzucker, H. J. (2014). Potassium and nitrogen poising: physiological changes and biomass gains in rice and barley. *Canad. J. Plant Sci.* 94, 1085–1089. doi: 10.4141/cjps2013-143

Britto, D. T., and Kronzucker, H. J. (2002). NH_4^+ toxicity in higher plants: a critical review. *J. Plant Physiol.* 159, 567–584. doi: 10.1078/0176-1617-0774

Chae, M.-J., Lee, J.-S., Nam, M.-H., Cho, K., Hong, J.-Y., Yi, S.-A., et al. (2007). A rice dehydration-inducible SNF1-related protein kinase 2 phosphorylates an abscisic acid responsive element-binding factor and associates with ABA signaling. *Plant Mol. Biol.* 63, 151–169. doi: 10.1007/s11103-006-9079-x

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 13, 1194–1202. doi: 10.1016/j.molp.2020.06.009

Chen, G., Guo, S., Kronzucker, H. J., and Shi, W. (2013). Nitrogen use efficiency (NUE) in rice links to NH_4^+ toxicity and futile NH_4^+ cycling in roots. *Plant Soil* 369, 351–363. doi: 10.1007/s11104-012-1575-y

Cheng, X., Sardana, R. K., and Altosaar, I. (1998). "Rice transformation by agrobacterium infection," in *Recombinant Proteins from Plants, Vol. 3*, eds C. Cunningham and A. J. R. Porter (Totowa: Humana Press) 1–9. doi: 10.1007/978-1-60327-260-5_1

Dai, X., Wang, Y., Yang, A., and Zhang, W. H. (2012). OsMYB2P-1, an R2R3 MYB transcription factor, is involved in the regulation of phosphate-starvation responses and root architecture in rice. *Plant Physiol.* 159, 169–183. doi: 10.1104/pp.112.194217

Dai, X., Wang, Y., and Zhang, W. H. (2016). OsWRKY74, a WRKY transcription factor, modulates tolerance to phosphate starvation in rice. *J. Exp. Bot.* 67, 947–960. doi: 10.1093/jxb/erv515

Dave, G., and Nilsson, E. (2005). Increased reproductive toxicity of landfill leachate after degradation was caused by nitrite. *Aquatic Toxicol.* 73, 11–30. doi: 10.1016/j.aquatox.2005.02.006

Di, D., Sun, L., Zhang, X., Li, G., Kronzucker, H. J., and Shi, W. (2018). Involvement of auxin in the regulation of ammonium tolerance in rice (Oryza sativa L.). *Plant Soil* 432, 373–387. doi: 10.1007/s11104-018-3813-4

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. doi: 10.1093/bioinformatics/bts635

Esteban, R., Ariz, I., Cruz, C., and Moran, J. F. (2016). Review: mechanisms of ammonium toxicity and the quest for tolerance. *Plant Sci.* 248, 92–101. doi: 10.1016/j.plantsci.2016.04.008

Forde, B. G. (2002). Local and long-range signaling pathways regulating plant responses to nitrate. *Annu. Rev. Plant Biol.* 53, 203–224. doi: 10.1146/annurev.arplant.53.100301.135256

Gerendás, J., Zhu, Z., Bendixen, R., Ratcliffe, R. G., and Sattelmacher, B. (1997). Physiological and biochemical processes related to ammonium toxicity in higher plants. *J. Plant Nutr. Soil Sci.* 160, 239–251. doi: 10.1002/jpln.19971600218

Gu, M., Zhang, J., Li, H., Meng, D., Li, R., Dai, X., et al. (2017). Maintenance of phosphate homeostasis and root development are coordinately regulated by MYB1, an R2R3-type MYB transcription factor in rice. *J. Exp. Bot.* 68, 3603–3615. doi: 10.1093/jxb/erx174

Hachiya, T., Inaba, J., Wakazaki, M., Sato, M., Toyooka, K., Miyagi, A., et al. (2021). Excessive ammonium assimilation by plastidic glutamine synthetase causes ammonium toxicity in *Arabidopsis thaliana*. *Nat. Commun.* 12, 4944. doi: 10.1038/s41467-021-25238-7

Hachiya, T., Watanabe, C. K., Fujimoto, M., Ishikawa, T., Takahara, K., Kawai-Yamada, M., et al. (2012). Nitrate addition alleviates ammonium toxicity without lessening ammonium accumulation, organic acid depletion and inorganic cation depletion in *Arabidopsis thaliana* shoots. *Plant Cell Physiol.* 53, 577–591. doi: 10.1093/pcp/pcs012

Hashimoto, K., and Kudla, J. (2011). Calcium decoding mechanisms in plants. *Biochimie* 93, 2054–2059. doi: 10.1016/j.biochi.2011.05.019

Hirano, T., Satoh, Y., Ohki, A., Takada, R., Arai, T., and Michiyama, H. (2008). Inhibition of ammonium assimilation restores elongation of seminal rice roots repressed by high levels of exogenous ammonium. *Physiol. Plant.* 134, 183–190. doi: 10.1111/j.1399-3054.2008.01117.x

Hu, H. C., Wang, Y. Y., and Tsay, Y. F. (2009). AtCIPK8, a CBL-interacting protein kinase, regulates the low-affinity phase of the primary nitrate response. *Plant J.* 57, 264–278. doi: 10.1111/j.1365-313X.2008.03685.x

Huang, D., Wang, S., Zhang, B., Shang-Guan, K., Shi, Y., Zhang, D., et al. (2015). A gibberellin-mediated DELLA-NAC signaling cascade regulates cellulose synthesis in rice. *Plant Cell* 27, 1681–1696. doi: 10.1105/tpc.15.00015

Jain, M., Tyagi, A. K., and Khurana, J. P. (2006). Genome-wide analysis, evolutionary expansion, and expression of early auxin-responsive SAUR gene family in rice (Oryza sativa). *Genomics* 88, 360–371. doi: 10.1016/j.ygeno.2006.04.008

Jeon, J., Lee, S., Jung, K., Jun, S., Jeong, D., Lee, J., et al. (2000). T-DNA insertional mutagenesis for functional genomics in rice. *Plant J.* 22, 561–570. doi: 10.1046/j.1365-313x.2000.00767.x

Kikuchi, K., Ueguchi-Tanaka, M., Yoshida, K. T., Nagato, Y., and Hirano, H. Y. (2000). Molecular analysis of the NAC gene family in rice. *Mol. General Genet.* 262, 1047–1051. doi: 10.1007/PL00008647

Kim, H. Y., Choi, E.-H., Min, M. K., Hwang, H., Moon, S.-J., Yoon, I., et al. (2015). Differential gene expression of two outward-rectifying shaker-like potassium channels OsSKOR and OsGORK in rice. *J. Plant Biol.* 58, 230–235. doi: 10.1007/s12374-015-0070-4

Kobayashi, Y., Yamamoto, S., Minami, H., and Hattori, K. T. (2004). Differential activation of the rice sucrose Nonfermenting1-related protein Kinase2 family by hyperosmotic stress and abscisic acid. *Plant Cell* 16, 1163–1177. doi: 10.1105/tpc.019943

Konishi, N., and Ma, J. F. (2021). Three polarly localized ammonium transporter 1 members are cooperatively responsible for ammonium uptake in rice under low ammonium condition. *New Phytol.* 232, 1778–1792. doi: 10.1111/nph.17679

Kronzucker, H. J., Britto, D. T., Davenport, R. J., and Tester, M. (2001). Ammonium toxicity and the real cost of transport. *Trends Plant Sci.* 6, 335–337. doi: 10.1016/S1360-1385(01)02022-2

Krouk, G., Lacombe, B., Bielach, A., Perrine-Walker, F., Malinska, K., Mounier, E., et al. (2010). Nitrate-regulated auxin transport by NRT1.1 defines a mechanism for nutrient sensing in plants. *Dev. Cell* 18, 927–937. doi: 10.1016/j.devcel.2010.05.008

Kudoyarova, G. R., Farkhutdinov, R. G., and Veselov, S. Y. (1997). Comparison of the effects of nitrate and ammonium forms of nitrogen on auxin content in roots and the growth of plants under different temperature conditions. *Plant Growth Regul.* 23, 207–208. doi: 10.1023/A:1005990725068

Kumar, V., Kim, S. H., Priatama, R. A., Jeong, J. H., Adnan, M. R., Saputra, B. A., et al. (2020). NH_4^+ suppresses NO_3^--dependent lateral root growth and alters gene expression and gravity response in OsAMT1 RNAi mutants of rice (Oryza sativa). *J. Plant Biol.* 63, 391–407. doi: 10.1007/s12374-020-09263-5

Lei, D., Li, Y., Ying, W., Gao, L., Wang, M., Francois, C., et al. (2016). Root ABA accumulation enhances rice seedling drought tolerance under

ammonium supply: interaction with aquaporins. *Front. Plant Sci.* 7, 1206. doi: 10.3389/fpls.2016.01206

Li, B., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12, 323. doi: 10.1186/1471-2105-12-323

Li, B., Li, G., Kronzucker, H. J., Baluska, F., and Shi, W. (2014). Ammonium stress in Arabidopsis: signaling, genetic loci, and physiological targets. *Trends Plant Sci.* 19, 107–114. doi: 10.1016/j.tplants.2013.09.004

Li, B., Li, Q., Xiong, L., Kronzucker, H. J., Kramer, U., and Shi, W. (2012). Arabidopsis plastid AMOS1/EGY1 integrates abscisic acid signaling to regulate global gene expression response to ammonium stress. *Plant Physiol.* 160, 2040–2051. doi: 10.1104/pp.112.206508

Liu, Y., Maniero, R. A., Giehl, R. F. H., Melzer, M., Steensma, P., Krouk, G., et al. (2022). PDX1.1-dependent biosynthesis of vitamin B6 protects roots from ammonium-induced oxidative stress. *Mol. Plant*. 15, 1–20. doi: 10.1016/j.molp.2022.01.012

Liu, Y., and Wirén, N. V. (2017). Ammonium as a signal for physiological and morphological responses in plants. *J. Exp. Bot.* 68, 2581–2592. doi: 10.1093/jxb/erx086

Lobet, G., and Draye, L. P. (2011). A novel image-analysis toolbox enabling quantitative analysis of root system architecture. *Plant Physiol.* 157, 29–39. doi: 10.1104/pp.111.179895

Meier, M., Liu, Y., Lay-Pruitt, K. S., Takahashi, H., and von Wirén, N. (2020). Auxin-mediated root branching is determined by the form of available nitrogen. *Nat. Plants* 6, 1136–1145. doi: 10.1038/s41477-020-00756-2

Miflin, B. J., and Habash, D. Z. (2002). The role of glutamine synthetase and glutamate dehydrogenase in nitrogen assimilation and possibilities for improvement in the nitrogen utilization of crops. *J. Exp. Bot.* 53, 979–987. doi: 10.1093/jexbot/53.370.979

Miller, A. J., and Cramer, M. D. (2005). Root nitrogen acquisition and assimilation. *Plant Soil* 274, 1–36. doi: 10.1007/1-4020-4099-7_1

Mizoi, J., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2012). AP2/ERF family transcription factors in plant abiotic stress responses. *Biochim. Biophys. Acta* 1819, 86–96. doi: 10.1016/j.bbagrm.2011.08.004

Romano, N., and Zeng, C. (2007). Acute toxicity of ammonia and its effects on the haemolymph osmolality, ammonia-N, pH and ionic composition of early juvenile mud crabs, Scylla serrata (Forskl). *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* 148, 278–285. doi: 10.1016/j.cbpa.2007.04.018

Ross, C. A., Liu, Y., and Shen, Q. J. (2007). The WRKY gene family in rice (Oryza sativa). *J. Integr. Plant Biol.* 49, 827–842. doi: 10.1111/j.1744-7909.2007.00504.x

Sánchez-Barrena, M. J., Martínez-Ripoll, M., Zhu, J.-K., and Albert, A. (2005). The structure of the *Arabidopsis thaliana* SOS3: molecular mechanism of sensing calcium for salt stress response. *J. Mol. Biol.* 345, 1253–1264. doi: 10.1016/j.jmb.2004.11.025

Straub, T., Ludewig, U., and Neuhauser, B. (2017). The kinase CIPK23 inhibits ammonium transport in *Arabidopsis thaliana*. *Plant Cell* 29, 409–422. doi: 10.1105/tpc.16.00806

Sun, L., Di, D., Li, G., Kronzucker, H. J., and Shi, W. (2017). Spatio-temporal dynamics in global rice gene expression (Oryza sativa L.) in response to high ammonium stress. *J. Plant Physiol.* 212, 94–104. doi: 10.1016/j.jplph.2017.02.006

Sun, L., Di, D. W., Li, G., Kronzucker, H. J., Wu, X., and Shi, W. (2020). Endogenous ABA alleviates rice ammonium toxicity by reducing ROS and free ammonium via regulation of the SAPK9-bZIP20 pathway. *J. Exp. Bot.* 71, 4562–4577. doi: 10.1093/jxb/eraa076

Tian, T., Liu, Y., Yan, H., You, Q., Yi, X., Du, Z., et al. (2017). agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res.* 45, W122–W129. doi: 10.1093/nar/gkx382

Tseng, I. C., Hong, C. Y., Yu, S. M., and Ho, T. (2013). Abscisic acid- and stress-induced highly proline-rich glycoproteins regulate root growth in rice. *Plant Physiol.* 163, 118–134. doi: 10.1104/pp.113.217547

Wang, D., Pei, K., Fu, Y., Sun, Z., Li, S., Liu, H., et al. (2007). Genome-wide analysis of the auxin response factors (ARF) gene family in rice (*Oryza sativa*). *Gene* 394, 13–24. doi: 10.1016/j.gene.2007.01.006

Xu, J., Li, H.-D., Chen, L.-Q., Wang, Y., Liu, L.-L., He, L., et al. (2006). A protein kinase, interacting with two calcineurin B-like proteins, regulates K^+ transporter AKT1 in arabidopsis. *Cell* 125, 1347–1360. doi: 10.1016/j.cell.2006.06.011

Xu, Y., Zhang, S., Guo, H., Wang, S., Xu, L., Li, C., et al. (2014). OsABCB14 functions in auxin transport and iron homeostasis in rice (*Oryza sativa* L.). *Plant J.* 79, 106–117. doi: 10.1111/tpj.12544

Xuan, Y. H., Kumar, V., Han, X., Kim, S. H., Jeong, J. H., Kim, C. M., et al. (2019). CBL-INTERACTING PROTEIN KINASE 9 regulates ammonium-dependent root growth downstream of IDD10 in rice (*Oryza sativa*). *Ann. Bot.* 124, 947–960. doi: 10.1093/aob/mcy242

Yamamoto, Y., Kamiya, N., Morinaka, Y., Matsuoka, M., and Sazuka, T. (2007). Auxin biosynthesis by the YUCCA genes in rice. *Plant Physiol.* 143, 1362–1371. doi: 10.1104/pp.106.091561

Yao, X., Horie, T., Xue, S., Leung, H. Y., Katsuhara, M., Brodsky, D. E., et al. (2010). Differential sodium and potassium transport selectivities of the rice OsHKT2;1 and OsHKT2;2 transporters in plant cells. *Plant Physiol.* 152, 341–355. doi: 10.1104/pp.109.145722

Yong, H. C., Pandey, G. K., Grant, J. J., Batistic, O., Li, L., Kim, B. G., et al. (2010). Two calcineurin B-like calcium sensors, interacting with protein kinase CIPK23, regulate leaf transpiration and root potassium uptake in Arabidopsis. *Plant J.* 52, 223–239. doi: 10.1111/j.1365-313X.2007.03236.x

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287. doi: 10.1089/omi.2011.0118

Yutaka, S., Akira, I., Satomi, S., von, W. N., Tomoyuki, Y., and Junji, Y. (2003). Distinct expression and function of three ammonium transporter genes (OsAMT1;1 – 1;3) in rice. *Plant Cell Physiol.* 7, 726–734. doi: 10.1093/pcp/pcg083

Zhang, M., Wang, Y., Chen, X., Xu, F., Ding, M., Ye, W., et al. (2021). Plasma membrane H^+-ATPase overexpression increases rice yield via simultaneous enhancement of nutrient uptake and photosynthesis. *Nat. Commun.* 12, 735. doi: 10.1038/s41467-021-20964-4

Zhang, T., You, J., Zhang, Y., Yao, W., Chen, W., Duan, Q., et al. (2021). LF1 regulates the lateral organs polarity development in rice. *New Phytol.* 231, 1265–1277. doi: 10.1111/nph.17220

Zhao, Q., Zhang, H., Wang, T., Chen, S., and Dai, S. (2013). Proteomics-based investigation of salt-responsive mechanisms in plant roots. *J. Proteomics* 82, 230–253. doi: 10.1016/j.jprot.2013.01.024

Zhu, Y., Di, T., Xu, G., Chen, X., Zeng, H., Yan, F., et al. (2009). Adaptation of plasma membrane $H(^+)$-ATPase of rice roots to low pH as related to ammonium nutrition. *Plant Cell Environ.* 32, 1428–1440. doi: 10.1111/j.1365-3040.2009.02009.x

Zou, N., Li, B., Chen, H., Su, Y., Kronzucker, H. J., Xiong, L., et al. (2013). GSA-1/ARG1 protects root gravitropism in Arabidopsis under ammonium stress. *New Phytol.* 200, 97–111. doi: 10.1111/nph.12365

Zou, N., Li, B., Dong, G., Kronzucker, H. J., and Shi, W. (2012). Ammonium-induced loss of root gravitropism is related to auxin distribution and TRH1 function, and is uncoupled from the inhibition of root elongation in Arabidopsis. *J. Exp. Bot.* 63, 3777–3788. doi: 10.1093/jxb/ers068

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 04 May 2022
doi: 10.3389/fpls.2022.817106

NtNAC053, A Novel NAC Transcription Factor, Confers Drought and Salt Tolerances in Tobacco

Xiaoxu Li[1,2†], Qi Wang[2,3†], Cun Guo[2,4†], Jinhao Sun[2], Zhiyuan Li[2], Yaofu Wang[1], Aiguo Yang[2], Wenxuan Pu[1], Yongfeng Guo[2], Junping Gao[1] and Liuying Wen[2*]*

[1] Technology Center, China Tobacco Hunan Industrial Co., Ltd., Changsha, China, [2] Key Laboratory for Tobacco Gene Resources, Tobacco Research Institute, Chinese Academy of Agricultural Sciences, Qingdao, China, [3] Shandong Peanut Research Institute, Qingdao, China, [4] Graduate School of Chinese Academy of Agricultural Sciences, Beijing, China

The NAC (NAM, ATAF1/2, and CUC2) family acts as one of the largest families of the transcription factor in the plant kingdom and was revealed to function as the important regulators in various environmental stresses. However, a few studies were reported about the biofunctions of the NAC transcription factor in tobacco. In the current study, we characterized a novel NAC transcription factor encoding the gene *NtNAC053* in tobacco, which was significantly up-regulated when exposed to salt and drought treatments. The results of *cis*-acting elements analysis suggested that the promoter region of *NtNAC053* gene possesses a number of stress-responsive elements, and this gene could be induced by exogenous abscisic acid (ABA) treatment. Moreover, the NtNAC053–GFP fusion protein was localized in the cell nucleus and possessed a transactivation domain in its C-terminal, implying that NtNAC053 may undertake as a transcriptional activator in tobacco. Notably, the overexpression of *NtNAC053* in tobacco resulted in hypersensitivity to ABA treatment. Furthermore, these overexpression lines showed significantly enhanced tolerances to drought and salt stresses. Under salt and drought stresses, these overexpression lines possessed higher superoxide dismutase (SOD), catalase (CAT), and peroxidase (POD) activities. Interestingly, the expressions of putative stress-related genes, including *NtCOR15A*, *NtRAB18*, *NtDREB1A*, *NtERF5*, *NtKAT2*, and *NtERD11*, were up-regulated in these overexpression lines when subjected to salt and drought stresses. The clues provided in our study suggested that the *NtNAC053* gene encodes a novel NAC transcription factor and could confer the drought and salt stress tolerances by inspiring the downstream stress-responsive genes and antioxidant system in tobacco.

Keywords: tobacco, NtNAC053, ABA, abiotic stress, antioxidant system

OPEN ACCESS

Edited by:
Mukesh Jain,
Jawaharlal Nehru University, India

Reviewed by:
X. Deng,
Rubber Research Institute, Chinese Academy of Tropical Agricultural Sciences, China
Zhan Qi Wang,
Huzhou University, China

****Correspondence:***
Junping Gao
gaojp0104@hngytobacco.com
Liuying Wen
wenliuying@caas.cn

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

Received: *17 November 2021*
Accepted: *21 March 2022*
Published: *04 May 2022*

Citation:
Li X, Wang Q, Guo C, Sun J, Li Z, Wang Y, Yang A, Pu W, Guo Y, Gao J and Wen L (2022) NtNAC053, A Novel NAC Transcription Factor, Confers Drought and Salt Tolerances in Tobacco. Front. Plant Sci. 13:817106. doi: 10.3389/fpls.2022.817106

INTRODUCTION

Transcription factors play a pivotal role in regulating the expression of target genes involved in plant growth, development, and response to environmental stresses (Qu and Zhu, 2006). The NAC protein family has been reported to be one of the largest plant-specific transcription factor families. Its name is based on three transcription factors: no apical meristem (NAM), *Arabidopsis*

transcription activation factor (ATAF1,2), and cup-shaped cotyledon (CUC) (Souer et al., 1996; Aida et al., 1997). NAC transcription factors harbor a highly conserved DNA-binding domain at the N-terminus and a variable transcription regulation region (TRR) at the C-terminus (Ernst et al., 2004; Olsen et al., 2005). Generally, the N-terminal DNA-binding domain of NAC protein possesses approximately 150 amino acid residues divided into five subdomains (A-E), which are related to the formation of dimer, functional diversity, and DNA-binding activity (Olsen et al., 2005). The C-terminal TRR of NAC protein functions in conferring regulation diversity of transcriptional activation activity (Yamaguchi et al., 2010; Puranik et al., 2011). Furthermore, the transmembrane domain is also found in several NAC proteins, which function in plasma membrane anchoring (Kim et al., 2010).

The *NAC* gene was first identified from petunias, in which it was involved in the shoot apical meristems development and primordium formation (Souer et al., 1996). In the plant kingdom, the NAC proteins have been reported to play important roles in various biological processes, including lateral root development (Xie et al., 2000; He et al., 2005; Zhang et al., 2018), vascular development (Zhao et al., 2016), development of shoot apical meristem (Takada et al., 2001; Larsson et al., 2012), the formation of the adventitious shoot (Hibara et al., 2003), seed germination (Park et al., 2011), flower formation (Hendelman et al., 2013), secondary wall synthesis (Mitsuda et al., 2007; Zhao et al., 2010), and leaf senescence (Guo and Gan, 2006). So far, the NAC family has been identified in a large number of plants, including *Arabidopsis* (105) (Ooka et al., 2003), tobacco (*Nicotiana tabacum*) (154) (Li et al., 2018), potato (*Solanum tuberosum*) (110) (Singh et al., 2013), white pear (*Pyrus bretschneideri*) (183) (Gong et al., 2019), maize (*Zea mays*) (152) (Shiriga et al., 2014), rice (*Oryza sativa*) (151) (Nuruzzaman et al., 2010), and peanut (*Arachis hypogaea*) (132) (Li P. et al., 2021).

Furthermore, several NAC members have been reported to participate in response to abiotic stresses in *Arabidopsis* and other species. Stress-inducible NAC transcription factors ANAC019, ANAC055, and ANAC072/RD26 were found to improve the drought tolerance *via* regulating the expression of the *ERD1* gene (Tran et al., 2004). Furthermore, *ANAC072* was identified to be involved in the ABA-dependent signaling pathway (Fujita et al., 2004). Notably, the expression of *ANAC002*/*ATAF1* was induced by drought and ABA treatments. Moreover, the *ataf1* knockout mutants displayed a lower survival ratio than wild-type plants under drought conditions, indicating that ATAF1 functions as a regulator in drought stress tolerance (Wu et al., 2009). The *ANAC062*/*NTL6* was reported to play an important role in inducing a drought-resistance response and the overexpression of *NTL6* plants exhibit enhanced resistance to water-deficit conditions (Kim et al., 2012). In addition, the loss of *ANAC035*/*LOV1* function results in hypersensitivity to cold temperature, whereas overexpression of this gene could enhance the cold tolerance (Yoo et al., 2007). In rice, the expression of *OsNAC5* gene could be significantly induced by various stresses, including drought, salt, cold, ABA, and Me-JA. Whereas the overexpression of *OsNAC5* resulted in the improvement of salt stress tolerance (Takasaki et al., 2010). Moreover, the ONAC066 was reported to promote the tolerances against drought and oxidative stresses by activating the transcription of *OsDREB2A* in transgenic rice (Yuan et al., 2019). In wheat, the overexpression of *TaNAC47* in Arabidopsis increases the tolerance to drought stress and sensitivity to ABA (Zhang et al., 2015). Besides, the overexpression of *TaNAC2a* improved drought tolerance in transgenic tobacco (Tang et al., 2012).

Tobacco is one of the most widely cultivated crops worldwide and serviced as a model plant that possesses various primary and secondary metabolisms. However, tobacco has been at a continuous risk from multiple environmental stresses, affecting both yield and quality. In our previous study, the systematic identification of the NAC family was performed in tobacco (Li et al., 2018). However, a few studies were reported about the biofunctions of the NAC transcription factor in tobacco. Notably, in our previous study, the expression of *StNAC053* was significantly up-regulated after salt and drought treatments. The overexpression of this potato gene could enhance the salt and drought tolerance in transgenic *Arabidopsis* (Wang et al., 2021). Considering the close genetic distance of potato and tobacco, the homolog of StNAC053 might also confer salt and drought tolerances in tobacco. To verify this hypothesis, we cloned and characterized an NAC transcription factor, NtNAC053, from tobacco, and systematic analysis was performed to explore the function of this gene in salt and drought stress responses.

MATERIALS AND METHODS

Plant Materials and Treatments

Cultivated tobacco K326 was used to analyze the expression pattern of *NtNAC053* in the current study. The root, root tip, stem, stem tip, young leaf, and senescent leaf were collected. For salt, drought, and ABA treatments, tobacco seedlings were transferred to MS liquid medium supplemented with 100 mM NaCl, 300 mM mannitol, and 10 μM ABA for 0, 1, 3, and 6 h, respectively. For cold treatment, tobacco seedlings were treated at 4°C for 0, 1, 3, and 6 h. All of the harvested samples were stored at −80°C before RNA extraction was performed. Three biological replicates were used for each sample.

RNA Extraction and qPCR Analysis

Ultrapure RNA Kit (cwbiotech, Beijing, China) was used to extract total RNA, then the first cDNA strand was synthesized using Evo M-MLV Mix Kit with gDNA Clean for qPCR (Accurate Biotechnology, Changsha, China). Expression of stress-responsive genes (*NtDREB1A*, *NtERF5*, *NtKAT2*, *NtCOR15A*, *NtRAB18*, *NtERD11*, *NtNHX1*, and *NtSOS1*) and ABA signaling genes from tobacco was detected. The qRT-PCR reactions were carried out in Roche LightCycler 480 Real-Time PCR instrument with SYBR Green Premix Pro Taq HS qPCR Kit (Accurate Biotechnology, Changsha, China). The tobacco ribosomal protein gene *L25* (GenBank No. L18908) was used as a control (Li Z. et al., 2021). All experimental data were obtained through three technical repetitions, and the relative expression level was calculated by $2^{-\Delta\Delta CT}$ method

(Livak and Schmittgen, 2001). The details of the primers in this study are provided in **Supplementary Table 1**.

Cloning and Sequence Analysis of *NtNAC053*

BLASTP search with an *E*-value cutoff of 0.0001 was carried out in SGN[1] using potato StNAC053 protein sequence as a query. The candidate gene (Nitab4.5_0000322g0110.1) was obtained and then named as *NtNAC053*. The CDS of *NtNAC053* was amplified by RT-PCR. The purified PCR products were subcloned into the pEASY-Blunt vector, and the constructed vector was then transformed into *Escherichia coli* 5α competent cells (Trans Gen, Beijing, China). After sequencing, the recombinant plasmid was named *Blunt-NtNAC053*. The physicochemical properties of the NtNAC053 protein, including molecular weight (MW) and isoelectric point (pI), were predicated using the ProtParam online tool[2]. The nuclear localization signal (NLS) of NAC proteins was predicted *via* cNLS Mapper[3]. The DNAMAN software (Lynnon Biosoft) was used to perform multiple sequence alignment. A neighbor-joining tree was generated using MEGA 6 with 1,000 replications. The NAC protein sequences were obtained from the TAIR[4], SGN (see text footnote 1), and NCBI. The conserved motifs of NAC protein sequences were analyzed *via* MEME[5] with default parameters. The collinearity relationship between NtNAC053 and StNAC053 was analyzed using TBtools (Chen et al., 2020).

To assess the promoter *cis*-acting elements of the *NtNAC053* gene, the sequences 2,000 bp upstream of the *NtNAC053* gene were extracted from the genome database. The obtained sequences were subjected to PlantCARE[6] platform analysis to further search for the putative *cis*-elements in their promoter regions.

Subcellular Localization and Transactivation Assay

The CDS (without stop codon) of *NtNAC053* gene was amplified from *Blunt-NtNAC053* using PCR. Then the purified PCR products was inserted into the PYG57 vector, which started by the CaMV-35S promoter and contained the GFP fragment. The construct was transformed into *Agrobacterium* competent cell GV3101, and transiently expressed in the leaves of *Nicotiana benthamiana*. Simultaneously, the empty vector injected leaves as a control. Three days after the injection, the leaves were soaked in a DAPI staining solution to determine the location of the nucleus. The confocal laser microscope (Leica SP8) was used to observe the fluorescence signal of the fusion protein.

The CDS of *NtNAC053* was amplified and inserted into the *Eco*RI site of the pBridge vector, using an In-fusion HD Cloning Kit (Takara, Shiga, Japan) to fuse with a GAL4 DNA binding domain. The constructs and the control vector were introduced into the yeast strain AH109 separately, followed by growing yeasts on SD/-Trp, and SD/-Trp supplemented with X-Gal for 4 days at 30°C. The transcriptional activation activities were evaluated based on the growth status of different transformants.

Generation of *NtNAC053* Transgenic Tobacco Plants

The coding sequence of *NtNAC053* gene was amplified from the cDNA and inserted into the pCHF3 vector, which was driven by the CaMV-35S promoter, to complete the construction of the overexpression vector. The pCHF3 plasmid containing the *NtNAC053* gene was transformed into tobacco leaves by the *Agrobacterium*-mediated method (Buschmann, 2016). T0 transgenic seeds were screened on MS medium containing 50 mg/L kanamycin to obtain *NtNAC053* overexpression plants. T3 generation homozygous seedlings were used for further experiments.

Seedling Growth Assays

For seed germination assays, T3 generation seeds and wild-type (K326) seeds were sterilized and evenly sown on MS medium containing 1 μM ABA, 100 mM NaCl, and 150 mM NaCl, respectively. After cultivation for 14 days in a growth chamber (25°C, continuous light), the germination rates were calculated with three replications.

For root elongation assays, tobacco seedlings were grown in vertical MS medium for 14 days, then transferred to MS medium with NaCl (100 and 150 mM) and ABA (1 μM), respectively. The length of the primary root was measured after 14 days with three replications.

Drought Stress Tolerance Assays of *NtNAC053* Transgenic Tobacco Plants

For the drought tolerance assays, T3 transgenic and wild-type (K326) tobacco seeds were sown into the soil mixture evenly and kept in normal conditions. Six weeks later, the tobacco seedlings were treated by drought stress for 10 days. Then the seedlings were re-watered for 3 days, and the survival rates of tobacco seedlings were calculated. The leaves of the seedlings were collected and air-dried for 3 h, then leaf weight was recorded at five time points (0, 45, 90, 135, and 180 min). The rate of water loss was calculated based on the leaves' weight changes.

Physiological Measurements

The physiological parameters, including contents of MDA and hydrogen peroxide, and the enzymatic activities of superoxide (SOD), peroxidase (POD), and catalase (CAT) were measured as previously described (He et al., 2019a). Three biological replicates were performed.

Yeast One Hybrid Assay

For yeast one-hybrid assay, the full-length CDS sequence of *NtNAC053* transcription factor was inserted into pGAD424 to construct a *pGAD424-NtNAC053* recombinant vector. Furthermore, the *pGAD424-NtNAC053* recombinant vector was co-transformed with different *LacZ* reporter gene constructs

[1]http://solgenomics.net/

[2]https://web.expasy.org/protparam/

[3]https://www.novopro.cn/tools/nls-signal-prediction.html

[4]https://www.arabidopsis.org/

[5]http://meme-suite.org/

[6]http://bioinformatics.psb.ugent.be/webtools/plantcare/html/

containing three tandem reporter sequences into yeast strain YM4271. The β-galactosidase activity was assessed according to the manufacturer's instruction (Clontech, United States).

Statistical Analysis

The GraphPad Prism 8 (*t*-test) was used to analyze the statistical significance. $P < 0.05$, 0.01, or 0.001 were regarded as significantly different from the control. All data were obtained from three replicates.

RESULTS

Isolation and Characterization of the *NtNAC053* Gene

The NAC transcription factors were reported to play multiple roles in response to abiotic stresses. In our previous study, the StNAC053 was found to confer salt and drought tolerance in transgenic *Arabidopsis* (Wang et al., 2021). To identify the homolog of StNAC053 in tobacco, the BLASTP search was carried out in SGN (see text footnote 1) using the StNAC053 protein sequence as a query. The candidate gene (Nitab4.5_0000322g0110.1) was obtained and then named as *NtNAC053*. Furthermore, the *NtNAC053* gene was cloned from cultivated tobacco K326. The full-length coding sequence of *NtNAC053* is 897 bp, which encodes a 298 amino acid protein. The molecular weight (MW) and theoretical isoelectric point (pI) of the protein are 34.247 kDa and 6.56, respectively. Multiple sequence alignments persuaded that NtNAC053 possesses a conserved NAC domain (8-162) in the N-terminal region and a variable domain at the C-terminal region. The NAC domain consists of five subdomains, including subdomain A-E, which have been predicted to function as the DNA-binding domain together. Furthermore, a novel nuclear localization signal (NLS) was identified in the subdomain D (**Supplementary Figure 1**).

Phylogenetic and Syntenic Analysis of the NtNAC053

To explore the phylogenetic relationship of NtNAC053, a neighbor-joining (NJ) tree containing the selected novel NAC transcription factor, including seven members from *Arabidopsis*, three from potato, and two from wheat. The results persuaded that NtNAC053 was clustered with StNAC053, and both of these two members fall into the ATAF subfamily (**Supplementary Figure 2A**). Furthermore, the conserved motifs of NtNAC053 and these selected NAC members were predicted by the MEME tools. As a result, 15 conserved motifs were identified (**Supplementary Figure 3**). Notably, motif 3, 1, 2, and 4 represented the NAC domain and were shared by all tested members (**Supplementary Figure 2B**). Besides, it was found that motif 12 may be specific to the dicotyledonous members of ATAF subfamily.

To further study the evolutionary relationship of *NtNAC053* gene, syntenic analysis was performed for tobacco and potato (**Supplementary Figure 2C**). The result showed that *NtNAC053* was the ortholog gene of *StNAC053*.

Expression Patterns and Promoter Analysis of *NtNAC053*

To explore the potential function of NtNAC053, the *cis*-acting elements and the expression pattern of *NtNAC053* in different tissues were investigated. Interestingly, the *cis*-acting elements analysis revealed that five elements related to ABA responsiveness (ABRE) were identified in the promoter region of the *NtNAC053* gene (**Supplementary Figure 4**). As shown in **Figure 1A**, the expression of *NtNAC053* had a higher level in tobacco root and young leaves, and had a lower level in the root tip, stem, and stem tip. Interestingly, the expression of *NtNAC053* had a high level in senescent leaf. Furthermore, we examined the expression pattern of *NtNAC053* under various treatments, including drought, salt, cold, and ABA. For the drought and ABA treatments, the expression levels of *NtNAC053* were up-regulated before the first 6 h, respectively (**Figures 1B,D**). For the salt treatment, the expression levels of *NtNAC053* were also up-regulated and reached the peak at 3 h (**Figure 1E**). For the cold treatment, no significant alteration was detected (**Figure 1C**).

Subcellular Localization and Transactivation Activity Assay

To further identify the function of NtNAC053, the subcellular localization was investigated. Generally, the coding sequence of *NtNAC053* gene was fused with GFP and transiently expressed in the *N. benthamiana* leaves. As a result, the fluorescence signals of 35S:GFP control displayed as spreading throughout the cells (**Figure 2**). Meanwhile, the green fluorescent signal of the fused *NtNAC053-GFP* was only observed in the nucleus. This result illustrated that NtNAC053 protein was a nuclear-localized protein.

To determine the transcriptional activity of NtNAC053, the full length of *NtNAC053* and two truncated variants, including N-terminal region and C-terminal region, were inserted into the pBridge vector, respectively (**Figure 3A**). The pBridge empty vector and these three recombinant constructs were then transformed into yeast competent cells. The transformants containing the full length of NtNAC053 and C-terminal region of NtNAC053 were found to show β-galactosidase activity and turn blue on the SD/-Trp/X-gal medium. Whereas the transformants with pBridge empty vector and N-terminal region of NtNAC053 did not show β-galactosidase activity (**Figure 3B**). Therefore, NtNAC053 could function as a transcriptional activator, and the C-terminal region is critical for this activity.

The Overexpression of *NtNAC053* in Tobacco Confers Abscisic Acid Hypersensitivity

To clarify the potential function of NtNAC053, the transgenic overexpression lines of *NtNAC053* gene were generated in tobacco. The qRT-PCR analysis was used to detect the expression levels of *NtNAC053* gene in these lines. The results revealed the expression levels of this gene in *OE6* and *OE8* reached 47.7- and 61.8-fold than that in wild type (WT) (**Supplementary Figure 5**). These two lines were then selected to perform phenotyping under different treatments. In our previous survey,

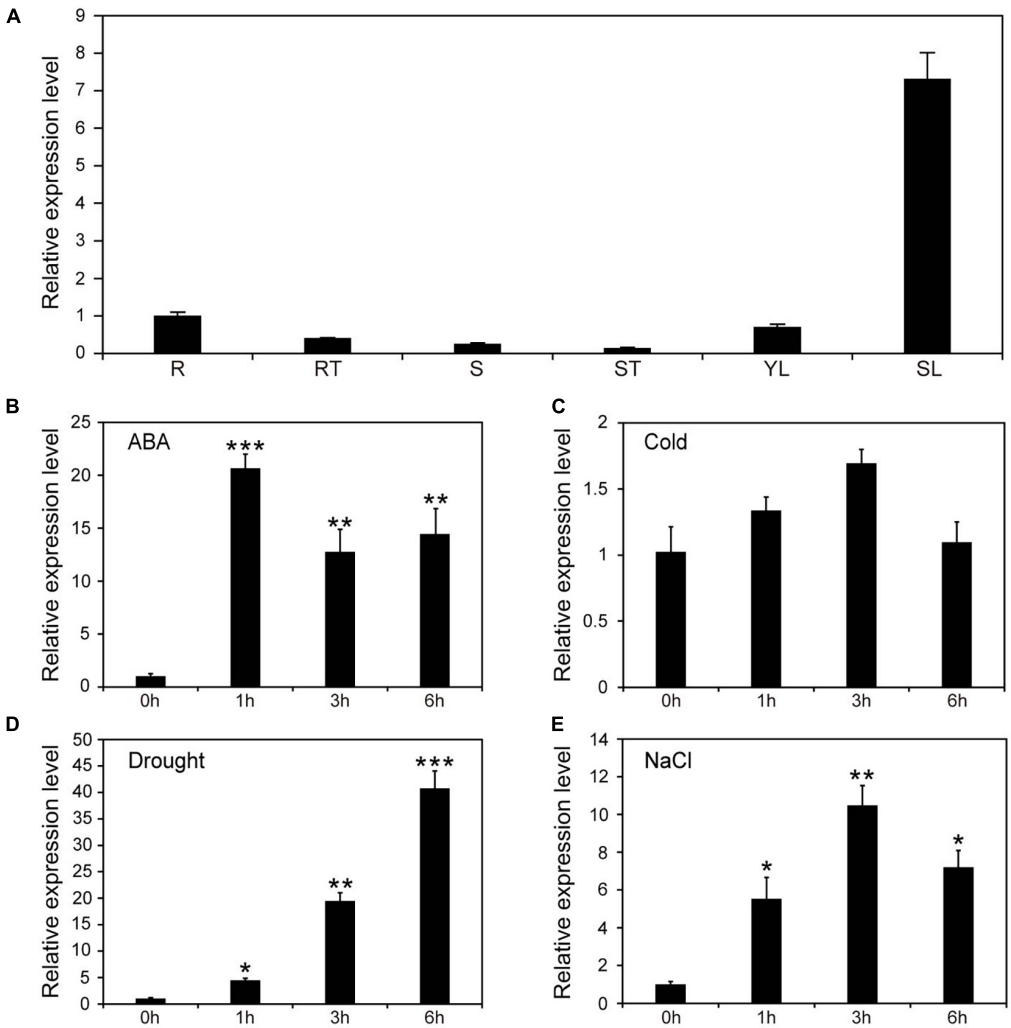


FIGURE 1 | Expression patterns of *NtNAC053* in tobacco. **(A)** The expression pattern of *NtNAC053* in tested tissues. R: root; RT: root tip; S: stem; ST: stem tip; YL: young leaf; SL: senescent leaf; Expression pattern of *NtNAC053* under ABA **(B)**, cold **(C)**, drought **(D)**, and salt **(E)** treatments. The data were means ± SD from three independent replications. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (*t*-tests).

the *NtNAC053* gene could be induced by exogenous ABA. Hence, the germination assay and root elongation assay were carried out to assess the ABA sensitivity in the overexpressed lines. As a result, the germination rates seemed to be identical in the germination assay between the overexpressed lines and WT plants. Interestingly, the overexpressed lines had significantly lower germination rates than WT plants when exposed under ABA treatment (**Figures 4A,B**). Furthermore, the root length seemed to be equivalent in the root elongation assay between the overexpressed lines and WT plants. Notably, the overexpressed lines seemed to be more sensitive than WT plants when treated by exogenous ABA (**Figures 4C,D**).

NtNAC053 Overexpression Enhances Salt Tolerance in Transgenic Tobacco

To explore whether NtNAC053 regulates the tolerance of tobacco plants to salt stress, we performed a seed germination assay and root elongation assay on MS media with salt treatment. As shown in **Figure 5**, the seed germination rates of both the overexpression lines and WT were decreased under salt treatment, whereas the overexpression lines displayed higher germination rates than WT seeds under salt treatments.

As shown in **Figure 6**, no significant difference was observed in root length between overexpression lines and WT seedlings. However, the primary roots of overexpression seedlings were longer than that of WT on the MS media supplemented with tested concentrations of NaCl.

NtNAC053 Overexpression Improves Drought Tolerance in Transgenic Tobacco

To further explore the function of *NtNAC053* gene, the overexpression lines and WT seedlings were treated by the

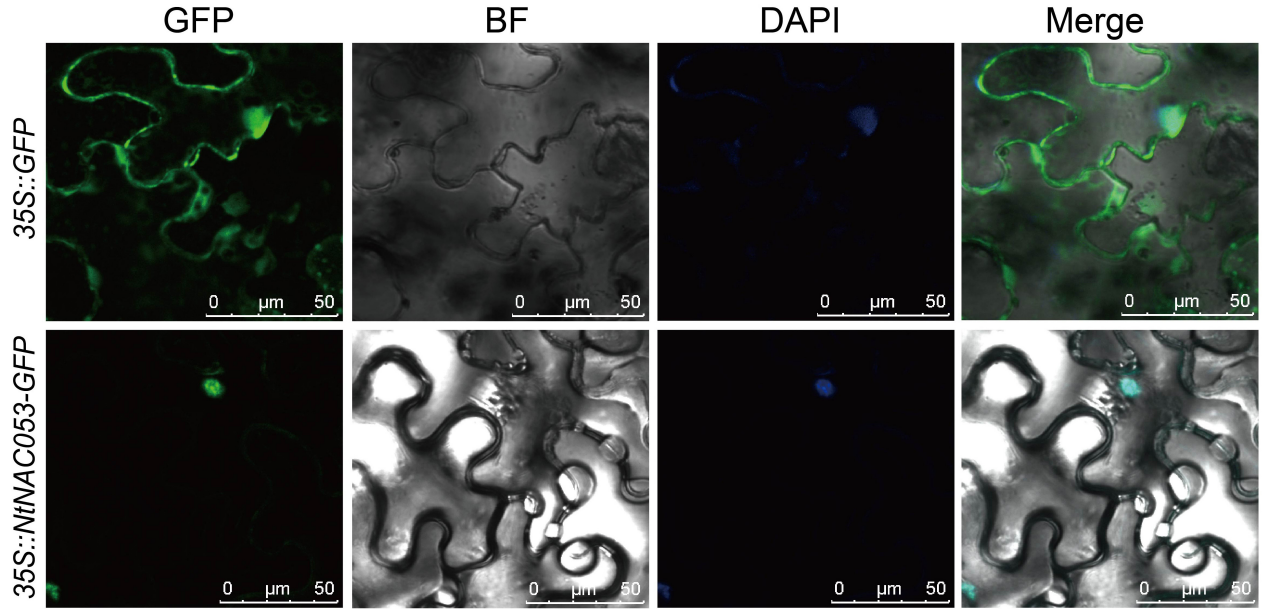


FIGURE 2 | The subcellular localization of the NtNAC053 protein in *Nicotiana benthamiana* epidermal cells. Transient expression of GFP and NtNAC053-GFP in *N. benthamiana* epidermal cells. Bar = 50 μM, DAPI, 4,6-diamidino-2-phenylindole for nuclear staining.

drought stress with no watering for approximately 10 days. The growth of overexpression lines and WT seedlings was almost the same under normal growth conditions. After the drought treatment, overexpression lines and WT seedlings displayed leaf wilting phenotype, whereas WT seedlings were much more extreme than that in overexpression lines (**Figure 7A**). After 3 days of re-watering, the survival rates of the *OE6* lines (72%) and *OE8* lines (85%) were significantly higher than those of the WT seedlings (35%) (**Figure 7B**). Additionally, the overexpression lines displayed higher water content than WT seedlings during the 3 h dehydration (**Figure 7C**). Besides, the drought stress-responsive gene *NtAQP1* was selected to explore the expression in wild type and overexpression lines. Under drought conditions, the expression of *NtAQP1* was significantly higher in overexpression lines (**Supplementary Figure 6**). These findings suggested that the overexpression of *NtNAC053* gene might enhance the drought tolerance of the tobacco plants.

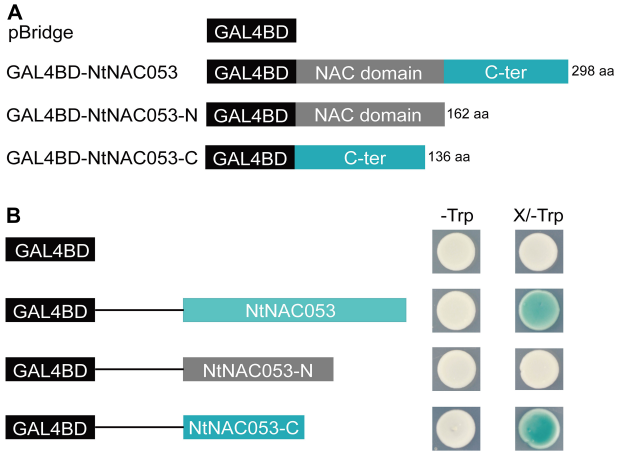


FIGURE 3 | The analysis of transcriptional activation activity of NtNAC053. **(A)** The fusion proteins with the GAL4 DNA-binding domain. **(B)** Transcriptional activation activity assay of NtNAC053 in the yeast strain AH109. The β-galactosidase activity was detected on the SD/-Trp/X-gal medium. NtNAC053-N represented the N-terminal region of NtNAC053 (1-162 aa); NtNAC053-C represented the C-terminal region of NtNAC053 (163-298 aa).

NtNAC053 Overexpression Enhances ROS-Scavenging Capability in Transgenic Tobacco

To further investigate the function of *NtNAC053*, the stress-related physiological parameters, including catalase (CAT), peroxidase (POD), superoxide (SOD), malonic dialdehyde (MDA), and hydrogen peroxide (H_2O_2), were detected (**Figure 8** and **Supplementary Figure 7**). After salt and drought treatments, the content of MDA and H_2O_2 in overexpression lines was significantly lower than those in WT plants. However, the activities of CAT, POD, and SOD in overexpression lines were significantly higher than those in WT plants.

NtNAC053 Regulates the Expression of Stress-Responsive Genes

To further explore the function of *NtNAC053* gene, the expression patterns of stress-responsive genes were detected in overexpression lines and WT plants using qRT-PCR. The expression pattern of the selected stress-responsive genes was calculated as folds relative to the expression level of

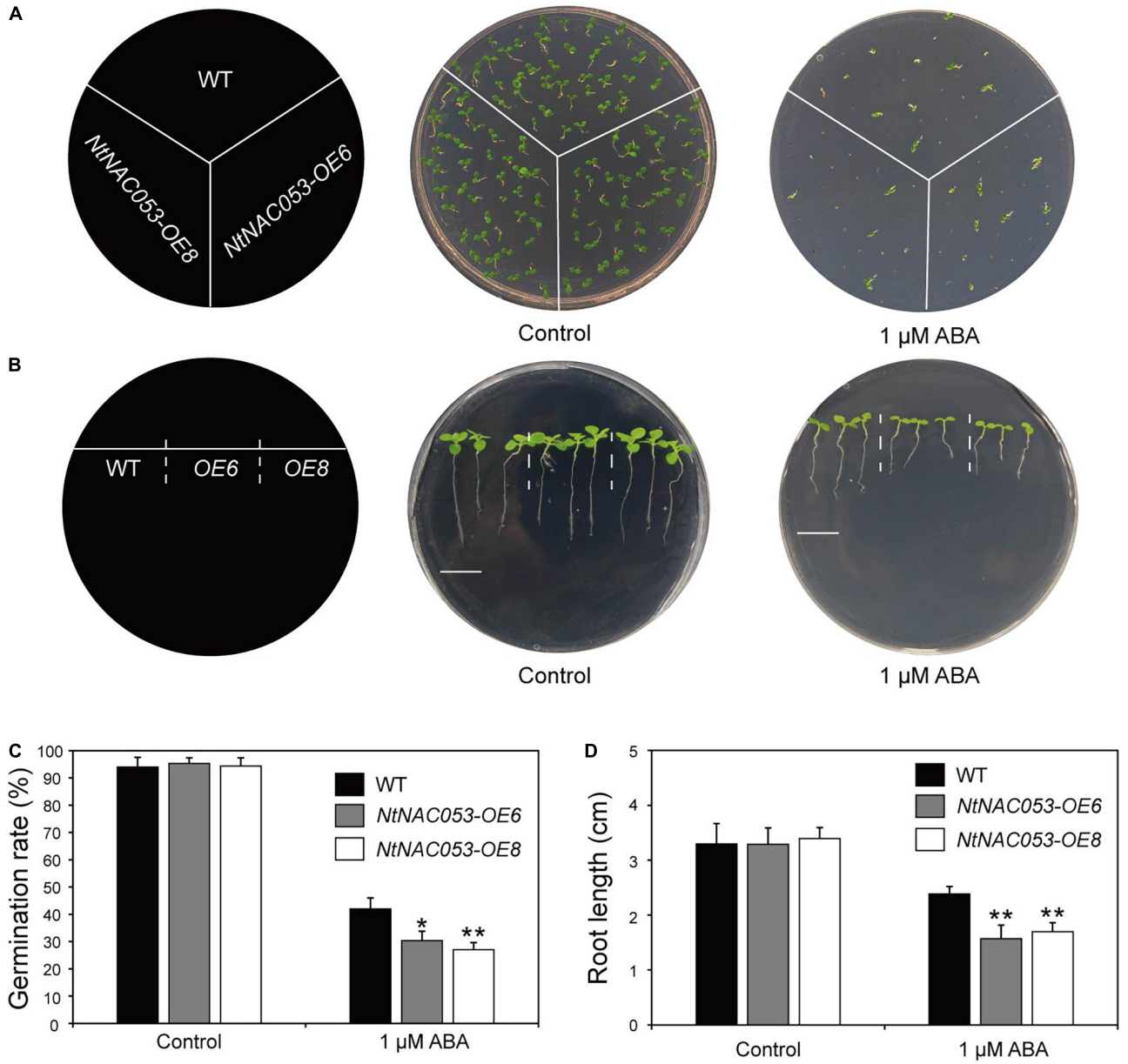


FIGURE 4 | Overexpression of the *NtNAC053* gene confers ABA hypersensitivity. **(A)** Seed germination assay between overexpression lines and wild type under 1 μM ABA treatment. **(B)** Root elongation assay between overexpression lines and wild type under 1 μM ABA treatment. Bar = 1.5 cm. **(C)** The statistics of germination rates under normal conditions and 1 μM ABA treatment. **(D)** The statistics of primary root length under normal conditions and 1 μM ABA treatment. WT, wild type. The data were means ± SD from three independent replications. * $p < 0.05$, ** $p < 0.01$ (*t*-tests).

none treatments. Exposed to salt and drought stresses, the expression levels of several genes were significantly induced in overexpression lines compared to WT plants, including *NtCOR15A*, *NtRAB18*, *NtDREB1A*, *NtERF5*, *NtKAT2*, and *NtERD11* (**Figure 9**). Besides, we selected two salt stress-responsive genes (*NtNHX1* and *NtSOS1*) and compared their expressions in overexpression lines and WT plants. Under salt conditions, the expressions of *NtNHX1* and *NtSOS1* were significantly higher in overexpression lines (**Supplementary Figure 8**). Notably, ABA signaling-related genes (*NtPYL6* and *NtSnRK2.4*) were selected for further study. In overexpression lines, these gene expressions were much higher compared to WT plants under drought and salt conditions (**Supplementary Figure 9**).

NtNAC053 Binds to NACRS Sequence

The NAC transcription factors mainly bind to the NAC recognition sequence (NACRS) with a core sequence of CACG to regulate gene expression, while the reverse complementary sequence is CGT[A/G]. To verify the putative DNA-binding

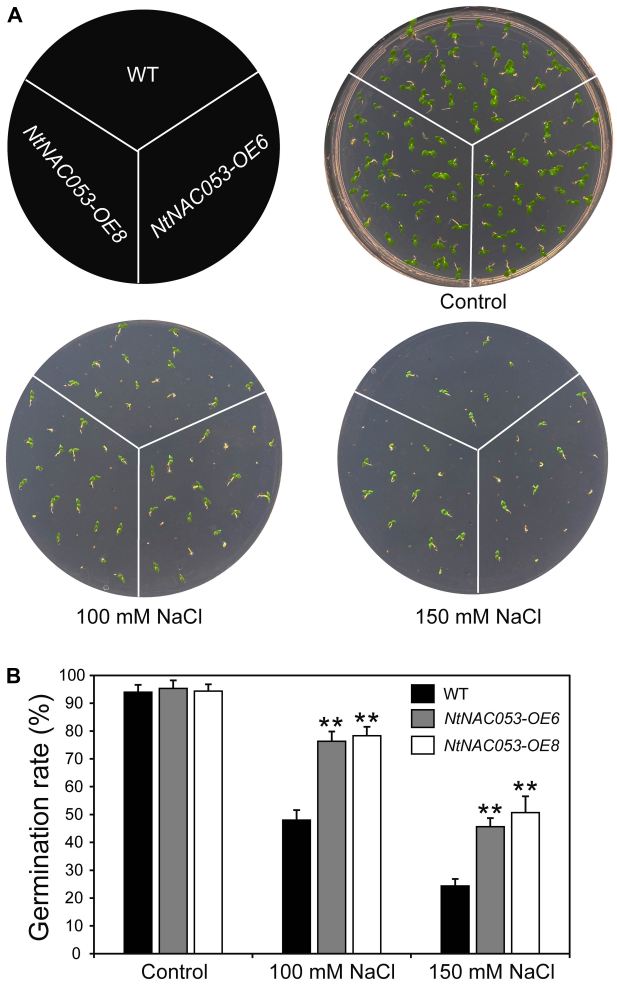


FIGURE 5 | Overexpression of the *NtNAC053* gene promotes seed germination under salt stress treatments. **(A)** Seed germination assay between overexpression lines and wild type under NaCl treatments. **(B)** The statistics of seed germination under normal conditions and NaCl treatments. WT, wild type. The data were means ± SD from three independent replications. ** $p < 0.01$ (*t*-tests).

sequences of *NtNAC053*, the NACRS sequence and the reverse complementary sequence were used as reporters (R1 and R2), while the base-substituted fragment was used as the mutant reporter (RM). The activator vector *pGAD-NtNAC053* was co-transformed into the yeast strain with three reporter vectors, respectively. If the NtNAC053 can bind to the tandem sequence of the reporter construct, the GAD will be able to direct *LacZ* expression, resulting in the blue color accumulation. As shown in **Figure 10**, these yeast one-hybrid experiments confirmed the ability of NtNAC053 to bind to the NACRS sequence *in vivo*.

DISCUSSION

Plants are suffering from a combination of abiotic and biotic stresses. The abiotic stresses, including salt and drought, affect plant growth and development. In previous studies, several NAC members have been revealed to participate in responses to abiotic stress through ABA signaling (Nakashima et al., 2012; Shao et al., 2015). In the current study, promoter analysis showed that the *NtNAC053* promoter possesses five ABRE motifs, which is an important *cis*-acting element related to ABA signal transduction, indicating this gene might be involved in the ABA signaling pathway (**Supplementary Figure 4**). To clarify this inference, the expression pattern analysis and overexpression survey were carried out. Interestingly, it was shown that *NtNAC053* was significantly induced by exogenous ABA treatments (**Figure 1B**). Furthermore, the seed germination rate and root lengths of overexpression lines were significantly lower than those of WT plants, suggesting that overexpression of *NtNAC053* conferred transgenic tobacco plants hypersensitive to ABA (**Figure 4**). Notably, the expressions of ABA signaling genes, including *NtPYL6* and *NtSnRK2.4*, were much higher in overexpression lines compared to WT plants under salt and drought conditions (**Supplementary Figure 9**).

It has been reported that NAC transcription factors play critical roles in various abiotic stress responses. However, a few studies were reported about the biofunctions of the NAC transcription factor in tobacco. In our previous work, a tobacco *NAC* gene was found to be dramatically induced by multiple stress treatments (Li et al., 2018), which was identified as the homolog of StNAC053. In our previous study, we found that the overexpression of *StNAC053* gene could enhance the salt and drought tolerance of transgenic *Arabidopsis* (Wang et al., 2021). Considering the close genetic distance of potato and tobacco, the homolog of StNAC053 might confer salt and drought tolerance in tobacco. To verify this hypothesis, a systematic analysis of NtNAC053 was performed. The StNAC053 had been reported to fall into the ATAF subfamily. Hence, several reported ATAF family members from potato, wheat, and Arabidopsis were selected to perform the phylogenetic analysis. Phylogenetic analysis showed that NtNAC053 was clustered with StNAC053 in ATAF subfamily, implying that NtNAC053 served as a homolog of StNAC053. In tobacco, we found that *NtNAC053* was significantly up-regulated when subjected to salt and drought stresses, suggesting that NtNAC053 may function as a regulator in response to salt and drought stresses (**Figure 1**). Furthermore, the seed germination rate and root lengths of WT plants were significantly lower than those of overexpression lines under salt treatment (**Figures 5**, **6**). Meanwhile, overexpression lines showed higher survival rates compared to WT plants when exposed to drought stress (**Figures 7A,B**). In addition, the leaves of overexpression lines had a lower water loss rate than those of WT plants under drought conditions (**Figure 7C**). Considering these results, the overexpression of *NtNAC053* gene could enhance the tolerance to salt and drought stresses in tobacco.

Additionally, massive amounts of ROS were produced when plants were exposed to multiple stresses. Generally, excessive ROS could destroy antioxidant systems and the integrity of cell membrane, resulting in the accumulation of MDA (Moore and Roberts, 1998; Gill and Tuteja, 2010). Plants have developed intricate defense mechanisms to cope with excess ROS, which include an antioxidant enzyme system.

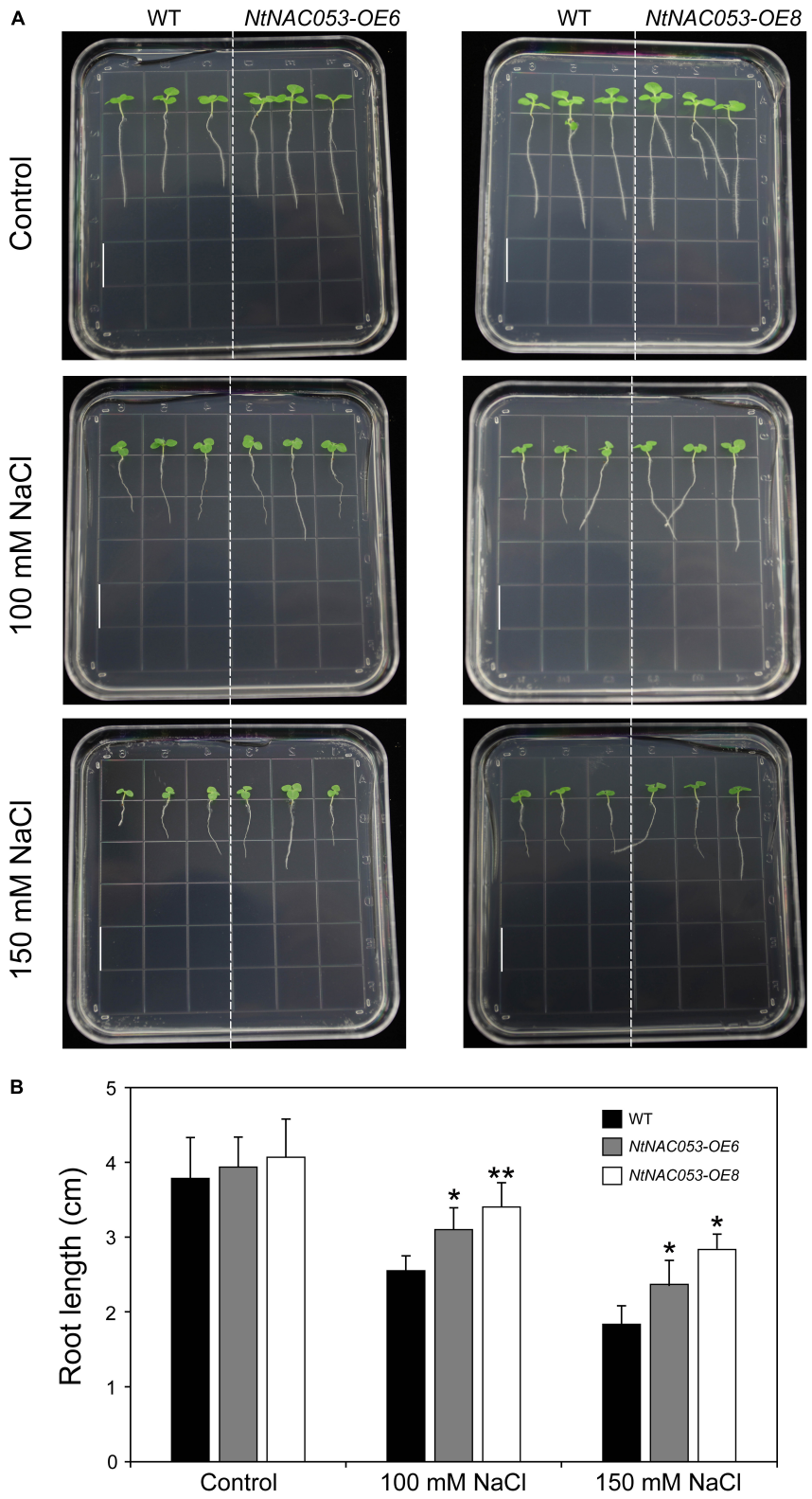


FIGURE 6 | Comparison of root growth between overexpression lines and wild-type tobacco under salt stress treatments. **(A)** The primary root length of overexpression lines and wild-type tobacco under 100 mM and 150 mM NaCl treatments. The primary root length was recorded 7 days after growth. Bar = 1.5 cm. **(B)** The statistics of primary root length under normal conditions and NaCl treatments. WT, wild type. The data were means ± SD from three independent replications. * $p < 0.05$, ** $p < 0.01$ (*t*-tests).

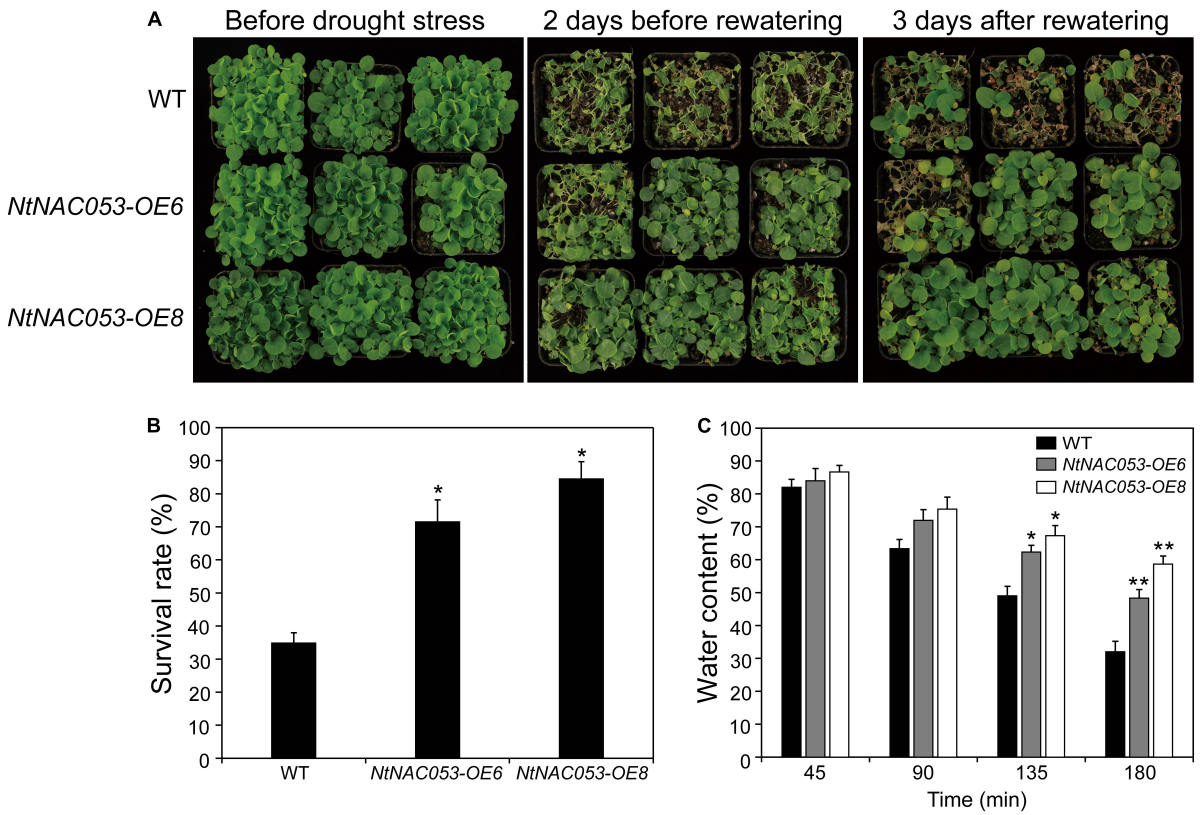


FIGURE 7 | Drought tolerance assay of transgenic *NtNAC053* tobacco. **(A)** Phenotypes of overexpression lines and wild-type tobacco under drought stress. Under normal conditions, the overexpression lines and wild-type tobacco were normally grown for 6 weeks. Overexpression lines displayed higher drought tolerance after 10 days of drought stress and 3 days of recovery. **(B)** The survival rate of the overexpression lines and wild-type tobacco were determined. **(C)** The water content of detached leaves of overexpression lines and wild-type tobacco. WT, wild type. The data were means ± SD from three independent replications. * $p < 0.05$, ** $p < 0.01$ (*t*-tests).

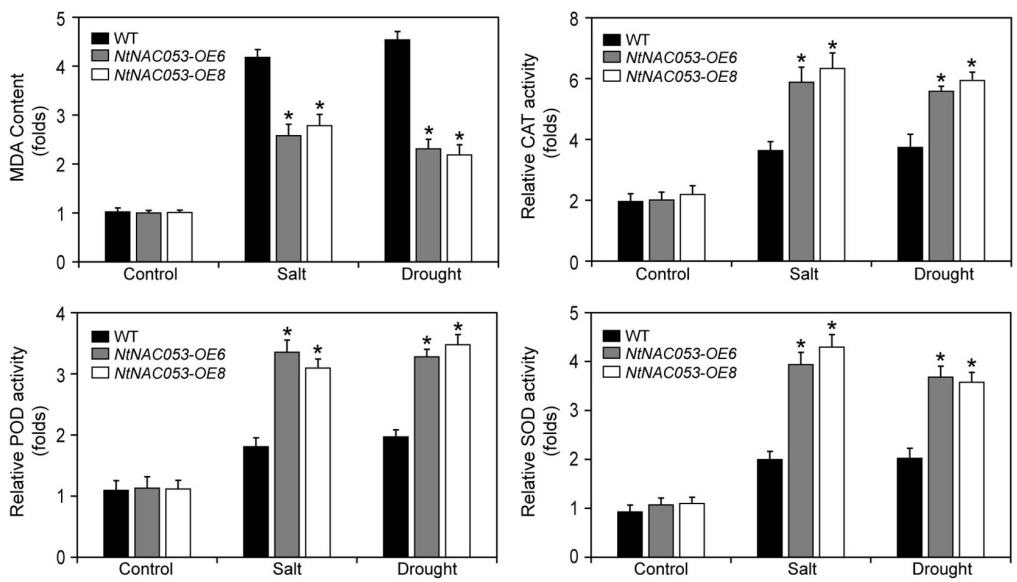


FIGURE 8 | Analysis of the antioxidant enzyme activities and MDA contents in overexpression lines and wild-type tobacco under salt and drought stresses. MDA, malonic dialdehyde; CAT, catalase; POD, peroxidase; SOD; superoxide. WT, wild type. The MDA content and enzyme activity were calculated as folds relative to the level of before treatments. The data were means ± SD from three independent replications. * $p < 0.05$ (*t*-tests).

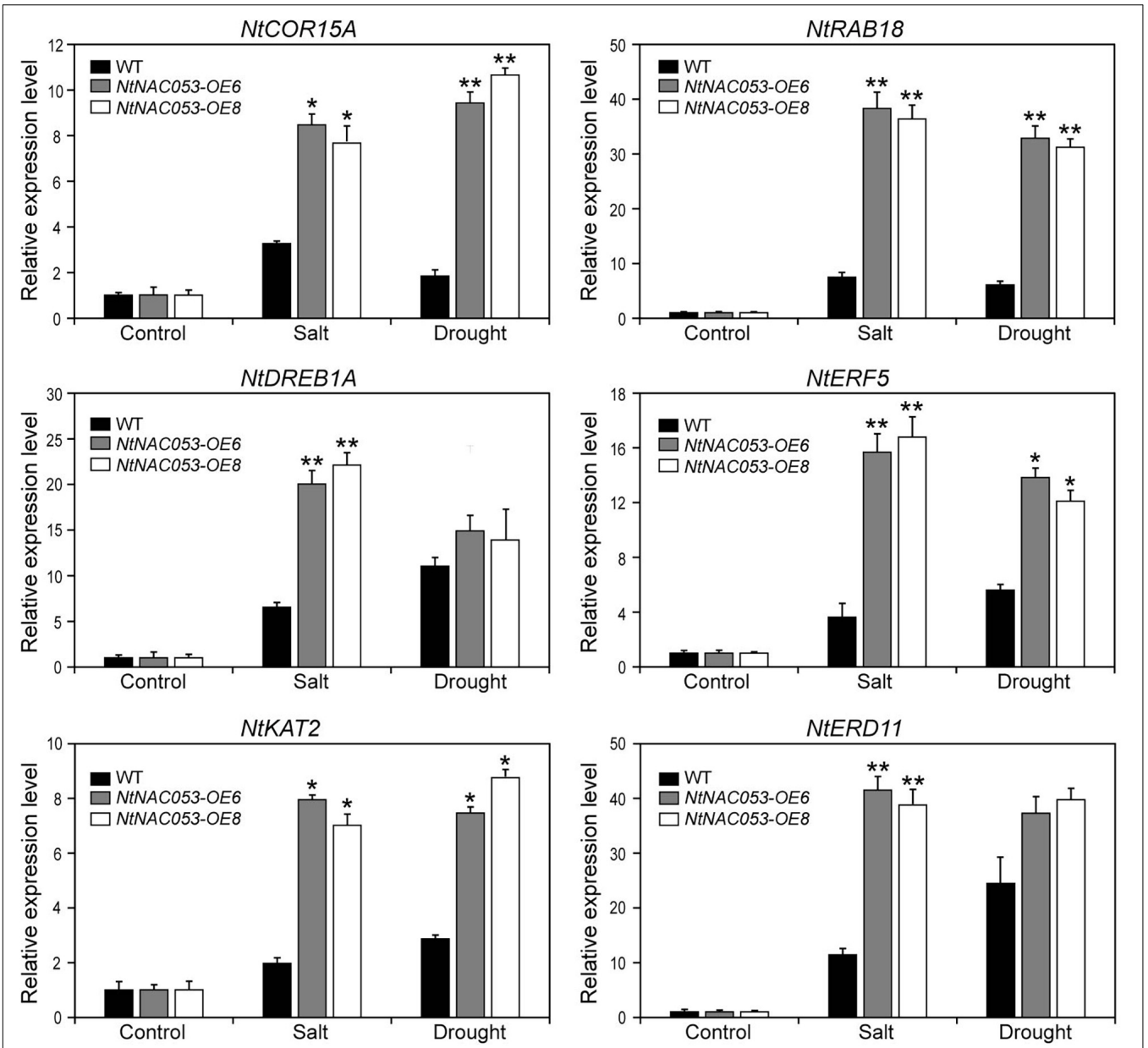


FIGURE 9 | Expression of stress-related genes in overexpression lines and wild-type tobacco under salt and drought stresses. The ratios of gene expression levels were calculated relative to the control, respectively. WT, wild type. The data were means ± SD from three independent replications. * $p < 0.05$, ** $p < 0.01$ (*t*-tests).

Notably, several enzymes, including CAT, POD, and SOD, played significant roles in the antioxidant enzyme system (Gill and Tuteja, 2010). Relative to wild type, the contents of MDA and hydrogen peroxide in overexpression lines were significantly decreased when exposed to salt and drought stresses, suggesting that there is lower oxidative damage in overexpression lines (**Figure 8** and **Supplementary Figure 7**). Furthermore, the enzyme activities of CAT, POD, and SOD in overexpression lines were significantly higher than those of wild type under salt and drought stresses, suggesting that overexpression lines increased the capacity of the ROS-scavenging system (**Figure 8**). It was reported that NAC transcription factors usually participate in response to abiotic stresses through regulating the expression of downstream stress-responsive genes (He et al., 2019b). AtNAP functions to negatively regulate salt stress tolerance through repressing *ARBE1* (Seok et al., 2017). In this study, we found that *NtNAC053* encoded a nuclear localization protein and possessed transcriptional activation ability (**Figures 2**, **3**). Many aquaporins genes were reported to be responsive to abiotic stress during different developing stages. In tobacco, overexpression of the *NtAQP1* gene improves water use efficiency, hydraulic conductivity, and yield production under abiotic stress

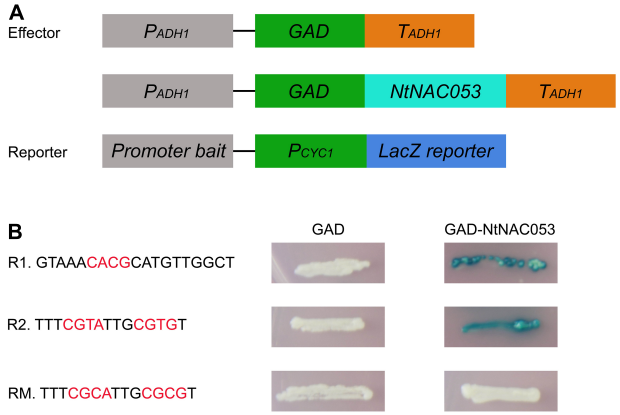


FIGURE 10 | Binding activity of NtNAC053. **(A)** Diagrams illustrating the structure of reporter and effector constructs. **(B)** The sequence of the NACRS (R1 and R2) and mutated NACRS (RM). Different combinations of effector and reporter constructs were co-expressed in yeast strain YM4271.

(Sade et al., 2010). Notably, under drought stress treatments, the drought stress-responsive gene *NtAQP1* was significantly higher in the *NtNAC053* overexpression lines (**Supplementary Figure 6**). Furthermore, NtNAC053 could significantly activate several stress-responsive genes, including *NtCOR15A, NtRAB18, NtDREB1A, NtERF5, NtKAT2*, and *NtERD11* (**Figure 9**). Besides, several ion channels encoding genes, including *NHX1* and *SOS1*, have been reported to be involved in the ion homeostasis and salt stress responses. These genes were significantly higher in overexpression lines (**Supplementary Figure 8**). In addition, promoter analysis revealed that all these stress-responsive genes promoters possessed NAC-binding sites (**Supplementary Table 2**). Yeast one-hybrid assay implied that NtNAC053 might be combined with NACRS (**Figure 10**), suggesting NtNAC053 could activate these stress-responsive genes by binding their promoter directly.

CONCLUSION

In this study, a typical *NAC* gene *NtNAC053* was cloned from tobacco K326 and functionally verified in overexpressing tobacco. *NtNAC053* overexpression lines showed significantly increased resistance to salt and drought stresses, while the overexpression lines seemed to be more sensitive to ABA treatment. Taken together, the transcription factor NtNAC053 might be involved in salt and drought stress responses *via* ABA-mediated signaling in tobacco.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

XL, QW, and CG conducted the research and participated in drafting the manuscript. JS, ZL, YW, AY, WP, and YG assisted in data collection and analysis. JG and LW conceived this research, designed the experiments, and drafted the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Tobacco Genome Project [110202001029 (JY-12), and 110202101003 (JY-03)], the China Tobacco Hunan Industrial Co., Ltd., Science and Technology Project (KY2022YC0010 and KY2021YC0005), and the Key Science and Technology Program of Hunan Provincial Tobacco Corporation (Grant No. 19-23Aa01).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.817106/full#supplementary-material

Supplementary Figure 1 | Multiple sequence alignment of NtNAC053 with reported NAC proteins from *Arabidopsis* and potato, including ANAC002, ANAC081, ANAC029, ANAC072, ANAC055, ANAC019, and StNAC053. The black lines indicated five subdomains **(A–E)** and nuclear location signal (NLS).

Supplementary Figure 2 | The phylogenetic, syntenic analysis and motif organization of NtNAC053. **(A)** The phylogenetic tree was generated based on the amino acid sequences of NtNAC053 and other reported NAC proteins. **(B)** The motif organizations were predicted by MEME. **(C)** The gray line in the background represents the collinear blocks between tobacco and potato, while the red line exhibits the syntenic NAC gene pairs.

Supplementary Figure 3 | The putative conserved motifs in NAC proteins.

Supplementary Figure 4 | The *cis*-acting element analysis of the promoter regions of *NtNAC053* gene.

Supplementary Figure 5 | The expression level of the *NtNAC053* gene in wild type and two overexpression lines, the ratios of gene expression levels were calculated relative to the wild type.

Supplementary Figure 6 | Expression of *NtAQP1* in overexpression lines and wild-type tobacco under drought stress. The ratios of gene expression levels were calculated relative to the control, respectively.

Supplementary Figure 7 | Analysis of H_2O_2 contents in overexpression lines and wild-type tobacco under salt and drought stresses. H_2O_2, hydrogen peroxide. WT, wild type.

Supplementary Figure 8 | Expression of *NtNHX1* and *NtSOS1* in overexpression lines and wild-type tobacco under salt stress. The ratios of gene expression levels were calculated relative to the control, respectively. WT, wild type.

Supplementary Figure 9 | The expressions of ABA signaling genes in overexpression lines and wild-type tobacco under salt and drought stresses. The ratios of gene expression levels were calculated relative to the control, respectively. WT, wild type.

REFERENCES

Aida, M., Ishida, T., Fukaki, H., Fujisawa, H., and Tasaka, M. (1997). Genes involved in organ separation in *Arabidopsis*: an analysis of the cup-shaped cotyledon mutant. *Plant Cell* 9, 841–857. doi: 10.1105/tpc.9.6.841

Buschmann, H. (2016). Plant cell division analyzed by transient Agrobacterium-mediated transformation of tobacco bY-2 cells. *Methods Mol. Biol.* 1370, 17–25. doi: 10.1007/978-1-4939-3142-2_2

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant.* 13, 1194–1202. doi: 10.1016/j.molp.2020.06.009

Ernst, H. A., Olsen, A. N., Larsen, S., and Lo Leggio, L. (2004). Structure of the conserved domain of ANAC, a member of the NAC family of transcription factors. *EMBO Rep.* 5, 297–303. doi: 10.1038/sj.embor.7400093

Fujita, M., Fujita, Y., Maruyama, K., Seki, M., Hiratsu, K., Ohme-Takagi, M., et al. (2004). A dehydration-induced NAC protein, RD26, is involved in a novel ABA-dependent stress-signaling pathway. *Plant J.* 39, 863–876. doi: 10.1111/j.1365-313X.2004.02171.x

Gill, S. S., and Tuteja, N. (2010). Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. *Plant Physiol. Biochem.* 48, 909–930. doi: 10.1016/j.plaphy.2010.08.016

Gong, X., Zhao, L., Song, X., Lin, Z., Gu, B., Yan, J., et al. (2019). Genome-wide analyses and expression patterns under abiotic stress of NAC transcription factors in white pear (*Pyrus bretschneideri*). *BMC Plant Biol.* 19:161. doi: 10.1186/s12870-019-1760-8

Guo, Y., and Gan, S. (2006). AtNAP, a NAC family transcription factor, has an important role in leaf senescence. *Plant J.* 46, 601–612. doi: 10.1111/j.1365-313X.2006.02723.x

He, K., Zhao, X., Chi, X., Wang, Y., Jia, C., Zhang, H., et al. (2019a). A novel Miscanthus NAC transcription factor MlNAC10 enhances drought and salinity tolerance in transgenic *Arabidopsis*. *J. Plant Physiol.* 233, 84–93. doi: 10.1016/j.jplph.2019.01.001

He, L., Bian, J., Xu, J., and Yang, K. (2019b). Novel maize NAC transcriptional repressor ZmNAC071 confers enhanced sensitivity to ABA and osmotic stress by downregulating stress-responsive genes in transgenic *Arabidopsis*. *J. Agric. Food Chem.* 67, 8905–8918. doi: 10.1021/acs.jafc.9b02331

He, X. J., Mu, R. L., Cao, W. H., Zhang, Z. G., Zhang, J. S., and Chen, S. Y. (2005). AtNAC2, a transcription factor downstream of ethylene and auxin signaling pathways, is involved in salt stress response and lateral root development. *Plant J.* 44, 903–916. doi: 10.1111/j.1365-313X.2005.02575.x

Hendelman, A., Stav, R., Zemach, H., and Arazi, T. (2013). The tomato NAC transcription factor SlNAM2 is involved in flower-boundary morphogenesis. *J. Exp. Bot.* 64, 5497–5507. doi: 10.1093/jxb/ert324

Hibara, K., Takada, S., and Tasaka, M. (2003). CUC1 gene activates the expression of SAM-related genes to induce adventitious shoot formation. *Plant J.* 36, 687–696. doi: 10.1046/j.1365-313x.2003.01911.x

Kim, M. J., Park, M. J., Seo, P. J., Song, J. S., Kim, H. J., and Park, C. M. (2012). Controlled nuclear import of the transcription factor NTL6 reveals a cytoplasmic role of SnRK2.8 in the drought-stress response. *Biochem. J.* 448, 353–363. doi: 10.1042/BJ20120244

Kim, S. G., Lee, S., Seo, P. J., Kim, S. K., Kim, J. K., and Park, C. M. (2010). Genome-scale screening and molecular characterization of membrane-bound transcription factors in *Arabidopsis* and rice. *Genomics* 95, 56–65. doi: 10.1016/j.ygeno.2009.09.003

Larsson, E., Sundstrom, J. F., Sitbon, F., and von Arnold, S. (2012). Expression of PaNAC01, a picea abies CUP-SHAPED COTYLEDON orthologue, is regulated by polar auxin transport and associated with differentiation of the shoot apical meristem and formation of separated cotyledons. *Ann. Bot.* 110, 923–934. doi: 10.1093/aob/mcs151

Li, P., Peng, Z., Xu, P., Tang, G., Ma, C., Zhu, J., et al. (2021). Genome-wide identification of NAC transcription factors and their functional prediction of abiotic stress response in peanut. *Front. Genet.* 12:630292. doi: 10.3389/fgene.2021.630292

Li, W., Li, X., Chao, J., Zhang, Z., Wang, W., and Guo, Y. (2018). NAC family transcription factors in tobacco and their potential role in regulating leaf senescence. *Front. Plant Sci.* 9:1900. doi: 10.3389/fpls.2018.01900

Li, Z., Chao, J., Li, X., Li, G., Song, D., Guo, Y., et al. (2021). Systematic analysis of the bZIP family in tobacco and functional characterization of NtbZIP62 involvement in salt stress. *Agronomy* 11:148. doi: 10.3390/agronomy11010148

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Mitsuda, N., Iwase, A., Yamamoto, H., Yoshida, M., Seki, M., Shinozaki, K., et al. (2007). NAC transcription factors, NST1 and NST3, are key regulators of the formation of secondary walls in woody tissues of *Arabidopsis*. *Plant Cell* 19, 270–280. doi: 10.1105/tpc.106.047043

Moore, K., and Roberts, L. J. (1998). Measurement of lipid peroxidation. *Free Radic. Res.* 28, 659–671. doi: 10.3109/10715769809065821

Nakashima, K., Takasaki, H., Mizoi, J., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2012). NAC transcription factors in plant abiotic stress responses. *Biochim. Biophys. Acta* 1819, 97–103. doi: 10.1016/j.bbagrm.2011.10.005

Nuruzzaman, M., Manimekalai, R., Sharoni, A. M., Satoh, K., Kondoh, H., Ooka, H., et al. (2010). Genome-wide analysis of NAC transcription factor family in rice. *Gene* 465, 30–44. doi: 10.1016/j.gene.2010.06.008

Olsen, A. N., Ernst, H. A., Leggio, L. L., and Skriver, K. (2005). NAC transcription factors: structurally distinct, functionally diverse. *Trends Plant Sci.* 10, 79–87. doi: 10.1016/j.tplants.2004.12.010

Ooka, H., Satoh, K., Doi, K., Nagata, T., Otomo, Y., Murakami, K., et al. (2003). Comprehensive analysis of NAC family genes in *Oryza sativa* and *Arabidopsis thaliana*. *DNA Res.* 10, 239–247. doi: 10.1093/dnares/10.6.239

Park, J., Kim, Y. S., Kim, S. G., Jung, J. H., Woo, J. C., and Park, C. M. (2011). Integration of auxin and salt signals by the NAC transcription factor NTM2 during seed germination in *Arabidopsis*. *Plant Physiol.* 156, 537–549. doi: 10.1104/pp.111.177071

Puranik, S., Bahadur, R. P., Srivastava, P. S., and Prasad, M. (2011). Molecular cloning and characterization of a membrane associated NAC family gene, SiNAC from foxtail millet [Setaria italica (L.) P. Beauv]. *Mol. Biotechnol.* 49, 138–150. doi: 10.1007/s12033-011-9385-7

Qu, L. J., and Zhu, Y. X. (2006). Transcription factor families in *Arabidopsis*: major progress and outstanding issues for future research. *Curr. Opin. Plant Biol.* 9, 544–549. doi: 10.1016/j.pbi.2006.07.005

Sade, N., Gebretsadik, M., Seligmann, R., Schwartz, A., Wallach, R., and Moshelion, M. (2010). The role of tobacco Aquaporin1 in improving water use efficiency, hydraulic conductivity, and yield production under salt stress. *Plant Physiol.* 152, 245–254. doi: 10.1104/pp.109.145854

Seok, H. Y., Woo, D. H., Nguyen, L. V., Tran, H. T., Tarte, V. N., Mehdi, S. M., et al. (2017). Arabidopsis AtNAP functions as a negative regulator via repression of AREB1 in salt stress response. *Planta* 245, 329–341. doi: 10.1007/s00425-016-2609-0

Shao, H., Wang, H., and Tang, X. (2015). NAC transcription factors in plant multiple abiotic stress responses: progress and prospects. *Front. Plant Sci.* 6:902. doi: 10.3389/fpls.2015.00902

Shiriga, K., Sharma, R., Kumar, K., Yadav, S. K., Hossain, F., and Thirunavukkarasu, N. (2014). Genome-wide identification and expression pattern of drought-responsive members of the NAC family in maize. *Meta. Gene* 2, 407–417. doi: 10.1016/j.mgene.2014.05.001

Singh, A. K., Sharma, V., Pal, A. K., Acharya, V., and Ahuja, P. S. (2013). Genome-wide organization and expression profiling of the NAC transcription factor family in potato (*Solanum tuberosum* L.). *DNA Res.* 20, 403–423. doi: 10.1093/dnares/dst019

Souer, E., van Houwelingen, A., Kloos, D., Mol, J., and Koes, R. (1996). The no apical meristem gene of petunia is required for pattern formation in embryos and flowers and is expressed at meristem and primordia boundaries. *Cell* 85, 159–170. doi: 10.1016/s0092-8674(00)81093-4

Takada, S., Hibara, K., Ishida, T., and Tasaka, M. (2001). The CUP-SHAPED COTYLEDON1 gene of *Arabidopsis* regulates shoot apical meristem formation. *Development* 128, 1127–1135. doi: 10.1007/s004290100164

Takasaki, H., Maruyama, K., Kidokoro, S., Ito, Y., Fujita, Y., Shinozaki, K., et al. (2010). The abiotic stress-responsive NAC-type transcription factor OsNAC5 regulates stress-inducible genes and stress tolerance in rice. *Mol. Genet. Genomics* 284, 173–183. doi: 10.1007/s00438-010-0557-0

Tang, Y., Liu, M., Gao, S., Zhang, Z., Zhao, X., Zhao, C., et al. (2012). Molecular characterization of novel TaNAC genes in wheat and overexpression of TaNAC2a confers drought tolerance in tobacco. *Physiol. Plant* 144, 210–224. doi: 10.1111/j.1399-3054.2011.01539.x

Tran, L. S., Nakashima, K., Sakuma, Y., Simpson, S. D., Fujita, Y., Maruyama, K., et al. (2004). Isolation and functional analysis of *Arabidopsis* stress-inducible NAC transcription factors that bind to a drought-responsive cis-element in the early responsive to dehydration stress 1 promoter. *Plant Cell* 16, 2481–2498. doi: 10.1105/tpc.104.022699

Wang, Q., Guo, C., Li, Z., Sun, J., Deng, Z., Wen, L., et al. (2021). Potato NAC transcription factor StNAC053 enhances salt and drought tolerance in transgenic *Arabidopsis*. *Int. J. Mol. Sci.* 22:2568. doi: 10.3390/ijms22052568

Wu, Y., Deng, Z., Lai, J., Zhang, Y., Yang, C., Yin, B., et al. (2009). Dual function of *Arabidopsis* ATAF1 in abiotic and biotic stress responses. *Cell Res.* 19, 1279–1290. doi: 10.1038/cr.2009.108

Xie, Q., Frugis, G., Colgan, D., and Chua, N. H. (2000). Arabidopsis NAC1 transduces auxin signal downstream of TIR1 to promote lateral root development. *Genes Dev.* 14, 3024–3036. doi: 10.1101/gad.852200

Yamaguchi, M., Ohtani, M., Mitsuda, N., Kubo, M., Ohme-Takagi, M., Fukuda, H., et al. (2010). VND-INTERACTING2, a NAC domain transcription factor, negatively regulates xylem vessel formation in *Arabidopsis*. *Plant Cell* 22, 1249–1263. doi: 10.1105/tpc.108.064048

Yoo, S. Y., Kim, Y., Kim, S. Y., Lee, J. S., and Ahn, J. H. (2007). Control of flowering time and cold response by a NAC-domain protein in *Arabidopsis*. *PLoS One* 2:e642. doi: 10.1371/journal.pone.0000642

Yuan, X., Wang, H., Cai, J., Bi, Y., Li, D., and Song, F. (2019). Rice NAC transcription factor ONAC066 functions as a positive regulator of drought and oxidative stress response. *BMC Plant Biol.* 19:278. doi: 10.1186/s12870-019-1883-y

Zhang, L., Yao, L., Zhang, N., Yang, J., Zhu, X., Tang, X., et al. (2018). Lateral root development in potato is mediated by stu-mi164 regulation of NAC transcription factor. *Front. Plant Sci.* 9:383. doi: 10.3389/fpls.2018.00383

Zhang, L., Zhang, L., Xia, C., Zhao, G., Jia, J., and Kong, X. (2015). The novel wheat transcription factor TaNAC47 enhances multiple abiotic stress tolerances in transgenic plants. *Front. Plant Sci.* 6:1174. doi: 10.3389/fpls.2015.01174

Zhao, J., Liu, J. S., Meng, F. N., Zhang, Z. Z., Long, H., Lin, W. H., et al. (2016). ANAC005 is a membrane-associated transcription factor and regulates vascular development in *Arabidopsis*. *J. Integr. Plant Biol.* 58, 442–451. doi: 10.1111/jipb.12379

Zhao, Q., Gallego-Giraldo, L., Wang, H., Zeng, Y., Ding, S. Y., Chen, F., et al. (2010). An NAC transcription factor orchestrates multiple features of cell wall development in *Medicago truncatula*. *Plant J.* 63, 100–114. doi: 10.1111/j.1365-313X.2010.04223.x

Conflict of Interest: XL, YW, WP, and JG were employed by China Tobacco Hunan Industrial Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 09 May 2022
doi: 10.3389/fpls.2022.854171

Genome-Wide Identification and Characterization of TCP Family Genes in Pak-Choi [*Brassica campestris* (syn. *Brassica rapa*) ssp. *chinensis* var. *communis*]

*Feiyi Huang†, Churan Shi†, Yuhang Zhang and Xilin Hou**

State Key Laboratory of Crop Genetics & Germplasm Enhancement, Key Laboratory of Biology and Genetic Improvement of Horticultural Crops (East China), Ministry of Agriculture and Rural Affairs of the P. R. China, Engineering Research Center of Germplasm Enhancement and Utilization of Horticultural Crops, Ministry of Education of the P. R. China, Nanjing Suman Plasma Engineering Research Institute, Nanjing Agricultural University, Nanjing, China

The TEOSINTE BRANCHED1/CYCLOIDEA/PROLIFERATING CELL FACTOR (TCP) gene family, a kind of plant specific transcription factor, is essential for stress response, cell growth, and cell proliferation. However, the characterization of TCP family is still not clear in Pak-choi [*Brassica campestris* (syn. *Brassica rapa*) ssp. *chinensis* var. *communis*]. In this study, genome-wide analysis of TCP gene family was performed and 26 *TCP* genes were identified in Pak-choi. Phylogenetic analysis demonstrated that the 26 BcTCPs were divided into two classes: Class I and Class II. Class II was further classified into two subclasses, CIN and CYC/TB1. The qPCR results suggested that most *BcTCPs* respond to abiotic stresses. The expressions of *BcTCP3*, *BcTCP12*, *BcTCP21*, and *BcTCP22* were significantly changed under ABA and cold treatment. *BcTCP3* and *BcTCP12* were also up-regulated under osmotic treatment. Subcellular localization showed that BcTCP3 and BcTCP21 were located in the nucleus. Our results will facilitate revealing the functions and regulatory mechanisms of *BcTCPs*.

Keywords: abiotic stress, expression analysis, subcellular localization, TCP, Pak-choi

OPEN ACCESS

Edited by:
Mukesh Jain,
Jawaharlal Nehru University, India

Reviewed by:
Xiao-Li Tan,
Jiangsu University, China
Jianjun Zhao,
Agricultural University of Hebei, China

***Correspondence:**
Xilin Hou
hxl@njau.edu.cn

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

Received: *13 January 2022*
Accepted: *17 March 2022*
Published: *09 May 2022*

Citation:
Huang F, Shi C, Zhang Y and Hou X (2022) Genome-Wide Identification and Characterization of TCP Family Genes in Pak-Choi [Brassica campestris (syn. Brassica rapa) ssp. chinensis var. communis]. Front. Plant Sci. 13:854171. doi: 10.3389/fpls.2022.854171

INTRODUCTION

The TCP gene family is a class of plant-specific transcription factors and plays important roles in plant growth and development. TCP was named from four genes members: *TEOSINTE BRANCHED 1* (*TB1*) from maize (*Zea mays*) (Doebley et al., 1995), *CYCLOIDEA* (*CYC*) from snapdragon (*Antirrhinum majus*) (Luo et al., 1996), and the *PROLIFERATING CELL FACTORS 1* (*PCF1*) and *PCF2* from rice (*Oryza sativa*) (Kosugi and Ohashi, 1997). *TB1* is a key factor to control apical dominance in maize (Doebley et al., 1997). *CYC* involves in the regulation of floral asymmetry in snapdragons. PCF1 and PCF2 bind to the promoter of *PROLIFERATING CELL NUCLEAR ANTIGEN* (*PCNA*) gene in rice, which involves in DNA replication and repair, maintenance of chromatin structure, chromosome segregation, and cell-cycle progression (Luo et al., 1996; Kosugi and Ohashi, 1997). TCP proteins contain a 59-amino-acid domain with a bHLH motif that involved in DNA binding and protein-protein interaction (Martin-Trillo and Cubas, 2010). According to the

conserved domain, TCP proteins can be divided into two classes: Class I (TCP-P Class) and Class II (TCP-C Class) (Kosugi and Ohashi, 2002; Navaud et al., 2007). Class II can be further divided into two subclades: CIN and CYC/TB1 (Martin-Trillo and Cubas, 2010). Class I includes *PCF1* and *PCF2* in rice, *TCP8*, *TCP9*, *TCP14*, *TCP20*, etc. in *Arabidopsis*. Class II contains *TB1* in maize, *TCP1*, *TCP2*, *TCP4*, *TCP10*, *BCR1*, *BCR2*, etc. in *Arabidopsis*.

Abiotic stress has an important effect on plant growth and productivity. Cold stress slows down the metabolic processes of plants, inhibits auxin transport and root elongation (Shibasaki et al., 2009). Salt stress leads to the accumulation of toxic ions in cells, which influences plant absorption of nutrients and water (Boudsocq and Laurière, 2005). Plants are subjected to dehydration under low temperature, high salinity, and drought environmental conditions, of which cause osmotic stress, inducing Abscisic acid (ABA) biosynthesis. ABA functions in multiple stresses, regulating genes related to dehydration and cold stress and controlling osmotic stress tolerance (Yamaguchi-Shinozaki and Shinozaki, 2006; Fujita et al., 2011). Moreover, ABA can reduce water loss by governing stomatal closure (Kim et al., 2010). The growth and development of plants will be severely affected when plants suffer from abiotic stresses. Hence, the study of response mechanisms during abiotic stresses is vital for species. At present, it is reported that many genes of the TCP family are related to various stresses. TCP20 responds to nitrate availability by interacting with NIN-like protein 6 (NLP6) and NLP9, controlling plants root growth (Guan et al., 2017). *OsTCP19* is up-regulated under drought, salt, and cold stress, indicating *OsTCP19* may involve in the stress tolerance (Mukhopadhyay and Tyagi, 2015). In *Cicer arietinum*, five TCP genes (*CaTCP3*/*13*/*15*/*20*/*21*), which contained the MYB *cis*-elements, were strongly induced under drought conditions, and similar results were found in other legumes (Ling et al., 2020). *Oryza sativa miR319* (*Osa-miR319*), an upstream gene of *OsPCF5*, *OsPCF6*, *OsPCF8*, and *OsTCP21*, is decreased in the cold condition in rice. Meanwhile, overexpressing *Osa-miR319* results in down-regulation of its downstream target genes and enhances cold resistance, impling that TCP21 could involve in cold stress (Yang et al., 2013; Wang et al., 2014). In creeping bentgrass (*Agrostis stolonifera*), overexpressing *Osa-miR319* also reduces the expression of its target genes (*AsPCF5*, *AsPCF6*, *AsPCF8*, and *AsTCP14*), enhancing salt and drought tolerance (Zhou et al., 2013). GmTCP8 interacts with GmPYL10 and involves in the ABA signal pathway in soybean (Feng et al., 2018). *ZmTCP42* is associated with ABA and drought stress, which plays an active role in drought tolerance (Ding et al., 2019). AtTCP14 represses *ABA1* (*ABA DEFICIENT 1*) and other ABA-related stress genes in *Arabidopsis* seeds (Rueda-Romero et al., 2012).

Until now, the study of the TCP family has been more and more comprehensive, which have been identified in many species including *Arabidopsis thaliana* (Li, 2015), rice (*Oryza sativa*) (Yao et al., 2007), turnips (*Brassica rapa* ssp. *rapa*) (Du et al., 2017), poplar (*Populus trichocarpa*) (Ma et al., 2016), tomato (*Solanum lycopersicum*) (Parapunova et al., 2014), Switchgrass (*Panicum virgatum* L.) (Huo et al., 2019), *Brassica rapa* and *Brassica oleracea* (Liu et al., 2019). Pak-choi [*Brassica campestris* (syn. *Brassica rapa*) ssp. *chinensis* var. *communis*] has gradually become worldwide vegetable crops. However, Pak-choi *TCP* genes and their regulatory mechanisms are still not clear.

In this study, genome-wide analysis of TCP family genes in Pak-choi was performed. 26 *BcTCP* genes were identified in Pak-choi. Phylogenetic relationship, subcellular localization, and conserved motifs were analyzed. QPCR was employed to illustrate the expression patterns of *BcTCPs* under salt, osmotic, cold, and ABA treatment. The results showed that *BcTCPs* may involve in multiple abiotic stresses. Our findings provide a theoretical basis for further research on the potential functions and regulatory mechanisms of *BcTCP* genes.

MATERIALS AND METHODS

Cloning and Identification of *BcTCP* Gene Family Members

Total RNA was extracted from Pak-choi using the RNAeasy Mini Kit (Tiangen, Beijing, China). The CDSs of the *BcTCPs* were amplified with gene specific primers based on the sequences of *AtTCPs* by homology cloning (Huang et al., 2019 and **Supplementary Table 1**). The PCR products were cloned into the pMD18-T vector before sequencing. Then, the obtained plasmids were sequenced by TSINGKE (Nanjing, China). All putative BcTCP proteins sequences were analyzed using the Pfam database[1]. Additionally, protein molecular weight, isoelectric point, and amino acid length of BcTCPs were computed by the ExPASy ProtParam tool[2].

Phylogenetic Tree and *Cis*-Acting Elements Analysis

24 *AtTCP* genes were retrieved from TAIR[3]. Multiple sequence alignments of putative BcTCPs and AtTCPs were performed using Clustal X 2.1. Phylogenetic tree was completed by the Maximum Likelihood (ML) method using MEGA 7.0 software. The bootstrap values were performed with 1,000 replications. PlantCARE[4] was used to analyze *cis*-acting elements in promoters of each *BcTCP* gene. After classifying and counting the elements on each promoter (**Supplementary Table 2**), the heat map was made by TBtools.

Conserved Domain Analysis

The conserved motifs were analyzed using MEME program[5] with the default settings except the maximum width was set to 200, and the minimum and maximum numbers of motifs were defifined as 2 and 10, respectively.

[1]http://pfam.xfam.org/

[2]https://web.expasy.org/protparam/

[3]https://www.arabidopsis.org/

[4]http://bioinformatics.psb.ugent.be/webtools/plantcare/html/

[5]http://meme-suite.org/

Plant Materials and Growth Conditions

The Pak-choi cultivar '49caixin' was selected in our experiment, which has a short life cycle and may flower within two months after sowing under long-day conditions (Tian et al., 2004). Plants were cultivated in pots with medium (Soil matrix and vermiculite, 1:1). All seedlings were grown in a controlled artificial climatic chamber under the same conditions (16 h light at 22°C/8 h dark at 18°C, 60-70% relative humidity). One-month-old seedlings were treated with 100 μM ABA, 250 mM NaCl and 20% PEG6000, respectively. For cold treatment, one-month-old seedlings were transferred to 4°C. The leaves were collected under each stress treatment in a continuous time course (0, 10, 20, 30, 40, 50 min, 1, 2, 3, 4, 8 h) and stored at −70°C immediately for RNA extraction.

Expression Analysis of *BcTCP* Genes by qPCR

Following the manufacturer's instructions, total RNA was isolated from leaves using the RNA easy mini kit (Tiangen, Beijing, China). Then first-strand cDNA was synthesized with a RevertAid First Strand cDNA Synthesis Kit (Thermo, Shanghai, China). Primers for qPCR were designed using Primer 5.0 (**Supplementary Table 1**). QPCR was performed using the SYBR® Premix Ex Taq kit (Takara, Dalian, China). *BcActin* was used as an internal reference gene. Results were calculated by the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001).

Subcellular Localization Analysis

WOLF PSORT[6] was used to predict the subcellular localization of the putative BcTCP proteins. To further confirm their subcellular localization, the protein-coding regions of *BcTCP3* and *21* without the termination codon were amplified and then cloned into pCambia 1,302 vector in fusion with GFP at C-terminal end (*35S: GFP*), generating the fusion constructs. The *35S: BcTCP3-GFP* and *35S: BcTCP21-GFP* were generated. Each plasmid was injected into the tobacco leaves by *Agrobacterium tumefaciens*-mediated transient transformation (strain GV3101). The tobacco was incubated for 12 h under darkness. 2-3 days after injection, the fluorescence of GFP was observed and photographed by confocal microscopy (Leica, TCS SP2, Wetzlar, Germany). DAPI (nucleus specific dye) was used as nuclei dye in the experiment.

RESULTS

Identification of *TCP*s in Pak-Choi

A total of 26 *TCP* genes in Pak-choi were obtained by homology cloning based on the *Arabidopsis thaliana TCP* genes. These 26 genes were named *BcTCP1*∼*BcTCP26*. The CDS length of the 26 *BcTCP*s was ranged from 651 to 1,395 bp. The physical and chemical properties were further analyzed using the ExPASy ProtParam tool. The molecular weight of 26 putative BcTCP proteins varied from 23.12 to 56.74 kDa. The isoelectric point was ranged from 5.49 to 9.99, and the amino acid length was ranged from 216 to 520 aa (**Supplementary Table 3**).

Phylogenetic Analysis

An unrooted phylogenetic tree was constructed with the Maximum Likelihood (ML) method (**Figure 1**). 26 *BcTCPs* were divided into two classes, Class I and Class II. The Class II *TCP*s were further categorized into two subgroups, CIN and CYC/TB1. PCF Class (Class I) contained 13 genes, CYC/TB1 Class contained 7 genes, CIN only contained 6 genes. Sequence analysis showed that *AtTCP3*, *6*, *11*, *13*, *16*, *17*, *20*, *23* had no ortholog in Pak-choi. *AtTCP1*, *7*, *15*, *18*, *21*, and *24* had more than one ortholog in the Pak-choi.

Conserved Domain Analysis

The conserved motifs of BcTCP proteins were analyzed using MEME. A total of 7 motifs were found in BcTCPs. The motifs were more similar in the same class. Furthermore, similar conserved motifs arrangements can be observed in the same clade. Motif 1 was conserved in all BcTCPs. Motif 3 and motif 6 only existed in Class II except for BcTCP5, 6, 11, and 13 (**Figure 1**), suggesting that these BcTCPs may have similar functions. However, proteins among different groups shared diverse motifs. Motif 2 is presented in Class I (PCF) except for BcTCP20. Motif 7 was only existed in BcTCP15, 16, and 17. Motif 4 and motif 5 both existed in BcTCP7, BcTCP8, BcTCP21, BcTCP22, and BcTCP23. These results indicated that TCP genes may perform different functions in Pak-choi.

Analysis of *Cis*-Acting Elements in *BcTCPs* Promoters

Cis-acting elements existed in gene promoters can affect gene expressions and functions. In this study, we found that *BcTCPs* promoters contained not only basic core elements, like TATA-box and CAAT-box, but also a variety of *cis*-acting elements, such as light response elements, hormone response elements, environment response elements, development response elements, and other functional elements (**Figure 2**). All Class II *BcTCPs* promoters contained G-box elements and ABRE elements except for *BcTCP20* and *BcTCP21*. In Class I *BcTCPs*, the promoters of *BcTCP2*, *3*, *11*, *12*, *13*, *25*, and *26* all contained the low temperature response element (LTR). AT∼TATA box1 was identified in most all *BcTCP* promoters except *BcTCP6* and *BcTCP23*. The promoters of *BcTCP2*, *3*, *4*, *10*, *11*, *12*, *14*, *15*, *20*, *21*, *22*, *26* contained MYB, MYC, and MYB-like sequence elements at the same time. As a result, we hypothesized that these *BcTCP* genes may be involved in stress (Abe et al., 2003; Xu et al., 2021).

Expression Analysis of *BcTCP* Genes Under Multiple Abiotic Stresses

To investigate the functions of BcTCPs, qPCR was performed to detect the expressions of 26 *BcTCPs* under four treatments (salt, osmotic, cold and ABA) (**Figure 3**). The results showed that the expressions of most *BcTCPs*, except for *BcTCP12*, were

[6]http://wolfpsort.org/

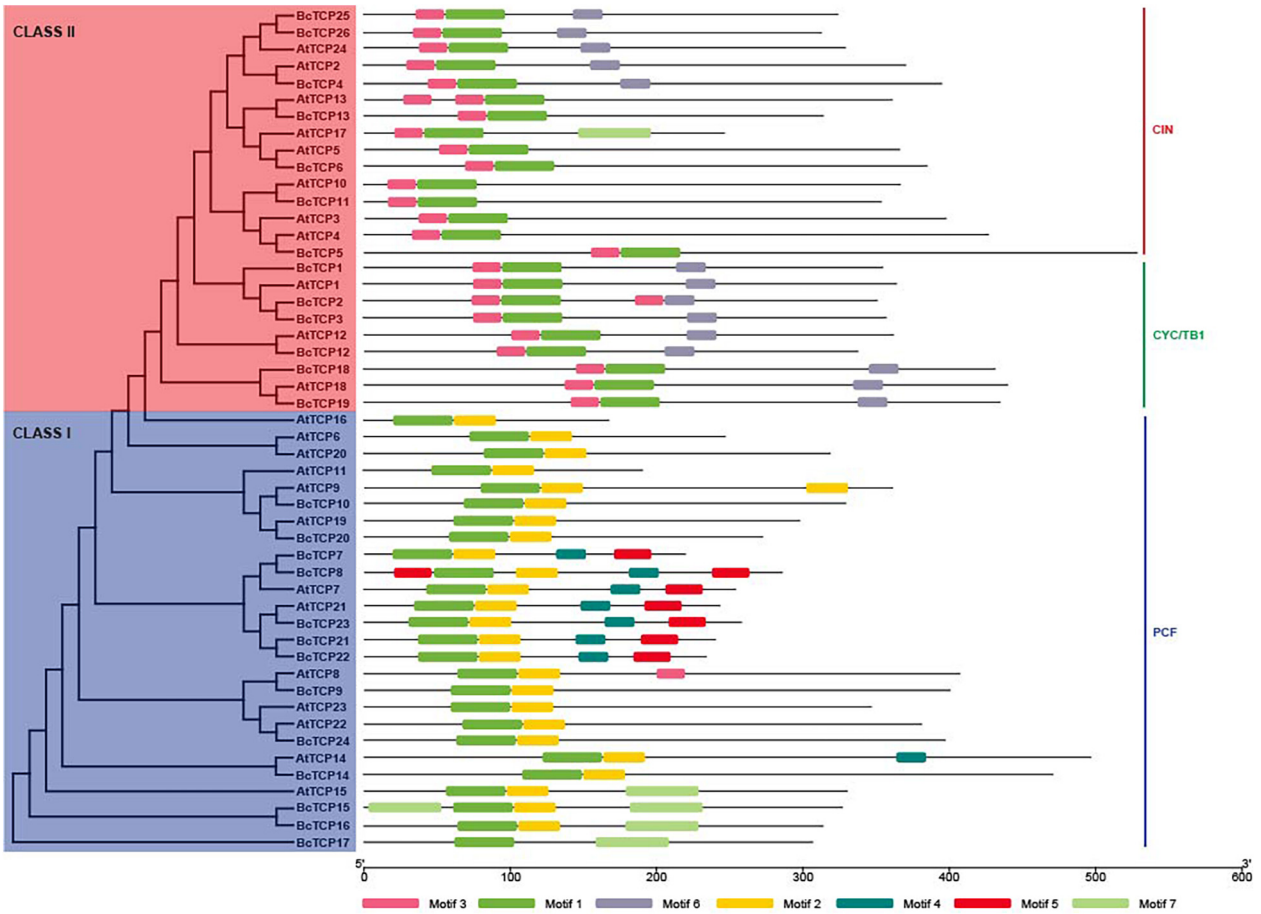


FIGURE 1 | The phylogenetic tree and conserved motifs of BcTCP and AtTCP proteins. Each motif represented with a specific color was shown.

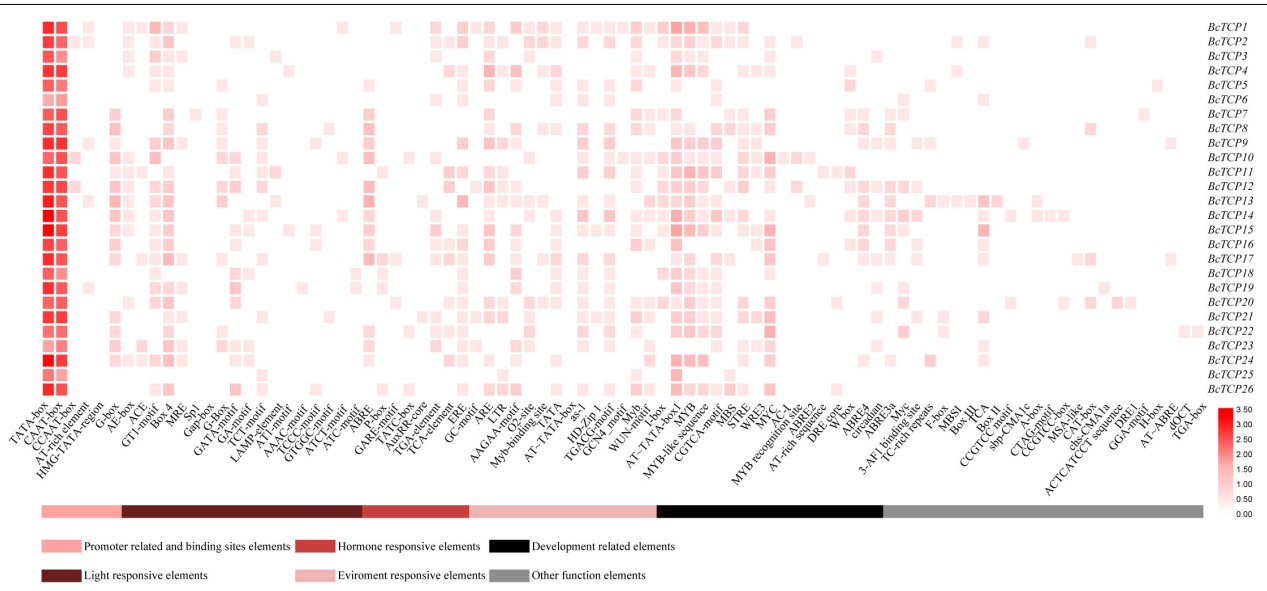


FIGURE 2 | *Cis*-acting elements on promoters of *BcTCP* genes. The color bar showed the number of *cis*-acting elements.

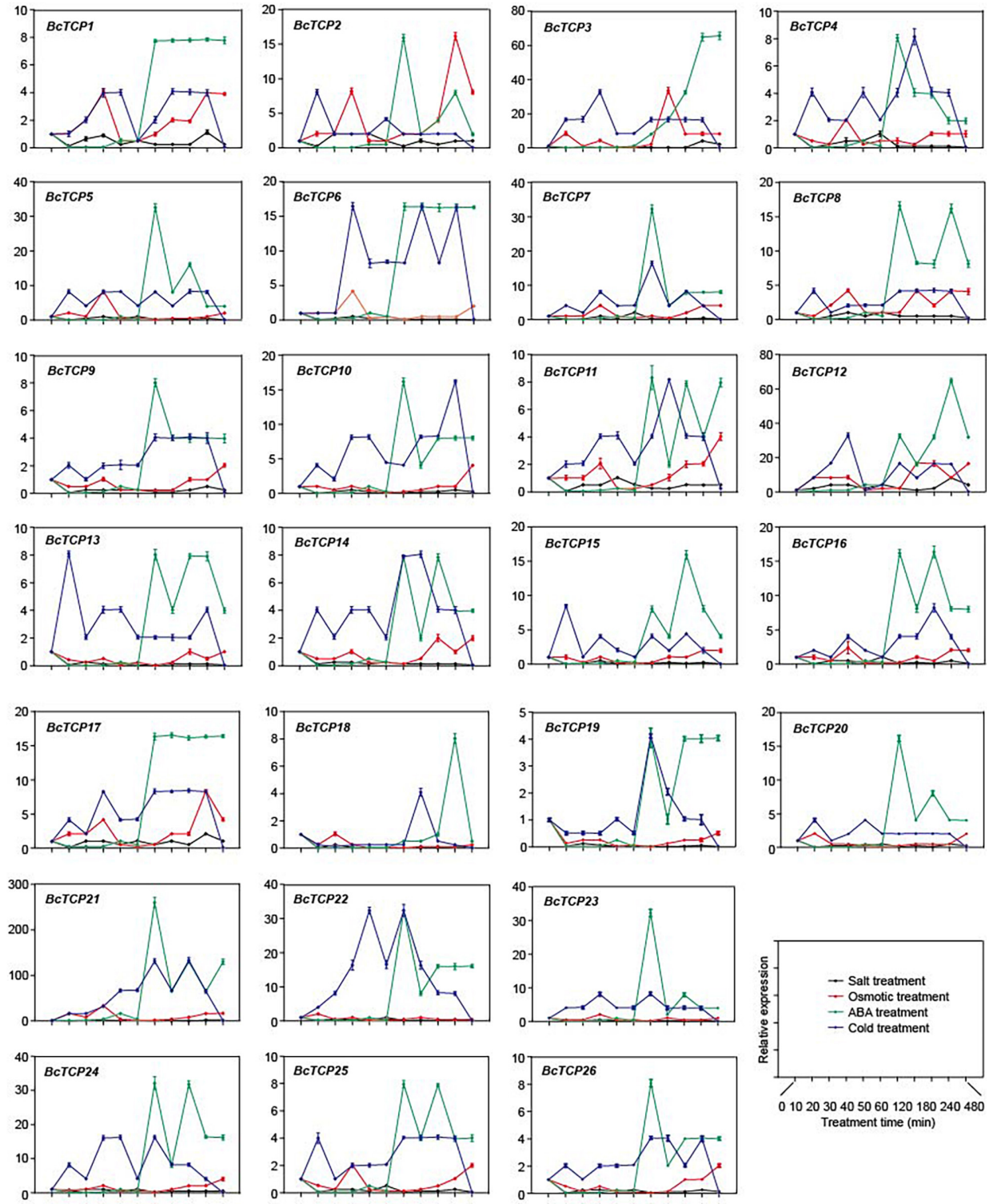


FIGURE 3 | Expression of *BcTCP* genes under different stress treatments. Data shown were means ± SE of three independent experiments.

not significantly changed under salt treatment. *BcTCP2*, *3*, *5*, *6*, *7*, *8*, *10*, *12*, *15*, *16*, *17*, *20*, *21*, *22*, *23*, and *24* all up-regulated after 1 h of the ABA treatment. The expression levels of *BcTCP2*, *5*, *6*, *7*, *8*, *10*, *17*, *20*, *21*, *22*, *23*, and *24* reached the maximum at 2 h, *BcTCP15* and *16* had the highest expression at 3h, the largest amount of *BcTCP12* expressed in 4 h, and the

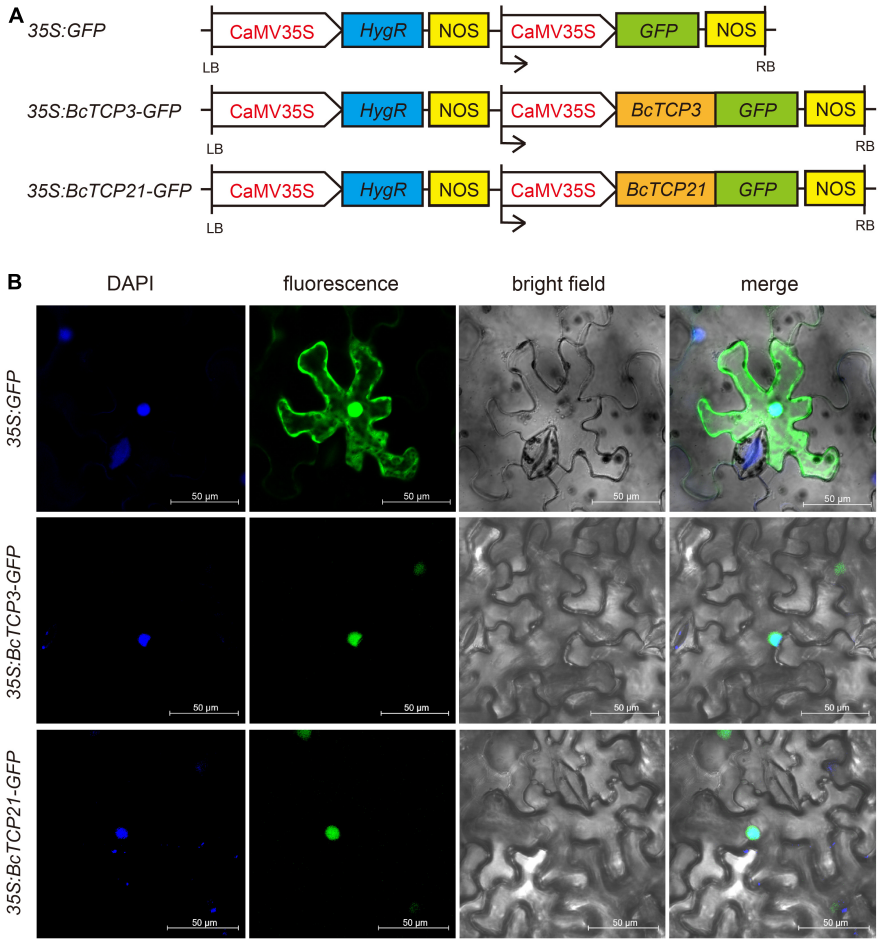


FIGURE 4 | Subcellular localization of BcTCP3 protein and BcTCP21 protein: **(A)** *35S: GFP*, *35S: BcTCP3*, and 35S: BcTCP21 construct. **(B)** Transient expression of 35S:GFP, 35S: 35S: BcTCP3-GFP, and 35S: BcTCP21-GFP fusion protein in tobacco leaves (Scale bars = 50 μm).

largest amount of *BcTCP3* expression quantity in 4 h. Under cold treatment, *BcTCP3*, *6*, *7*, *10*, *12*, *21*, *22*, and *24* showed a significant increase, indicating that these genes, especially *BcTCP21*, may involve in cold stress. Meanwhile, *BcTCP2*, *3*, and *12* were notably up-regulated under osmotic treatment. *BcTCP2* reached its peak at 4 h, and *BcTCP3* and *12* reached their peak at 2 h. These suggested that most *BcTCP* genes were involved in abiotic stresses.

Subcellular Localization

WoLF PSORT predicted that BcTCP proteins were localized in the nucleus (**Supplementary Table 3**). To test this, *35S: GFP*, *35S: BcTCP3-GFP*, and *35S: BcTCP21-GFP* constructs were transiently overexpressed in tobacco leaves (**Figure 4A**). As seen in **Figure 4B**, *35S: GFP* was localized in the nucleus and cytoplasm. The GFP signals emitted by BcTCP3 -GFP and BcTCP21 -GFP fusion protein were detected in the nucleus and overlapped with the DAPI staining. This results indicated that BcTCP3 and BcTCP21 were nuclear-localized proteins and may function as transcription factors.

DISCUSSION

TCP gene family involves in multiple biological processes, including the circadian clock, abiotic stresses, leaf sizes and shapes, flower development and flowering (Takeda et al., 2006; Pruneda-Paz et al., 2009; Mimida et al., 2011; Niwa et al., 2013; Liu et al., 2017; Feng et al., 2018). The roles of TCP genes under abiotic stresses have been reported in some model plants, such as rice, *Zea mays*, *Arabidopsis thaliana*, and soybean (Rueda-Romero et al., 2012; Mukhopadhyay and Tyagi, 2015; Feng et al., 2018; Ding et al., 2019). However, virtually less systematic and comprehensive information of the TCP gene family in Pak-choi was reported, which is a nutritious and economically important vegetable crop and widely cultivated in Asia. Unlike *A. thaliana*, *Brassicas* crops not only underwent this complex evolutionary history, but also the whole-genome triplication (WGT) event between 13 and 17 million years ago (MYA). These events were essential for evolution and resulted in differences among species (Wang et al., 2011). *A. thaliana* has been identified 24 *TCP* genes (Yao et al., 2007). In this study, 26 *BcTCP* genes were isolated and studied in Pak-choi. These *BcTCP*s were divided into Class I

and Class II, Class II was then divided into CIN and CYC/TB1, which was consistent with the previously described in *Arabidopsis*, rice, tomato, and strawberry (Yao et al., 2007; Parapunova et al., 2014; Wei et al., 2016). The phylogenetic analysis found that 16 *AtTCP*s had orthologs in Pak-choi. Among them, *AtTCP1*, *7*, *15*, *18*, *21*, and *24* had more than one ortholog. This indicated that BcTCP family genes and AtTCP family genes may have some similar functions. According to conserved domain, we found that the similarity of BcTCPs in the same class is higher than that of different classes of BcTCPs. For example, CYB/TB1 members all contained motif 1, motif 3, and motif 6. Analysis of promoter regions showed that some *BcTCPs* promoters contained MBS *cis*-regulatory elements, MYB *cis*-regulatory elements, ABRE elements, and G-box, suggesting that the *BcTCP* genes might play significant roles in stress responses (Abe et al., 2003; Xu et al., 2021).

QPCR analysis demonstrated that some *BcTCP*s were involved in multiple abiotic stress. For instance, *BcTCP3*, *12*, *21*, *22*, and *24* all responded to cold and ABA treatment. *BcTCP3* had no response to salt stress but had relatively obvious responses under cold and ABA treatments. *BcTCP12* responded to salt, osmotic, ABA, and cold treatments. Under salt treatment, the expression of *BcTCP12* was significantly activated, indicating that *BcTCP12* may enhance salt tolerance, which is same as previous studies in Oryza sativa (Xiao et al., 2009; Almeida et al., 2017). The expressions of *BcTCP3*, *5*, *7*, *12*, *21*, *22*, *23*, and *24* displayed more than 30 fold up-regulation under ABA treatment. During drought treatment, *BcTCP2*, *3*, and *12* were up-regulated by more than 15 fold. *BcTCP3*, *12*, *21*, and *22* were also up-regulated by more than 30 fold under cold treatment. These results were consistent with previous research results in rice (Yang et al., 2013; Wang et al., 2014). Subcellular localization analysis demonstrated BcTCP3 and BcTCP21 were located in the nucleus, indicating they may function as transcription factors.

These results revealed that *BcTCP*s were associated with multiple abiotic stresses. Our study may be helpful for improving plant stress tolerance. This will provide a piece of vital evidence for detecting the molecular mechanisms of *TCP*s under abiotic stresses. The systematic characterization of *BcTCP*s in Pak-choi will provide a better foundation for further functional studies of this gene family in plant growth and development.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

FH, CS, and YZ completed the experiments and wrote the manuscript. XH revised and approved the manuscript. All authors have read and agreed to the published version of the manuscript.

FUNDING

This work was supported by grants from the National Natural Science Foundation of China (32102410), the Fundamental Research Funds for the Central Universities (KYQN2022054), the Natural Science Foundation of Jiangsu province (BK20190513), the Key Projects of National Key Research and Development Plan (2017YFD0101803) and the China Agriculture Research System (CARS-23-A-06).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.854171/full#supplementary-material

REFERENCES

Abe, H., Urao, T., Ito, T., Seki, M., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2003). *Arabidopsis* AtMYC2 (bHLH) and AtMYB2 (MYB) Function as Transcriptional Activators in Abscisic Acid Signaling. *Plant Cell* 155, 63–78. doi: 10.1105/tpc.006130

Almeida, D. M., Gregorio, G. B., Oliveira, M. M., and Saibo, N. J. M. (2017). Five noveln transcription factors as potential regulators of *OsNHX1* gene expression in a salt tolerant rice genotype. *Plant Mol. Biol.* 93, 61–77. doi: 10.1007/s11103-016-0547-7

Boudsocq, M., and Laurière, C. (2005). Osmotic Signaling in Plants. Multiple Pathways Mediated by Emerging Kinase Families. *Plant Physiol.* 138, 1185–1194. doi: 10.1104/pp.105.061275

Ding, S., Cai, Z., Du, H., and Wang, H. (2019). Genome-Wide Analysis of TCP Family Genes in *Zea mays* L. Identified a Role for *ZmTCP42* in Drought Tolerance. *Int. J. Mol. Sci.* 20:2762. doi: 10.3390/ijms20112762

Doebley, J., Stec, A., and Gustus, C. (1995). *teosinte branched1* and the Origin of Maize: evidence for Epistasis and the Evolution of Dominance. *Genetics* 141, 333–346. doi: 10.1093/genetics/141.1.333

Doebley, J., Stec, A., and Hubbard, L. (1997). The evolution of apical dominance in maize. *Nature* 386, 485–488. doi: 10.1038/386485a0

Du, J., Hu, S., Yu, Q., Wang, C., Yang, Y., Sun, H., et al. (2017). Genome-Wide Identification and Characterization of BrrTCP Transcription Factors in *Brassica rapa* ssp. *rapa*. *Front. Plant Sci.* 8:1588. doi: 10.3389/fpls.2017.01588

Feng, Z. J., Xu, S. C., Liu, N., Zhang, G. W., Hu, Q. Z., and Gong, Y. M. (2018). Soybean TCP transcription factors: evolution, classification, protein interaction and stress and hormone responsiveness. *Plant Physiol. Biochem.* 127, 129–142. doi: 10.1016/j.plaphy.2018.03.020

Fujita, Y., Fujita, M., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2011). ABA-mediated transcriptional regulation in response to osmotic stress in plants. *J. Plant Res.* 124, 509–525. doi: 10.1007/s10265-011-0412-3

Guan, P., Ripoll, J. J., Wang, R., Vuong, L., Bailey-Steinitz, L. J., Ye, D., et al. (2017). Interacting TCP and NLP transcription factors control plant responses to nitrate availability. *Proc. Natl. Acad. Sci. U. S. A.* 114, 2419–2424. doi: 10.1073/pnas.1615676114

Huang, F. Y., Liu, T. K., Tang, J., Duan, W. K., and Hou, X. L. (2019). BcMAF2 activates *BcTEM1* and represses flowering in Pak-choi (*Brassica rapa* ssp. *chinensis*). *Plant Mol. Biol.* 100, 19–32. doi: 10.1007/s11103-019-00867-1

Huo, Y., Xiong, W., Su, K., Li, Y., Yang, Y., Fu, C., et al. (2019). Genome-Wide Analysis of the *TCP* Gene Family in Switchgrass (*Panicum virgatum* L.). *Int. J. Genomics* 2019, 1–13. doi: 10.1155/2019/8514928

Kim, T. H., Böhmer, M., Hu, H., Nishimura, N., and Schroeder, J. I. (2010). Guard Cell Signal Transduction Network: advances in Understanding Abscisic Acid, CO_2, and Ca^{2+} Signaling. *Annu. Rev. Plant Biol.* 61, 561–591. doi: 10.1146/annurev-arplant-042809-112226

Kosugi, S., and Ohashi, Y. (1997). PCF1 and PCF2 Specifically Bind to *cis* Elements in the Rice Proliferating Cell Nuclear Antigen Gene. *Plant Cell* 9, 1607–1619. doi: 10.2307/3870447

Kosugi, S., and Ohashi, Y. (2002). DNA binding and dimerization specificity and potential targets for the TCP protein family. *Plant J.* 30, 337–348. doi: 10.1046/j.1365-313X.2002.01294.x

Li, S. (2015). The *Arabidopsis thaliana* TCP transcription factors: a broadening horizon beyond development. *Plant Signal. Behav.* 10:e1044192. doi: 10.1080/15592324.2015.1044192

Ling, L., Zhang, W., An, Y., Du, B., Wang, D., and Guo, C. (2020). Genome-wide analysis of the *TCP* transcription factor genes in five legume genomes and their response to salt and drought stresses. *Funct. Integr. Genomics* 20, 537–550. doi: 10.1007/s10142-020-00733-0

Liu, J., Cheng, X., Liu, P., Li, D., Chen, T., Gu, X., et al. (2017). MicroRNA319-regulated TCPs interact with FBHs and PFT1 to activate *CO* transcription and control flowering time in *Arabidopsis*. *PLoS Genet.* 13:e1006833. doi: 10.1371/journal.pgen.1006833

Liu, M. M., Wang, M. M., Yang, J., Wen, J., Guo, P. C., Wu, Y. W., et al. (2019). Evolutionary and Comparative Expression Analyses of TCP Transcription Factor Gene Family in Land Plants. *Int. J. Mol. Sci.* 20:3591. doi: 10.3390/ijms20143591

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of Relative Gene Expression Data Using Realtime Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods* 25, 402–408. doi: 10.1006/meth.2001

Luo, D., Carpenter, R., Vincent, C., Copsey, L., and Coen, E. (1996). Origin of floral asymmetry in *Antirrhinum*. *Nature* 383, 794–799. doi: 10.1038/383794a0

Ma, X., Ma, J., Fan, D., Li, C., Jiang, Y., and Luo, K. (2016). Genome-wide Identification of TCP Family Transcription Factors from *Populus euphratica* and their Involvement in Leaf Shape Regulation. *Sci. Rep.* 6:32795. doi: 10.1038/srep32795

Martin-Trillo, M., and Cubas, P. (2010). TCP genes: a family snapshot ten years later. *Trends Plant Sci.* 15, 31–39. doi: 10.1016/j.tplants.2009.11.003

Mimida, N., Kidou, S., Iwanami, H., Moriya, S., Abe, K., Voogd, C., et al. (2011). Apple FLOWERING LOCUS T proteins interact with transcription factors implicated in cell growth and organ development. *Tree Physiol.* 31, 555–566. doi: 10.1093/treephys/tpr028

Mukhopadhyay, P., and Tyagi, A. K. (2015). *OsTCP19* influences developmental and abiotic stress signaling by modulating ABI4-mediated pathways. *Sci. Rep.* 5:9998. doi: 10.1038/srep09998

Navaud, O., Dabos, P., Carnus, E., Tremousaygue, D., and Hervé, C. (2007). TCP Transcription Factors Predate the Emergence of Land Plants. *J. Mol. Evol.* 65, 23–33. doi: 10.1007/s00239-006-0174-z

Niwa, M., Daimon, Y., Kurotani, K., Higo, A., Pruneda-Paz, J. L., Breton, G., et al. (2013). BRANCHED1 Interacts with FLOWERING LOCUS T to Repress the Floral Transition of the Axillary Meristems in *Arabidopsis*. *Plant Cell* 25, 1228–1242. doi: 10.1105/tpc.112.109090

Parapunova, V., Busscher, M., Busscher-Lange, J., Lammers, M., Karlova, R., Bovy, A. G., et al. (2014). Identification, cloning and characterization of the tomato TCP transcription factor family. *BMC Plant Biol.* 14:157. doi: 10.1186/1471-2229-14-157

Pruneda-Paz, J. L., Breton, G., Para, A., and Kay, S. A. (2009). A functional genomics approach reveals CHE as a component of the *Arabidopsis* circadian clock. *Science* 323, 1481–1485. doi: 10.1126/science.1167206

Rueda-Romero, P., Barrero-Sicilia, C., Gómez-Cadenas, A., Carbonero, P., and Oñate-Sánchez, L. (2012). *Arabidopsis thaliana* DOF6 negatively affects germination in non-after-ripened seeds and interacts with TCP14. *J. Exp. Bot.* 63, 1937–1949. doi: 10.1093/jxb/err388

Shibasaki, K., Uemura, M., Tsurumi, S., and Rahman, A. (2009). Auxin Response in *Arabidopsis* under Cold Stress: underlying Molecular Mechanisms. *Plant Cell* 21, 3823–3838. doi: 10.1105/tpc.109.069906

Takeda, T., Amano, K., Ohto, M. A., Nakamura, K., Sato, S., Kato, T., et al. (2006). RNA interference of the *Arabidopsis* putative transcription factor *TCP16* gene results in abortion of early pollen development. *Plant Mol. Biol.* 61, 165–177. doi: 10.1007/s11103-006-6265-9

Tian, C., Wan, P., Sun, S., Li, J., and Chen, M. (2004). Genome-wide analysis of the GRAS gene family in rice and *Arabidopsis*. *Plant Mol. Biol.* 54, 519–532. doi: 10.1023/B:PLAN.0000038256.89809.57

Wang, S. T., Sun, X. L., Hoshino, Y., Yu, Y., Jia, B., Sun, Z. W., et al. (2014). *MicroRNA319* Positively Regulates Cold Tolerance by Targeting *OsPCF6* and *OsTCP21* in Rice (*Oryza sativa* L.). *PLoS One* 9:e91357. doi: 10.1371/journal.pone.0091357

Wang, X. W., Wang, H. Z., Wang, J., Sun, R., Wu, J., Liu, S. Y., et al. (2011). The genome of the mesopolyploid crop species *Brassica rapa*. *Nat. Genet.* 43, 1035–1039. doi: 10.1038/ng.919

Wei, W., Hu, Y., Cui, M. Y., Han, Y. T., Gao, K., and Feng, J. Y. (2016). Identification and Transcript Analysis of the TCP Transcription Factors in the Diploid Woodland Strawberry *Fragaria vesca*. *Front. Plant Sci.* 7:1937. doi: 10.3389/fpls.2016.01937

Xiao, B. Z., Chen, X., Xiang, C. B., Tang, N., Zhang, Q. F., and Xiong, L. Z. (2009). Evaluation of Seven Function-Known Candidate Genes for their Effects on Improving Drought Resistance of Transgenic Rice under Field Conditions. *Mol. Plant* 2, 73–83. doi: 10.1093/mp/ssn068

Xu, J. L., Zhang, W. J., and Xiang, F. N. (2021). Advances in stress inducible promoter and *cis-acting* elements in higher plants. *J. Plant Physiol.* 57, 759–766. doi: 10.13592/j.cnki.ppj.2020.0221

Yamaguchi-Shinozaki, K., and Shinozaki, K. (2006). Transcriptional Regulatory Networks in Cellular Responses and Tolerance to Dehydration and Cold Stresses. *Annu. Rev. Plant Biol.* 57, 781–803. doi: 10.1146/annurev.arplant.57.032905.105444

Yang, C., Li, D., Mao, D., Liu, X., Ji, C., Li, X., et al. (2013). Overexpression of microRNA319 impacts leaf morphogenesis and leads to enhanced cold tolerance in rice (*Oryza sativa* L.). *Plant Cell Environ.* 36, 2207–2218. doi: 10.1111/pce.12130

Yao, X., Ma, H., Wang, J., and Zhang, D. B. (2007). Genome-Wide Comparative Analysis and Expression Pattern of TCP Gene Families in *Arabidopsis thaliana* and *Oryza sativa*. *J. Integr. Plant Biol.* 49, 885–897. doi: 10.1111/j.1744-7909.2007.00509.x

Zhou, M., Li, D., Li, Z., Hu, Q., Yang, C., Zhu, L., et al. (2013). Constitutive Expression of a *miR319* Gene Alters Plant Development and Enhances Salt and Drought Tolerance in Transgenic Creeping Bentgrass. *Plant Physiol.* 161, 1375–1391. doi: 10.1104/pp.112.208702

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 10 May 2022
doi: 10.3389/fpls.2022.870586

The Wheat Gene *TaVQ14* Confers Salt and Drought Tolerance in Transgenic *Arabidopsis thaliana* Plants

Xinran Cheng[1,2†], Hui Yao[1†], Zuming Cheng[1], Bingbing Tian[1], Chang Gao[1], Wei Gao[1], Shengnan Yan[1], Jiajia Cao[1], Xu Pan[1], Jie Lu[1], Chuanxi Ma[1], Cheng Chang[1] and Haiping Zhang[1]**

[1] College of Agronomy, Anhui Agricultural University, Key Laboratory of Wheat Biology and Genetic Improvement on Southern Yellow and Huai River Valley, Ministry of Agriculture and Rural Affairs, Hefei, China, [2] National Key Laboratory for Crop Genetics and Germplasm Enhancement, Jiangsu Plant Gene Engineering Research Center, Nanjing Agricultural University, Nanjing, China

OPEN ACCESS

Edited by:
Rajeev K. Varshney, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), India

Reviewed by:
Deepak Kumar, Banaras Hindu University, India
Ágnes Szepesi, University of Szeged, Hungary

***Correspondence:**
Cheng Chang
hangtgw@163.com
Haiping Zhang
zhanghaiping@ahau.edu.cn

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

Received: 07 February 2022
Accepted: 04 April 2022
Published: 10 May 2022

Citation:
Cheng X, Yao H, Cheng Z, Tian B, Gao C, Gao W, Yan S, Cao J, Pan X, Lu J, Ma C, Chang C and Zhang H (2022) The Wheat Gene TaVQ14 Confers Salt and Drought Tolerance in Transgenic Arabidopsis thaliana Plants. Front. Plant Sci. 13:870586. doi: 10.3389/fpls.2022.870586

Wheat is one of the most widely cultivated food crops worldwide, and the safe production of wheat is essential to ensure food security. Soil salinization and drought have severely affected the yield and quality of wheat. Valine-glutamine genes play important roles in abiotic stress response. This study assessed the effect of the gene *TaVQ14* on drought and salt stresses resistance. Sequence analysis showed that TaVQ14 encoded a basic unstable hydrophobic protein with 262 amino acids. Subcellular localization showed that TaVQ14 was localized in the nucleus. *TaVQ14* was upregulated in wheat seeds under drought and salt stress. Under NaCl and mannitol treatments, the percentage of seed germination was higher in Arabidopsis lines overexpressing *TaVQ14* than in wild-type lines, whereas the germination rate was significantly lower in plants with a mutation in the *atvq15* gene (a *TaVQ14* homolog) than in WT controls, suggesting that *TaVQ14* increases resistance to salt and drought stress in Arabidopsis seeds. Moreover, under salt and drought stress, Arabidopsis lines overexpressing *TaVQ14* had higher catalase, superoxide dismutase, and proline levels and lower malondialdehyde concentrations than WT controls, suggesting that *TaVQ14* improves salt and drought resistance in Arabidopsis by scavenging reactive oxygen species. Expression analysis showed that several genes responsive to salt and drought stress were upregulated in Arabidopsis plants overexpressing *TaVQ14*. Particularly, salt treatment increased the expression of *AtCDPK2* in these plants. Moreover, salt treatment increased Ca^{2+} concentrations in plants overexpressing *TaVQ14*, suggesting that *TaVQ14* enhances salt resistance in Arabidopsis seeds through calcium signaling. In summary, this study demonstrated that the heterologous expression of *TaVQ14* increases the resistance of Arabidopsis seeds to salt and drought stress.

Keywords: wheat, salt, drought, subcellular localization, expression analysis, *TaVQ14*

Abbreviations: VQ, valine-glutamine; ABA, abscisic acid; qRT-PCR, quantitative real-time PCR; RWC, plant water content; MDA, malondialdehyde; CAT, catalase; SOD, superoxide dismutase; PRO, proline; GUS, β-glucuronidase activity; APX, ascorbate peroxidase; AsA, ascorbic acid; GR, glutathion reductases; GSH, l-glutathione; GSSG, L-glutathione oxidized.

INTRODUCTION

Wheat (*Triticum aestivum* L.) is a widely cultivated grain crop and one of the three most important cereal grains worldwide (Brenchley et al., 2012). Wheat is a good energy source and provides essential nutrients such as protein, vitamins, and trace elements. Therefore, the safe production of wheat is crucial to ensure food security globally.

Droughts, soil salinization, and other environmental stresses are the main factors limiting the increase in wheat yield and quality. Resistance to drought and salinity is a complex biological trait involving multiple molecular, physiological, biochemical, and morphological changes (Zhou et al., 2018; Wang and Huang, 2019; Zelm et al., 2020). Therefore, identifying key genes and elucidating the mechanisms regulating crop resistance to drought and salinity stress are useful for increasing food production while allowing the sustainable management of ecological resources (Ingram and Bartels, 1996; Xiong et al., 2002; Jakab et al., 2005; Verslues et al., 2006; Yamaguchi-Shinozaki and Shinozaki, 2006).

Valine-glutamine-motif proteins have attracted increasing attention because of their interaction with WRKY transcription factors (Lai et al., 2011). A total of 34, 40, and 61 VQ proteins have been identified in Arabidopsis, rice, and maize, respectively (Cheng et al., 2012; Li N. et al., 2014; Wang et al., 2014). These proteins regulate plant growth and development and the response to biotic and abiotic stresses (Andreasson et al., 2005; Petersen et al., 2010; Wang et al., 2010; Fill and Petersen, 2011; Hu P. et al., 2013; Hu Y. et al., 2013; Kim et al., 2013; Li Y. et al., 2014; Wang et al., 2014; Pascal et al., 2014; Wang H. et al., 2015; Wang M. et al., 2015; Song et al., 2016; Li et al., 2020). For instance, *AtVQ14 (IKU1)* is strongly expressed in the embryo and endosperm and controls endosperm development and seed size. Given that the nutrients required for seed germination are affected by seed size, thus it also affects resistance to adverse environments (Wang et al., 2010). *AtVQ29* is involved in the photomorphogenesis of *Arabidopsis* seedlings. The hypocotyls of plants overexpressing *AtVQ29* are longer than those of wild-type (WT) plants under far-red light or low light and regulate flowering time (Li Y. et al., 2014). *AtVQ21 (MSK1)* transgenic plants positively regulate the resistance of the pathogen *Pseudomonas syringae* and negatively mediate the resistance of the pathogen *Botrytis cinerea* (Petersen et al., 2010; Fill and Petersen, 2011). *AtVQ22* improves JA-mediated disease resistance, mutant plants overexpressing this gene are more resistant to necrotizing pathogens, and transgenic lines were extremely sensitive to pathogen infection. In addition, the analysis of the rice transcriptome showed that *VQ22* expression increased after infection with *Magnaporthe grisea*, indicating that this gene plays an important role in disease resistance (Hu Y. et al., 2013). *AtCaMBP25* (*AtVQ15*) reduced osmotic stress during seed germination and growth in *Arabidopsis thaliana*. Under salt and osmotic stress, transgenic lines are highly sensitive to seed germination, growth, and development (Wang M. et al., 2015). Seed germination and seedling growth were inhibited in plants overexpressing *AtVQ9* under salt stress (Hu Y. et al., 2013). *ZmVQ54*, *ZmVQ19*, *OsVQ2*, *OsVQ16*, and *OsVQ20* were highly expressed under drought stress (Kim et al., 2013). *GmVQ6* and *GmVQ53* were highly expressed in roots and stems under low-nitrogen conditions. *Arabidopsis* lines overexpressing (Wang et al., 2014). *PeVQ28* were salt tolerant (Cheng et al., 2020). These data demonstrate that *VQ* genes have multiple roles in regulating plant growth and development and resistance to biotic and abiotic stresses.

Little is known about the functions of *VQ* genes in wheat. Our previous study showed that *TaVQ14* was related to salt stress response (Cheng et al., 2021). Therefore, *TaVQ14* was selected as the target gene for further functional analysis. First, the subcellular location of TaVQ14 was determined. Second, *Arabidopsis* plants overexpressing *TaVQ14* were obtained by genetic transformation, and molecular, physiological, and phenotypic analyses were carried out. This study elucidated the functions of VQ genes and provided useful information for genetically improving salt and drought resistance in wheat crops.

MATERIALS AND METHODS

Experimental Materials and Stress Treatment

Wheat varieties Jing 411 (J411) and Hongmangchun 21 (HMC21) were provided by Shihe Xiao from the Chinese Academy of Agricultural Sciences, both of which were moderately salt and drought tolerant (Ren, 2012). Wheat seeds were treated with 300 mM NaCl or 300 mM mannitol. Samples were collected at 0, 4, 6, 10, 48, and 72 h, frozen in liquid nitrogen and stored immediately at −80°C.

Tobacco (*Nicotiana tabacum*) and Colombian ecotype *Arabidopsis* [Columbia-0 (Col-0), wild-type (WT)] seeds were provided by the State Key Laboratory of Crop Resistance of Anhui Agricultural University and were expanded and preserved in our laboratory. The seeds of WT and mutant *Arabidopsis* lines were cultivated in Murashige and Skoog (MS) medium at 24°C under a 16-h light/8-h dark cycle and transferred to square pots (diameter 6 cm) containing black soil and verstone (1:3, v/v) (Chen D. et al., 2017; Gao et al., 2017). Tobacco seeds were planted in square pots (6 cm in diameter) containing black soil and verstone (1:2, v/v).

Bioinformatics Analysis of *TaVQ14*

TaVQ14 sequences were obtained from the Ensembl Plants database. The isoelectric point (pI), molecular weight (MW), and other properties of TaVQ14 protein were analyzed using ExPASy (Cheng et al., 2018). The number of exons and introns of *TaVQ14* was analyzed using Gene Structure Display Server[1] (Zhu et al., 2018). The phylogenetic analysis of *TaVQ14* was performed using MEGA 7.0 (Cheng et al., 2019a). The promoter region of *TaVQ14* was analyzed using PlantCARE (Toufighi et al., 2005; Cheng et al., 2019b).

[1]http://gsds.gao-lab.org/

Total RNA Extraction and Real-Time PCR Analysis

Total RNA was extracted from wheat grains using RNzol Universal Total RNA Extraction Reagent (Tiangen, Beijing, China). Primers were designed using Primer Premier 6, and TaActin was used as a reference gene (**Supplementary Table 1**; Ji et al., 2011). RT-PCR analysis was performed using the TransStart Tip Green qPCR SuperMix kit (Transgen, Beijing, China). Each treatment included three biological replicates and three technical replicates. Data were transferred to Excel spreadsheets and analyzed using GraphPad Prism version 6.0 (Livak and Schmittgen, 2001; Zhao et al., 2012; Wang et al., 2017).

Cloning and Expression of *TaVQ14*

Specific primers were designed to clone the coding sequence of *TaVQ14* (**Supplementary Table 1**).

The vector pCAMBIA1305 (*p1305*) containing a GPF reporter gene was used for subcellular localization. Primers containing *Xba*I and *Bam*HI restriction sites were designed, and the stop codon was removed to construct the *p1305-CaMV35S-TaVQ14-GFP* expression vector, which served as a control (Dai et al., 2007).

The vector pCAMBIA1301a (*p1301a*) was used in overexpression experiments. Primers containing *Bam*HI and *Xba*I restriction sites were designed, and the *p1301a-TaVQ14* expression vector was constructed.

Subcellular Localization of TaVQ14

The subcellular localization of TaVQ14 was predicted using CELLO version 2.5 (Yu et al., 2004, 2006). The *Agrobacterium* suspension containing the *p1301a* vector was injected into tobacco leaves, and tissue sections were observed under a confocal microscope.

Analysis of *TaVQ14* Overexpression in *Arabidopsis*

Arabidopsis ecotype Col-0 (WT) was transformed with the p1301-TaVQ14 vector using the floral dip method to obtain mutant T0 seeds. Transformed seeds were screened on MS medium supplemented with hygromycin. SYBR Green I fluorescence quantitative PCR was used to detect the copy number of the exogenously introduced TaVQ14 gene relative to the dxr gene, the dxr gene encoding terpenoid synthase is a single copy in the Arabidopsis genome, setting the internal reference gene (**Supplementary Table 1**). Positive plants were obtained by cloning GUS gene fragments (**Supplementary Table 1**) and GUS staining and propagated to T3 overexpressed *Arabidopsis* plants. Seeds from WT and transgenic plants were grown on MS medium supplemented with mannitol (0, 150, or 300 mM) or NaCl (0, 100, or 150 mM) under the same experimental conditions, and the rates of germination were calculated. The calculation formula is: germination rate (GP) = number of normal germinated seeds at the end of germination/number of tested seeds × 100%. The effect of *TaVQ14* overexpression on stress resistance was assessed.

Analysis of TaVQ14 Homologs

The *atvq15* mutant (SALK_005722), with a mutation in position 17114641 of *AtVQ15* (*AT2G41010*, chr2: 17113798-17115047), was obtained from the Arashare platform. Mutants were screened by three-primer PCR to obtain homozygous plants. Seeds from WT and transgenic plants were grown on MS medium supplemented with mannitol (0, 150, or 300 mM) or NaCl (0, 100, or 150 mM) under the same experimental conditions, and the rates of germination were calculated.

Measurement of Stress-Related Physiological Indexes and Gene Expression Analysis

Seedlings of WT and transgenic lines were grown in a greenhouse for 24 h in plates containing MS medium supplemented with 300 mM mannitol or 150 mM NaCl, and plates containing MS medium were used as controls. Relative water content (RWC) (Bates et al., 1973; Zhao et al., 2014), and the levels of catalase (CAT) (Rorth and Jensen, 1967; Goldstein, 1968; Del Rio et al., 1977), proline (PRO) (Bates et al., 1973), malondialdehyde (MDA) (Heath and Packer, 1968), superoxide dismutase (SOD) (Sequeira and Mineo, 1966; Kochba et al., 1977), Ascorbate peroxidase (APX) (Zhao et al., 2018), Ascorbic acid (AsA) (Ji et al., 2019), Glutathion reductases (GR) (Zhang et al., 2018), L-Glutathione (GSH) (Xin et al., 2019), and L-Glutathione oxidized (GSSG) (Zhang et al., 2019) were determined, and the expression levels of genes induced by drought stress (*AtRD29A*, *AtRD29B*, *AtP5CS1*, *AtOST1, AtDi19-3, and AtWRKY46*) and salt stress (*AtSHM1*, *AtSOS2*, *AtCDPK2, AtDi19-3, and AtPP2C49*) were quantified (Huang et al., 2012; Qin et al., 2014; Chen J. et al., 2017; Chu et al., 2021).

Seedlings of WT and transgenic lines were grown in square pots for 20 days and treated with 150 mM NaCl or 300 mM mannitol every 3 days. The concentrations of Na^+, K^+, and Ca^{2+} were measured on treatment day 12.

Statistical Analysis

Statistical analysis was performed using SPSS version 19.0 and Origin software.

RESULTS

Bioinformatics Analysis of *TaVQ14*

TaVQ14 has 789 base pairs and encodes a 262-amino acid protein. The protein has a pI of 9.164, MW of 27481.07 Da, hydrophobicity index of 64.39, hydrophilicity index of -0.234, and instability index (II) of 74.16. The percentage of alanine, serine, and proline was 17.9, 14.5, and 11.8%, respectively. In addition, the most common *cis*-acting elements in the promoter region were associated with response to light (64.32%), abscisic acid (21.4%), and auxin (7.14%), and with zein metabolism regulation (7.14%) (**Supplementary Table 2**).

Fourteen homologous genes were obtained from the Ensembl Plants database, and a phylogenetic tree was constructed. The

results showed that *TaVQ14* was highly homologous with sequences from *Aegilops tauschii, Triticum turgidum, Hordeum vulgare*, and *Brachypodium distachyon* (**Supplementary Table 3** and **Figure 1**).

Expression Pattern Analysis of *TaVQ14*

Valine-glutamine genes are involved in abiotic stress response in plants (Hu Y. et al., 2013; Kim et al., 2013; Wang M. et al., 2015; Cheng et al., 2020). Our previous study showed that *TaVQ14* was associated with salt stress response (Cheng et al., 2021). Based on these findings, wheat seeds were treated with NaCl or mannitol to assess the role of *TaVQ14* in salinity and drought stress. The results showed that salt and mannitol treatment increased the expression of *TaVQ14* in a time-dependent manner (**Figure 2**). Furthermore, the expression levels of *TaVQ14* after treatment were similar between the two wheat varieties. These results indicated that *TaVQ14* expression was induced by NaCl and mannitol, suggesting that *TaVQ14* might play essential roles in salinity and drought stress.

Subcellular Localization Analysis of *TaVQ14*

TaVQ14 was predicted to be found in the nucleus. To confirm this prediction, tobacco leaves were transformed with *Agrobacterium tumefaciens* harboring the *p1305-CaMV35S-TaVQ14-GFP* fusion protein and observed on a confocal microscope. The results showed that the protein signal was detected in the cell nucleus, whereas the GFP control vector was distributed throughout the cell (**Figure 3**).

Obtaining of *Arabidopsis* Plants Overexpressing *TaVQ14*

Twenty T1 lines overexpressing *TaVQ14* were obtained. GUS activity results showed that these lines were genetically transformed (stained blue), whereas WT lines (negative control) were not transformed and stained yellow (**Supplementary Figure 1a**). DNA was extracted from transgenic (T1) and WT plants. The results showed that the transgenic line and positive control, but not the WT line, presented a 650-bp band (**Supplementary Figure 1b**). We randomly selected five lines from the *TaVQ14* transgenic Arabidopsis positive seedlings to detect the expression level of the target gene *TaVQ14*. The results showed that *TaVQ14* was not expressed in WT, and highly expressed in four lines (**Supplementary Figure 1c**). In WT, and Lines 1/-2/-3, the copy number of *TaVQ14* gene was detected. The results showed that the *Ct* value of *TaVQ14* gene in WT was greater than 40, indicated that there was no *TaVQ14* gene in this material. The *TaVQ14* gene of Lines 1/-2/-3 were all

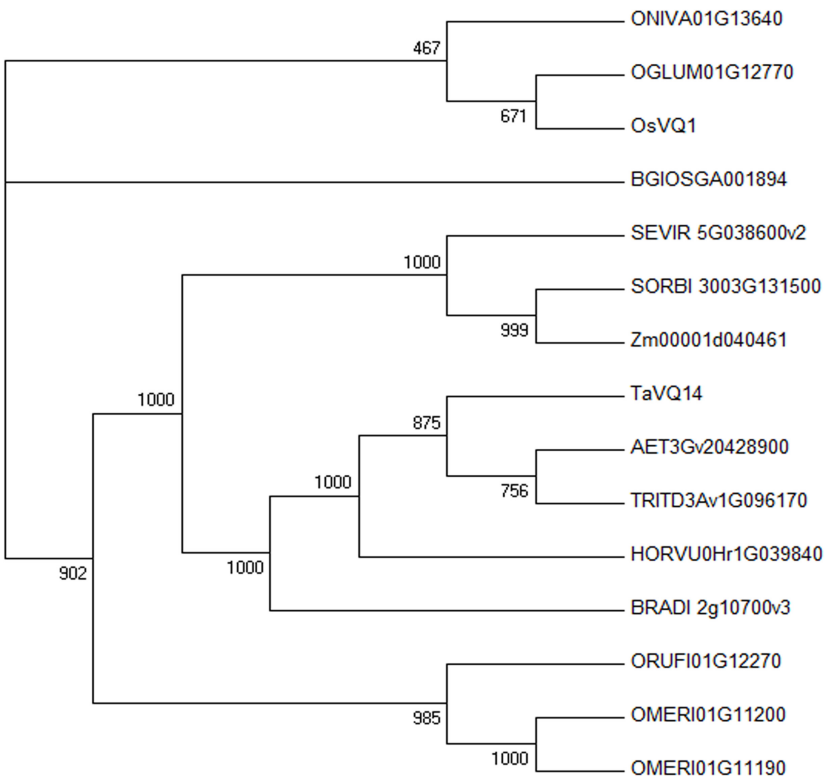


FIGURE 1 | Phylogenetic analysis of *TaVQ14* and its homologs. Homologs*: Aegilops tauschii, AET3Gv20428900; Brachypodium distachyon, BRADI_2g10700v3; Hordeum vulgare, HORVU0Hr1G039840; Oryza glumipatula, OGLUM01G12770; Oryza sativa Indica Group, BGIOSGA001894; Oryza meridionalis, OMERI01G11190; Oryza meridionalis, OMERI01G11200; Oryza nivara, ONIVA01G13640; Oryza rufipogon, ORUFI01G12270; Oryza sativa Japonica Group, OsVQ1 (Os01g0278000); Setaria viridis, SEVIR_5G038600v2; Sorghum bicolor, SORBI_3003G131500; Triticum turgidum, TRITD3Av1G096170; Zea mays, Zm00001d040461.*

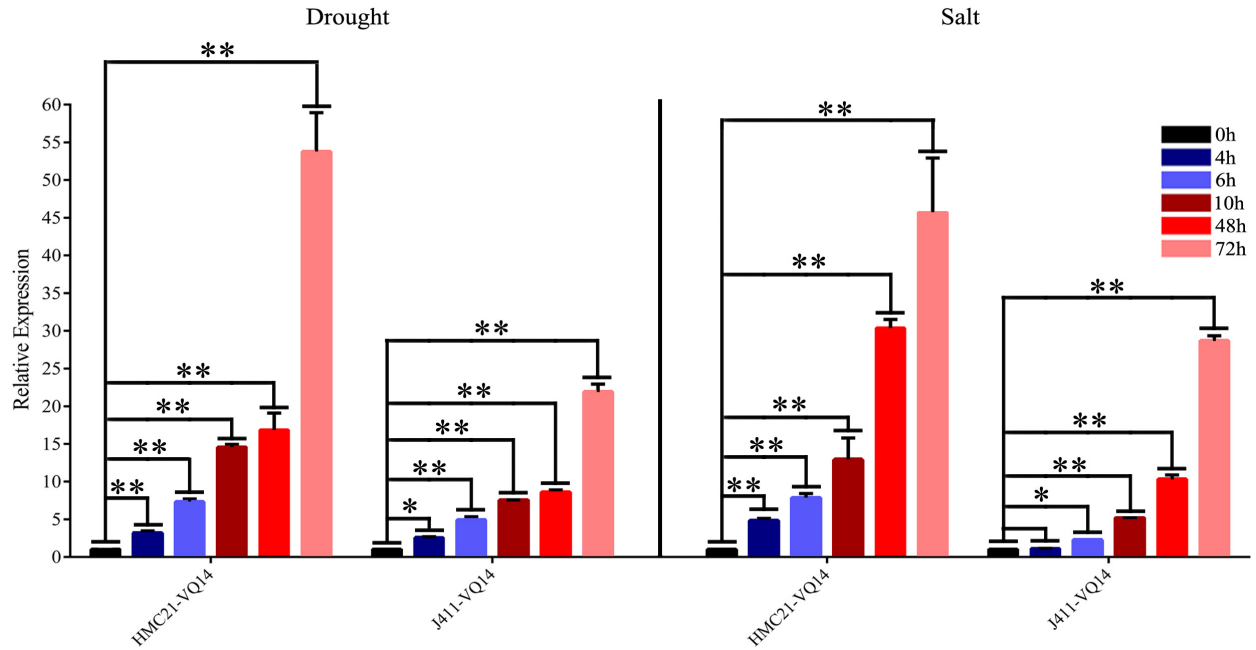


FIGURE 2 | Analysis of *TaVQ14* expression by reverse-transcription quantitative PCR in wheat seeds subjected to drought and NaCl stress. $^{*}P < 0.05$, $^{**}P < 0.01$.

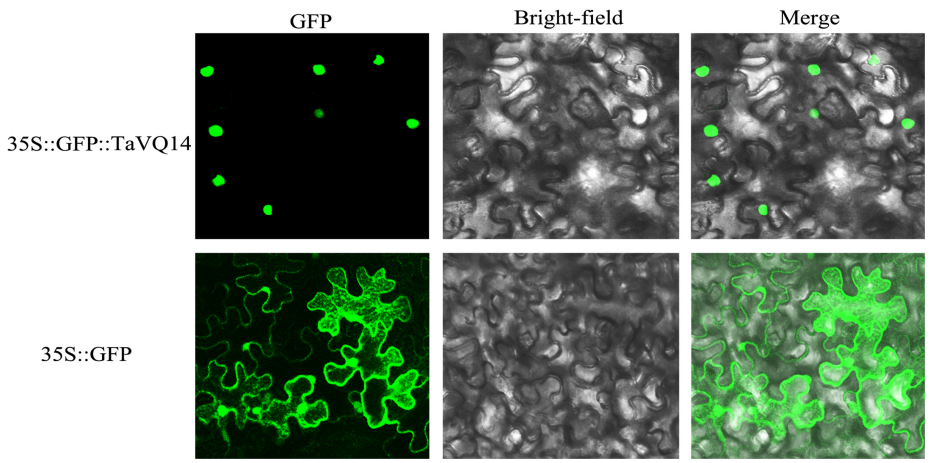


FIGURE 3 | Nuclear localization of TaVQ14. The construct 35S:GFP:TaVQ14 and the control vector 1305 (35S:GFP) were transformed into *Nicotiana tabacum* leaves. The GFP signals in root cells were observed by confocal microscopy.

single copy material (**Supplementary Table 4**). Transgenic and WT plants were self-crossed for three generations, and their seeds were harvested.

Resistance to Salt Stress

Seedlings were treated with NaCl (0, 100, or 150 mM), and the rates of germination were calculated. There was no significant difference in germination rate between transgenic and WT plants (100 vs. 99%) before treatment. Treatment with 100 and 150 mM NaCl decreased the rate of germination in WT and transgenic plants; nonetheless, germination was lower in the former (78 and 34% vs. 90 and 51%) (**Figures 4A,B**).

We measured several physiological indexes in seedlings treated with 150 mM NaCl. Treatment decreased the RWC of transgenic and WT plants; however, the rate of decline was significantly higher in WT plants. Treatment increased the concentrations of CAT, PRO, SOD, APX, GR, GSH, and GSSG in both lines. The increase was more significant in the transgenic line. The increase in MDA levels was less pronounced. The increase in GR, GSH, and GSSG was no obvious changed. Treatment reduced the concentrations of AsA in both lines, the reduce was more significant in WT (**Figures 4C–L**).

These results showed that transgenic plants had stronger salt tolerance than controls, indicating that the overexpression of

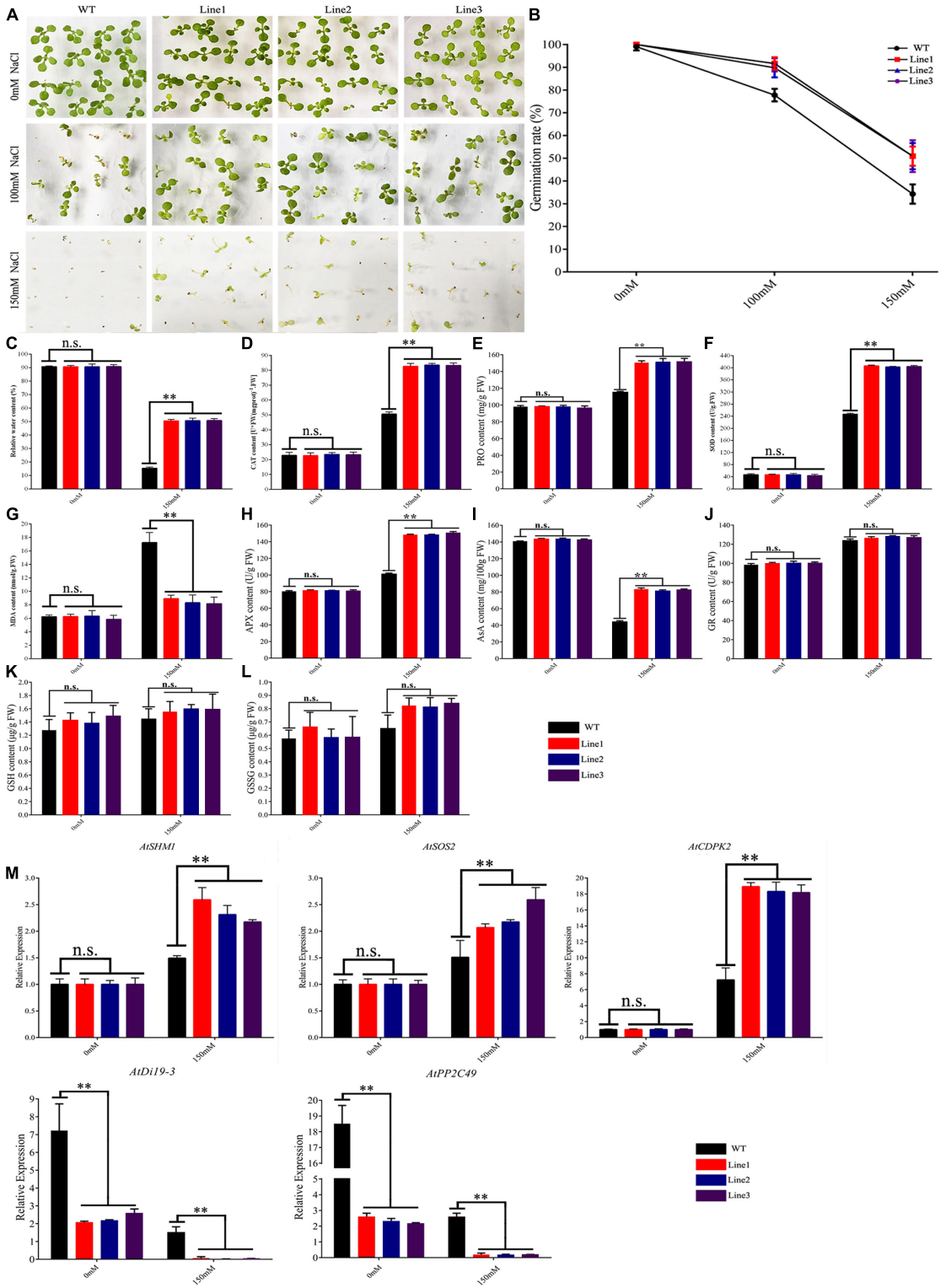


FIGURE 4 | Germination phenotypes of *TaVQ14* in transgenic *Arabidopsis* plants under salt stress. **(A)** Germination performance in *TaVQ14*-overexpressing and wild-type (WT) seeds grown on Murashige and Skoog medium containing 0, 100, or 150 mM of NaCl. **(B)** Rate of germination in transgenic and WT seeds. **(C–L)** Relative water content **(C)** and levels of catalase **(D)**, proline **(E)**, superoxide dismutase **(F)**, malondialdehyde **(G)**, Ascorbate Peroxidase **(H)**, Ascorbic acid **(I)**, Glutathion Reductases **(J)**, L-Glutathione **(K)**, and L-Glutathione oxidized **(L)** in transgenic and WT plants after salt treatment. Values are means ± SE (n = 3). $^{*}P < 0.05$ $^{**}P < 0.01$ (t-test). **(M)** Relative expression levels of salt-responsive genes in transgenic *A. thaliana* plants under normal salinity. Leaves of transgenic and WT plants were collected after salt stress. Y-axis: relative expression levels; X-axis: the time course of stress treatments; Error bars, 6 ± SE.

TaVQ14 improved salt tolerance. To further assess the effect of *TaVQ14* on salt tolerance in *Arabidopsis*, we analyzed the expression of salt stress genes (*AtSHM1*, *AtSOS2*, *AtCDPK2, AtDi19-3, and AtPP2C49*) (Huang et al., 2012; Qin et al., 2014; Chu et al., 2021). Salt treatment increased the expressions of *AtSHM1*, *AtSOS2* and *AtCDPK2,* decreased the expression of *AtDi19-3 and AtPP2C49* (**Figure 4M**). In addition, the relative expression of *AtCDPK2* increased in both groups, suggesting that, as a crucial Ca^{2+} sensor, *AtCDPK2* enhances salt tolerance in *Arabidopsis* seeds.

Resistance to Drought Stress

Seedlings were treated with mannitol (0, 150, or 300 mM), and the rates of germination were measured. There was no significant difference in the percentage of germination between transgenic and WT plants (100 vs. 99%). Treatment with 150 and 300 mM mannitol decreased the rate of germination in both groups; however, germination was less affected in transgenic plants (95 and 85% vs. 86 and 47%) (**Figures 5A,B**).

Mannitol treatment decreased RWC in both groups, the rate of decline was higher in the WT group. Treatment increased the levels of CAT, SOD, PRO, MDA, APX, GR, GSH, and GSSG. The rate of increase in catalase and SOD was higher in transgenic plants, whereas the rate of increase in MDA was higher in WT plants. The increase in GR, GSH, and GSSG was no obvious changed. Treatment reduced the concentrations of AsA in both lines, the reduce was more significant in WT (**Figures 5C–L**).

These results showed that transgenic plants had stronger drought tolerance than WT plants, indicating that the overexpression of *TaVQ14* increased drought resistance in *Arabidopsis* seeds. To further evaluate the effect of *TaVQ14* on drought tolerance, the expression levels of drought-related genes (*AtRD29A*, *AtRD29B*, *AtP5CS1, AtOST1, AtDi19-3, and AtWRKY46*) were quantified (Huang et al., 2012; Qin et al., 2014; Chen J. et al., 2017). Drought treatment increased the expression of *AtRD29A, AtRD29B, AtP5CS1,* and *AtOST1*, decreased the expression of *AtDi19-3* and *AtWRKY46* (**Figure 5M**); suggesting that TaVQ14 improves drought tolerance by enhancing the expression of these genes.

Measurement of Ion Concentrations

The concentration of K^+ and Ca^{2+} was significantly higher in *TaVQ14*-overexpressing lines, whereas Na^+ content was similar between the groups (**Figure 6A**). Treatment with 150 mM NaCl increased Na^+ contents in both groups; however, the increase was more pronounced in the WT group (**Figure 6B**). Treatment with 300 mM mannitol increased K^+ and Ca^{2+} concentrations in both groups. However, the increase was higher in the transgenic group (**Figure 6C**).

Response of *atvq15* Mutants to Salt and Drought Stresses

Homologous genes have similar functions across species (Wang et al., 2017). To further investigate the role of *TaVQ14* homologs in regulating salt and drought resistance, a phylogenetic analysis of *TaVQ14* and VQ family members was performed (**Supplementary Figure 2a**). The *atvq15* sequence was obtained from the Arashare platform, and detected homozygous plants of *atvq15* mutant by screening leaf DNA (**Supplementary Figure 2b**). Seeds of *atvq15* mutant and WT plants were treated with mannitol (0, 150, or 300 mM) or NaCl (0, 100, or 150 mM), and the rates of germination were calculated. There was no significant difference in the rate of germination rate between *atvq15* mutant and WT plants (94% in both groups) before treatment (**Figure 7**). Treatment with 100 and 150 mM NaCl decreased the germination rate in both groups, but the rates were lower in *atvq15* mutants (20 and 0% vs. 64 and 14%) (**Figures 7A,B**). Treatment with 150 and 300 mM Mannitol decreased the germination rate in both groups; nonetheless, the effect was stronger in *atvq15* mutants (rates of 45 and 5% vs. 78 and 16%) (**Figures 7C,D**). These results support that *AtVQ15* and *TaVQ14* regulate salt and drought resistance in *Arabidopsis* seeds.

DISCUSSION

Valine-glutamine proteins are widely found in *Arabidopsis*, rice, maize, soybean, grapes, and other plant species (Cheng et al., 2012; Kim et al., 2013; Li N. et al., 2014; Wang et al., 2014; Wang M. et al., 2015). However, few studies have assessed the functions of these proteins. Wheat is one of the most widely cultivated crops; nonetheless, the functions of *VQ* genes in wheat are incompletely understood. Our previous study has shown that *TaVQ14* encodes an unstable basic hydrophobic protein (Cheng et al., 2021). Therefore, this gene was selected for further functional analysis. Our results showed that *TaVQ14* expression was significantly upregulated under high salinity and drought conditions, indicating that *TaVQ14* was involved in salt and drought stress responses.

Arabidopsis thaliana is an excellent model for research in plant biology is an excellent model for research in plant biology, which can obtain transgenic plants in short time. Therefore, transgenic *A. thaliana* lines were used to assess the function of *TaVQ14*. Under drought and salt stress, the rate of germination in *TaVQ14*-overexpressing lines was significantly higher than that of WT plants, indicating that the tolerance of the former to drought and salinity stress was improved.

A plant mutant for *AtVQ15*, a *TaVQ14* homolog, was produced. The results revealed that the percentage of seed germination was lower in these mutants than in WT controls under drought and salt stress. These findings indicate that the *AtVQ15* mutation reduces stress tolerance and that *TaVQ14* and its homolog *AtVQ15* regulate tolerance to drought and salinity. Furthermore, gene expression analysis showed that several genes responsive to drought (*AtRD29A, AtRD29B, AtP5CS1, AtOST1, AtWRKY46,* and *AtDi19-3*) and salinity (*AtSHM1, AtSOS2, AtCDPK2, AtPP2C29,* and *AtDi19-3*) (Huang et al., 2012) were differentially expressed after treatment, suggesting that *TaVQ14* enhances the resistance of *Arabidopsis* seeds to drought and salinity by regulating the expressions of these genes.

The concentrations of K^+, Na^+, and Ca^{2+} were measured in *TaVQ14*-overexpressing and WT lines. Ca^{2+} concentration

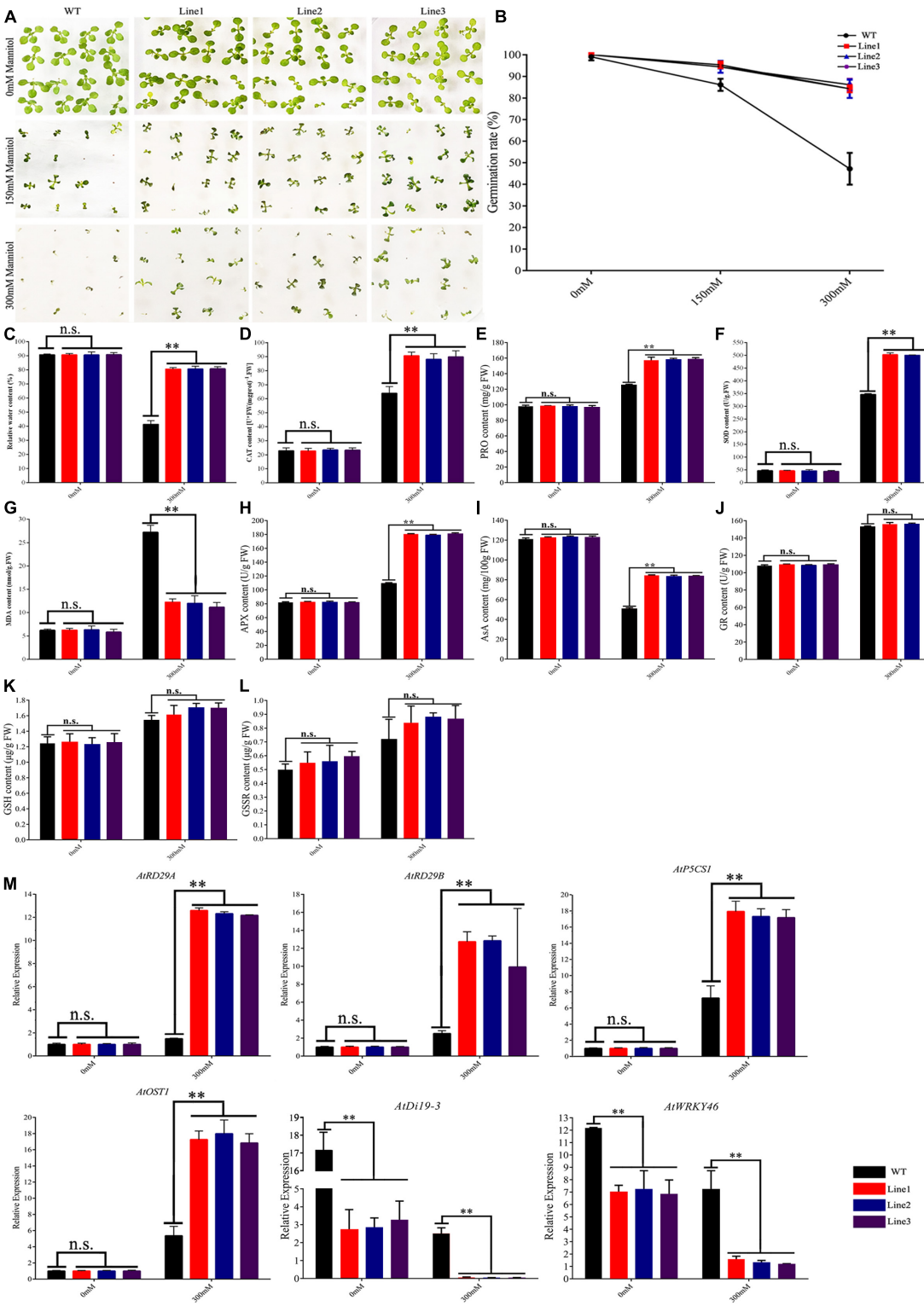


FIGURE 5 | Germination phenotypes of *TaVQ14* in transgenic *Arabidopsis* plants under drought stress. **(A)** Germination performance of *TaVQ14*-overexpressing and wild-type (WT) seeds on Murashige and Skoog medium containing 0, 150, or 300 mM of mannitol. **(B)** Rate of germination in transgenic and WT seeds. **(C–L)** Relative water content **(C)**, and levels of catalase **(D)**, proline **(E)**, superoxide dismutase **(F)**, malondialdehyde **(G)**, Ascorbate Peroxidase **(H)**, Ascorbic acid **(I)**, Glutathion Reductases **(J)**, L-Glutathione **(K)**, and L-Glutathione oxidized **(L)** in transgenic and WT plants after drought treatment. Values are means ± SE (n = 3). $^{*}P < 0.05$ $^{**}P < 0.01$ (t-test). **(M)** Relative expression levels of drought-responsive genes in transgenic *A. thaliana* plants under normal water conditions. Leaves of transgenic and WT plants were collected after drought stress. Y-axis: relative expression levels; X-axis: the time course of stress treatments; Error bars, 6 ± SE.

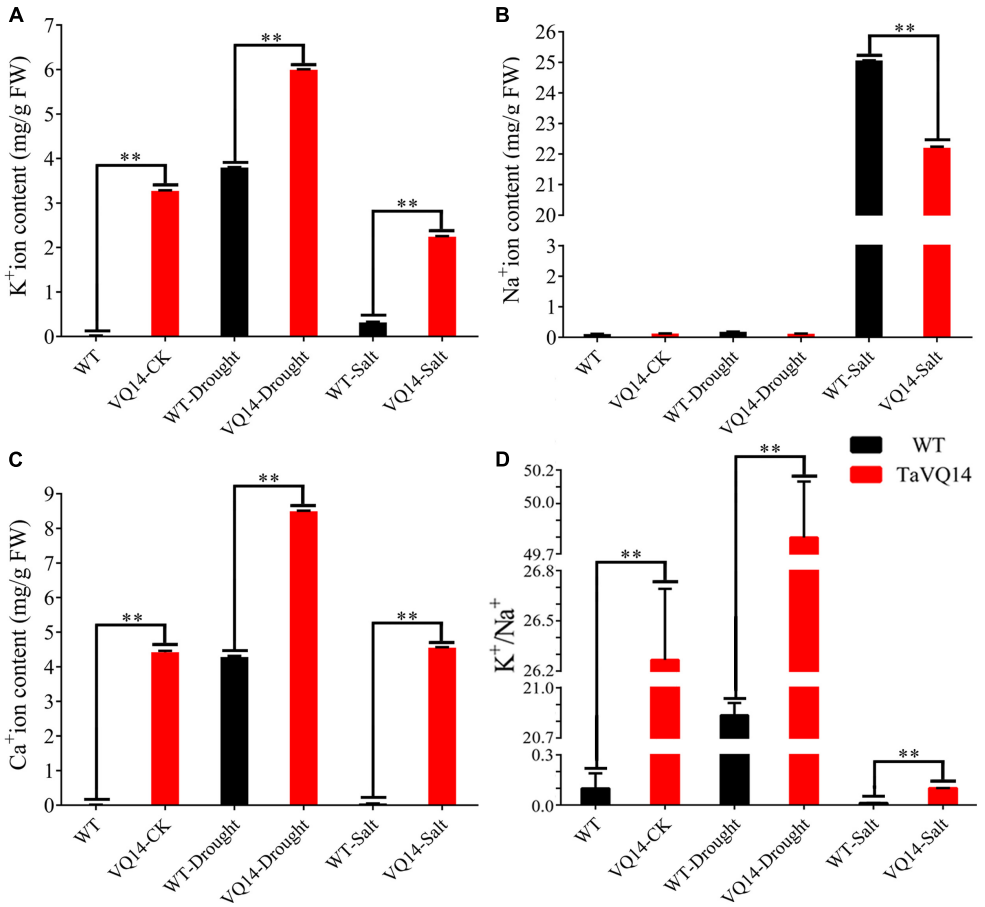


FIGURE 6 | Concentration of Na^+, K^+, and Ca^{2+} in transgenic *Arabidopsis thaliana* plants under salt and drought stress. **(A)** K^+ ion content in transgenic *Arabidopsis thaliana* plants under salt and drought stress. **(B)** Na^+ ion content in transgenic *Arabidopsis thaliana* plants under salt and drought stress. **(C)** Ca^{2+} ion content in transgenic *Arabidopsis thaliana* plants under salt and drought stress. **(D)** K^+/Na^+ in transgenic *Arabidopsis thaliana* plants under salt and drought stress. $^*P < 0.05$, $^{**}P < 0.01$.

and K^+/Na^+ ratio was significantly higher in the transgenic line before treatment, suggesting that *TaVQ14* overexpression improves drought and salt resistance by increasing Ca^{2+} and K^+ concentrations. Treatment with 300 mM NaCl increased Ca^{2+}, K^+, and K^+/Na^+ ratio in the transgenic line, suggesting that *TaVQ14* increases resistance to salt stress by excreting Na^+ and increasing the uptake of Ca^{2+} and K^+ (**Figures 6A–D**). Treatment with 300 mM mannitol increased the uptake of K^+ and Ca^{2+} in both groups; nonetheless, the increase was more pronounced in the transgenic line, suggesting that *TaVQ14* overexpression improves drought resistance by increasing K^+ and Ca^{2+} concentrations. *AtCDPK2* was upregulated in the transgenic line under salt stress, suggesting that *TaVQ14* improves salt tolerance by increasing *AtCDPK2* expression through Ca^{2+} signaling. Salinity stress increases cytosolic Ca^{2+} levels. Calcium-dependent protein kinases (CPKs or CDPKs) are strongly implicated in Ca^{2+} signaling in plants and play an important role in salinity stress (Singh et al., 2017). In rice, *OsCPK21* genes regulated the ABA-dependent salt stress signaling pathway (Asano et al., 2011). *OsCPK12* conferred tolerance to salt stress through regulation of ROS homeostasis (Asano et al., 2012). In addition, salinity stress tolerance is stronger in plants overexpressing *OsCPK4* (Campo et al., 2014). These findings reveal that Ca^{2+} signaling, together with ROS signaling and hormonal regulation, mediates the response to salinity stress. However, the mechanisms underlying the regulation of salt tolerance by *TaVQ14* and CDPK need to be further investigated.

Stress resistance is improved by modulating gene expression and physiological and biochemical processes, including the accumulation of osmotic substances and the increase in active oxygen scavenging activity (Zhou et al., 2018; Zhang et al., 2021). Free proline levels in plants are low under normal conditions; nonetheless, under stress conditions, including drought, low temperature, high salinity, and high alkalinity, proline is stored in large quantities, and storage levels are positively correlated with stress resistance. Therefore, proline is used as a biochemical index of stress resistance in plants (Xiang et al., 2007; Hnilickova et al., 2021; Qian et al., 2021; Rajametov et al., 2021). Our results showed that proline level was significantly higher in transgenic

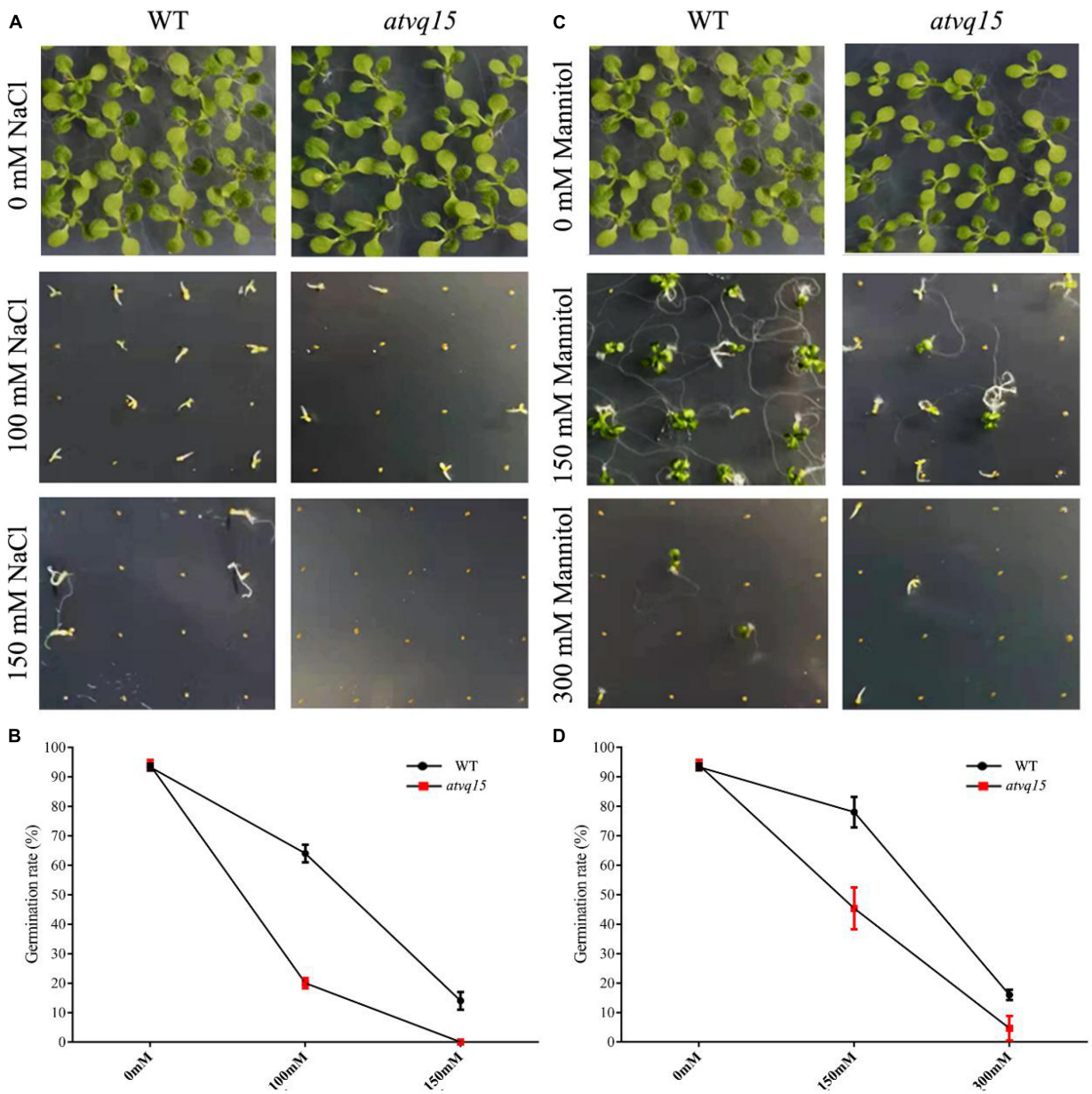


FIGURE 7 | Germination phenotypes of *Arabidopsis thaliana* seeds with a mutation in the gene *atvq15*, a *TaVQ14* homolog, under salt and drought stress. **(A)** Germination performance in *atvq15* mutant and wild-type (WT) seeds grown on Murashige and Skoog (MS) medium containing 0, 100, or 150 mM of NaCl. **(B)** Rate of germination in *atvq15* mutant and WT seeds. **(C)** Germination performance in *atvq15* mutant and WT seeds grown on MS medium supplemented with 0, 150, or 300 mM of mannitol. **(D)** Rate of germination in *atvq15* mutant and WT seeds. Relative water content, and levels of catalase, proline, superoxide dismutase, malondialdehyde, Ascorbate Peroxidase, Ascorbic acid, Glutathion Reductases, L-Glutathione, and L-Glutathione oxidized in atvq14 and WT plants after salt treatment. Relative water content, and levels of catalase, proline, superoxide dismutase, malondialdehyde, Ascorbate Peroxidase, Ascorbic acid, Glutathion Reductases, L-Glutathione, and L-Glutathione oxidized in atvq14 and WT plants after drought treatment.

plants than in WT controls. Consistent with our results, the large increase in proline concentration increases intracellular osmotic pressure and decreases water potential and water content (Song et al., 2011). MDA is the main product of membrane lipid peroxidation, leading to membrane damage and plant damage. Thus, MDA levels are positively correlated with the degree of membrane lipid peroxidation and can serve as an indicator of cellular reactive oxygen species stress (Xiong et al., 2002; Mittler et al., 2004). Our results showed that MDA concentration was significantly lower in transgenic plants than in WT controls, indicating that *TaVQ14* overexpression increased resistance to oxidative stress. CAT, SOD, and APX are important protective enzymes and reduce oxidative stress by decreasing the production of active oxygen and hydrogen peroxide (May, 2008). In this

study, CAT, SOD, and APX levels were significantly higher in the transgenic line than in WT plants, which may explain why the MDA concentration was lower in the former.

The analysis of gene expression, physiological, biochemical, and phenotypic data demonstrated the role of *TaVQ14* overexpression in Arabidopsis improving its salt tolerance and drought tolerance, indicating that *TaVQ14* plays important roles in improving salt tolerance and drought resistance in wheat, and these data provide a basis for the functional analysis of *TaVQ14* in wheat.

China's Bohai Rim region has more than 40 million mu of medium and low yield farmland and more than 10 million mu of saline alkali wasteland, which has been suffering from drought, waterlogging and alkali disasters for a long time. This experiment proved that *TaVQ14* had the functions of salt tolerance and drought tolerance, and was an excellent salt tolerance and drought tolerance gene. Overexpression and knockout of *TaVQ14* gene in wheat is the direction of our subsequent experimental work, and we hope to widely apply this gene to the cultivation of new wheat varieties with salt tolerance and drought tolerance for the increase of agricultural production and income.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

XC and HY projected the study, put into effect the main bioinformatics analysis, and drew up the manuscript. ZC, CG, and WG carried out the software and helped to handle figures and tables. BT, JJC, and XP participated in the experimental test. SY processed experimental data and joined to amend the manuscript. JJC and XP took part in the software and draw up the manuscript. JL and CXM had a hand in the project of the study and helped to revamp the manuscript. CC and HPZ conceived and guided the experiment, were involved in its project and coordination, and helped to draw up the manuscript. All authors read and accepted the final manuscript.

FUNDING

This work was supported by grants from the China's Agricultural Research System (CARS-03), Jiangsu Collaborative Innovation Center for Modern Crop Production (JCIC-MCP), and The Agriculture Research System of Anhui province (AHCYTX-02).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.870586/full#supplementary-material

REFERENCES

Andreasson, E., Jenkins, T., Brodersen, P., Thorgrimsen, S., Petersen, N. H. T., Zhu, S., et al. (2005). The MAP kinase substrate *MKS1* is a regulator of plant defense responses. *EMBO J.* 24, 2579–2589. doi: 10.1038/sj.emboj.7600737

Asano, T., Hakata, M., Nakamura, H., Aoki, N., Komatsu, S., Ichikawa, H., et al. (2011). Functional characterisation of *OsCPK21*, a calcium-dependent protein kinase that confers salt tolerance in rice. *Plant Mol. Biol.* 75, 179–191. doi: 10.1007/s11103-010-9717-1

Asano, T., Hayashi, N., Kikuchi, S., and Ohsugi, R. (2012). CDPK-mediated abiotic stress signaling. *Plant Signal. Behav.* 7, 817–821. doi: 10.4161/psb.20351

Bates, L. S., Waldren, R. P., and Teare, I. D. (1973). Rapid determination of free proline for water-stress studies. *Plan Soil* 39, 205–207. doi: 10.1007/BF00018060

Brenchley, R., Spannagl, M., Pfeifer, M., Barker, G. L. A., Damore, R., Allen, A. M., et al. (2012). Analysis of the bread wheat genome using whole genome shotgun sequencing. *Nature* 491, 705–710. doi: 10.1038/nature11650

Campo, S., Baldrich, P., Messeguer, J., Lalanne, E., Coca, M., and San Segundo, B. (2014). Overexpression of a calcium-dependent protein kinase confers salt and drought tolerance in rice by preventing membrane lipid peroxidation. *Plant Physiol.* 165, 688–704. doi: 10.1104/pp.113.230268

Chen, D., Chen, Z., Wu, M., Wang, Y., Wang, Y., Yan, H., et al. (2017). Genome-Wide Identification and Expression Analysis of the *HD-Zip* Gene Family in Moso Bamboo (*Phyllostachys edulis*). *J. Plant Growth Regul.* 36, 323–337. doi: 10.1007/s00344-016-9642-x

Chen, J., Nolan, T., Ye, H., Zhang, M., Tong, H., Xin, P., et al. (2017). Arabidopsis WRKY46, WRKY54, and WRKY70 Transcription Factors Are Involved in Brassinosteroid-Regulated Plant Growth and Drought Responses. *Plant Cell* 29, 1425–1439. doi: 10.1105/tpc.17.00364

Cheng, X. R., Gao, C., Liu, X., Xu, D. M., Pan, X., Gao, W., et al. (2021). Identification of the Wheat VQ protein family and expression analysis of candidate genes associated with seed dormancy and germination. *BMC Plant Biol.* 22:119. doi: 10.1186/s12870-022-03430-1

Cheng, X., Wang, S., Xu, D. M., Liu, X., Li, X., Xiao, W., et al. (2019a). Identification and analysis of the *GASR* gene family in Common Wheat (*Triticum aestivum* L.) and characterization of *TaGASR34*, a gene associated with seed dormancy and germination. *Front. Genet.* 10:980. doi: 10.3389/fgene.2019.00980

Cheng, X., Xiong, R., Yan, H., Gao, Y., Liu, H., Wu, M., et al. (2019b). The trihelix family of transcription factors: functional and evolutionary analysis in Moso bamboo (*Phyllostachys edulis*). *BMC Plant Biol.* 19:154. doi: 10.1186/s12870-019-1744-8

Cheng, X. R., Wang, Y., Xiong, R., Gao, Y., Yan, H., and Xiang, Y. (2020). A Moso bamboo gene *VQ28* confers salt tolerance to transgenic Arabidopsis plants. *Planta* 251:99. doi: 10.1007/s00425-020-03391-5

Cheng, X. R., Xiong, R., Liu, H., Wu, M., Chen, F., Yan, H., et al. (2018). Basic helix-loop-helix gene family: genome wide identification, phylogeny, and expression in Moso bamboo. *Plant Physiol. Biochem.* 132, 104–119. doi: 10.1016/j.plaphy.2018.08.036

Cheng, Y., Zhou, Y., Yang, Y., Chi, Y. J., Zhou, J., Chen, J. Y., et al. (2012). Structural and functional analysis of VQ motif-containing proteins in Arabidopsis as interacting proteins of *WRKY* transcription factors. *Plant Physiol.* 159, 810–825. doi: 10.1104/pp.112.196816

Chu, M., Chen, P., Meng, S., Xu, P., and Lan, W. (2021). The Arabidopsis phosphatase PP2C49 negatively regulates salt tolerance through inhibition of AtHKT1;1. *J. Integr. Plant Biol.* 63, 528–542. doi: 10.1111/jipb.13008

Dai, X., Xu, Y., Ma, Q., Xu, W., Wang, T., Xue, Y., et al. (2007). Overexpression of an *R1R2R3 MYB* gene, *OsMYB3R-2*, increases tolerance to freezing, drought, and salt stress in transgenic Arabidopsis. *Plant Physiol.* 143, 1739–1751. doi: 10.1104/PP.106.094532

Del Rio, L. A., Ortega, M. G., Lopez, A. L., and Gorgo, J. L. (1977). A more sensitive modification of the catalase assay with the Clark oxygen electrode:

Application to the kinetic study of the pea leaf enzyme. *Anal. Biochem*. 80, 409–415. doi: 10.1016/0003-2697(77)90662-5

Fill, B. K., and Petersen, M. (2011). Constitutive expression of *MKS1*confers susceptibility to *Botrytis cinerea* infection independent of *PAD3* expression. *Plant Signal. Behav*. 6, 1425–1427. doi: 10.4161/psb.6.10.16759

Gao, Y., Liu, H., Wang, Y., Li, F., and Xiang, Y. (2017). Genome-wide identification of PHD-finger genes and expression pattern analysis under various treatments in moso bamboo (*Phyllostachys edulis*). *Plant Physiol. Biochem*. 123, 378–391. doi: 10.1016/j.plaphy.2017.12.034

Goldstein, D. B. (1968). A method for assay of catalase with the oxygen cathode. *Anal. Biochem*. 24, 431–437. doi: 10.1016/0003-2697(68)90148-6

Heath, R. L., and Packer, L. (1968). Photoperoxidation in isolated chloroplasts. I. Kinetics and stoichiometry of fatty acid peroxidation. *Arch. Biochem. Biophys.* 125, 189–198. doi: 10.1016/0003-9861(68)90654-1

Hnilickova, H., Kraus, K., Vachova, P., and Hnilicka, F. (2021). Salinity stress affects photosynthesis, malondialdehyde formation, and proline content in *Portulaca oleracea* L. *Plants* 10:845. doi: 10.3390/plants10050845

Hu, P., Zhou, W., Cheng, Z., Fan, M., Wang, L., and Xie, D. (2013). *JAV1* controls jasmonate-regulated plant defense. *Mol. Cell* 50, 504–515. doi: 10.1016/j.molcel.2013.04.027

Hu, Y., Chen, L., Wang, H., Zhang, L., Wang, F., and Yu, D. (2013). Arabidopsis transcription factor *WRKY8* functions antagonistically with its interacting partner *VQ9* to modulate salinity stress tolerance. *Plant J.* 74, 730–745. doi: 10.1111/tpj.12159

Huang, X., Zhang, Y., Jiao, B., Chen, G., Huang, S., Guo, F., et al. (2012). Overexpression of the wheat salt tolerance-related gene *TaSC* enhances salt tolerance in Arabidopsis. *J. Exp. Bot*. 63, 5463–5473. doi: 10.1093/jxb/ers198

Ingram, J., and Bartels, D. (1996). The molecular basis of dehydration tolerance in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 47, 377–403. doi: 10.1146/annurev.arplant.47.1.377

Jakab, G., Zimmerli, L., Métraux, J. P., and Mauch-Mani, B. (2005). Enhancing Arabidopsis salt and drought Stress tolerance by chemical priming for its abscisic acid responses. *Plant Physiol*. 139, 267–274. doi: 10.1104/pp.105.065698

Ji, X., Dong, B., Shiran, B., Shiran, B., Talbot, M. J., Edlington, J. E., et al. (2011). Control of abscisic acid catabolism and abscisic acid homeostasis is important for reproductive stage stress tolerance in cereals. *Plant Physiol.* 156, 647–662. doi: 10.1104/pp.111.176164

Ji, Y., Zhang, P., Xing, X., Jia, L., Zhang, Y., Jia, T., et al. (2019). Effect of 1α, 25-dihydroxyvitamin D3 on the osteogenic differentiation of human periodontal ligament stem cells and the underlying regulatory mechanism. *Int. J. Mol. Med*. 43, 167–176. doi: 10.3892/ijmm.2018.3947

Yamaguchi-Shinozaki, K., and Shinozaki, K. (2006). Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. *Annu Rev. Plant Biol.* 57, 781–803. doi: 10.1146/ANNUREV.ARPLANT.57.032905.105444

Kim, D. Y., Kwon, S. I., Choi, C., Lee, H., Ahn, I., Park, S. R., et al. (2013). Expression analysis of rice *VQ* genes inresponse to biotic and abiotic stresses. *Gene* 529, 208–214. doi: 10.1016/j.gene.2013.08.023

Kochba, J., Lavee, S., and Spiegelroy, P. (1977). Differences in peroxidase activity and isoenzymes in embryogenic and non-embryogenic'Shamouti'orange ovular callus lines. *Plant Cell Physiol*. 18, 463–467. doi: 10.1093/oxfordjournals.pcp.a075455

Lai, Z., Li, Y., Wang, F., Cheng, Y., Fan, B., Yu, J. Q., et al. (2011). Arabidopsis sigma factor binding proteins are activators of the *WRKY33* transcription factor in plant defense. *Plant Cell* 23, 3824–3841. doi: 10.1105/tpc.111.090571

Li, N., Li, X., Xiao, J., and Wang, S. (2014). Comprehensive analysis of VQ motif-containing gene expression in rice defense responses to three pathogens. *Plant Cell Rep*. 33, 1493–1505. doi: 10.1007/s00299-014-1633-4

Li, Y., Jing, Y., Li, J., Xu, G., and Lina, R. (2014). Arabidopsis VQ-motif-containing protein 29 represses seedling de-etiolation by interacting with *PIF1*. *Plant Physiol*. 164, 2068–2080. doi: 10.1104/pp.113.234492

Li, X., Qin, R., Du, Q., Cai, L., Hu, D., Du, H., et al. (2020). Knockdown of *GmVQ58* encoding a VQ motif-containing protein enhances soybean resistance to the common cutworm (*Spodoptera litura Fabricius*). *J. Exp. Bot*. 71, 3198–3210. doi: 10.1093/jxb/eraa095

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(T) (-Delta Delta C) method. *Methods* 25, 402–408. doi: 10.1006/meth.2001

May, G. (2008). Duplicated *P5CS* genes of Arabidopsis play distinct roles in stress regulation and developmental control of proline biosynthesis. *Plant J. Cell Mol. Biol.* 53, 11–28. doi: 10.1111/j.1365-313X.2007.03318.x

Mittler, R., Vanderauwera, S., Gollery, M., and Van, B. F. (2004). Reactive oxygen gene network of plants. *Trends Plant Sci.* 9, 490–498. doi: 10.1007/978-90-481-3112-9_5

Pascal, P., Lennart, E. L., Siska, H., Katja, K., Kai, N., Gerit, B., et al. (2014). The *Arabidopsis thaliana mitogen*-activated protein kinases *MPK3* and *MPK6* target a subclass of VQ-motif-containing proteins to regulate immune responses. *New Phytol*. 203, 592–606. doi: 10.1111/nph.12817

Petersen, K., Qiu, J., Lutje, J., Katrine, F. B., Sidsel, H., John, M., et al. (2010). Arabidopsis *MKS1* is involved in basal immunity and requires an intact N-terminaldomain for proper function. *PLoS One* 5:e14364. doi: 10.1371/journal.pone.0014364

Qian, M., Wang, L., Zhang, S., Sun, L., Luo, W., Posny, D., et al. (2021). Investigation of proline in superficial scald development during low temperature storage of 'Dangshansuli' pear fruit. *Postharvest Biol. Technol.* 181:111643. doi: 10.1016/j.postharvbio.2021.111643

Qin, L. X., Li, Y., Li, D. D., Xu, W. L., Zheng, Y., and Li, X. B. (2014). Arabidopsis drought-induced protein *Di19-3* participates in plant response to drought and high salinity stresses. *Plant Mol. Biol.* 86, 609–625. doi: 10.1007/s11103-014-0251-4

Rajametov, S., Yang, E. Y., Cho, M. C., Chae, S. Y., and Chae, W. B. (2021). Heat-tolerant hot pepper exhibits constant photosynthesis via increased 4 transpiration rate, high proline content and fast recovery in heat stress 5 condition. *Sci. Rep.* 11:14328. doi: 10.20944/preprints202107.0348.v1

Ren, Y. (2012). Effects of salt stress on different wheat varieties at seedling stage. *Hubei Agric. Sci.* 51, 3702–3705. doi: 10.14088/j.cnki.issn0439-8114.2012.17.054

Rorth, M., and Jensen, P. K. (1967). Determination of catalase activity by means of the Clark oxygen electrode. *Biochim. Biophys. Acta* 139, 171–173. doi: 10.1016/0005-2744(67)90124-6

Sequeira, L., and Mineo, L. (1966). Partial purification and kinetics of indoleacetic acid oxidase from tobacco roots. *Plant Physiol*. 41, 1200–1208. doi: 10.1104/pp.41.7.1200

Singh, A., Sagar, S., and Biswas, D. K. (2017). Calcium dependent protein kinase, a versatile player in plant stress management and development. *Crit. Rev. Plant Sci.* 36, 336–352. doi: 10.1080/07352689.2018.1428438

Song, S., Chen, Y., Zhao, M., and Zhang, W. H. (2011). A novel *Medicago truncatula* HD-Zip gene, *MtHB2*, is involved in abiotic stress responses. *Environ. Exp. Bot*. 69, 1–9. doi: 10.1016/j.envexpbot.2012.02.001

Song, W., Zhao, H., Zhang, X., Lei, L., and Lai, J. (2016). Genome-wideidentification of VQ motif-containing proteins and their expressionprofiles under abiotic stresses in maize. *Front. Plant Sci.* 6:1177. doi: 10.3389/fpls.2015.01177

Toufighi, K., Brady, S. M., Austin, R. S., Ly, E., and Provart, N. J. (2005). The botany array resource: e-northerns, expression angling, and promoter analyses. *Plant J.* 43, 153–163. doi: 10.1111/j.1365-313X.2005.02437

Verslues, P. E., Agarwal, M., Surekha, K. A., Zhu, J. H., and Zhu, J. K. (2006). Methods and concepts in quantifying resistance to drought, salt and freezing, abiotic stresses that affect plant water status. *Plant J.* 45, 523–539. doi: 10.1111/j.1365-313X.2005.02593.x

Wang, A., Garcia, D., Zhang, H., Feng, K., Chaudhury, A., Berger, F., et al. (2010). The VQ motif protein *IKU1* regulates endosperm growth and seed size in Arabidopsis. *Plant J.* 63, 670–679. doi: 10.1111/j.1365-313X.2010.04271.x

Wang, H., Hu, Y., Pan, J., and Yu, D. (2015). Arabidopsis VQ motif-containing proteins *VQ12* and *VQ29* negatively modulatebasal defense against Botrytis cinerea. *Sci. Rep.* 5:14185. doi: 10.1038/srep14185

Wang, M., Vannozzi, A., Wang, G., Zhong, Y., Corso, M., Cavallini, E., et al. (2015). A comprehensive survey of the grapevine *VQ* gene family and its transcriptional correlation with WRKY proteins. *Front. Plant Sci.* 6:417. doi: 10.3389/fpls.2015.00417

Wang, J., and Huang, R. (2019). Modulation of ethylene and ascorbic acid on reactive oxygen species scavenging in plant salt response. *Front. Plant Sci.* 10:319. doi: 10.3389/fpls.2019.00319

Wang, X., Zhang, H., Sun, G., Jin, Y., and Qiu, L. (2014). Identification of active VQ motif-containing genes and the expression patterns under low nitrogen treatment in soybean. *Gene* 543, 237–243. doi: 10.1016/j.gene.2014.04.012

Wang, Y., Liu, H., Zhu, D., Gao, Y., Yan, H., and Yan, X. (2017). Genome-wide analysis of VQ motif-containing proteins in Moso bamboo (*Phyllostachys edulis*). *Planta* 246, 165–181. doi: 10.1007/s00425-017-2693-9

Xiang, Y., Huang, Y., and Xiong, L. (2007). Characterization of stress-responsive CIPK genes in rice for stress tolerance improvement. *Plant Physiol.* 144, 1416–1428. doi: 10.1104/pp.107.101295

Xin, G. S., Zhao, Y. Q., Jiang, B., Xin, Z. W., and Liu, X. Y. (2019). Inhibition of MUC1-C regulates metabolism by AKT pathway in esophageal squamous cell carcinoma. *J. Cell. Physiol.* 234, 12019–12028. doi: 10.1002/jcp.27863

Xiong, L., Schumaker, K. S., and Zhu, J. K. (2002). Cell signaling during cold, drought, and salt stress. *Plant Cell* 14 (Suppl), S165–S183. doi: 10.1105/TPC.000596

Yu, C. S., Chen, Y. C., Lu, C. H., and Hwang, J. K. (2006). Prediction of protein subcellular localization. *Proteins* 64, 643–651. doi: 10.1002/prot.21018

Yu, C. S., Lin, C. J., and Huang, J. K. (2004). Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. *Protein Sci.* 13, 1402–1406. doi: 10.1110/ps.03479604

Zelm, E. V., Zhang, Y., and Testerink, C. (2020). Salt tolerance mechanisms of plants. *Annu. Rev. Plant Biol.* 71, 1–24. doi: 10.1146/annurev-arplant-050718-100005

Zhang, H., Zhu, J., Gong, Z., and Zhu, J. K. (2021). Abiotic stress responses in plants. *Nat. Rev.* 23, 104–119. doi: 10.1038/s41576-021-00413-0

Zhang, J., Ahmad, S., Wang, L. Y., Han, Q., Zhang, J. C., and Luo, Y. P. (2019). Cell death induced by α-terthienyl via reactive oxygen species-mediated mitochondrial dysfunction and oxidative stress in the midgut of *Aedes aegypti* larvae. *Free Radic. Biol. Med.* 137, 87–98. doi: 10.1016/j.freeradbiomed.2019.04.021

Zhang, Z., Liu, H., Sun, C., Ma, Q., Bu, H., Chong, K., et al. (2018). A C2H2 zinc-finger protein OsZFP213 interacts with OsMAPK3 to enhance salt tolerance in rice. *J. Plant Physiol.* 43, 167–176. doi: 10.1016/j.jplph.2018.07.003

Zhao, H., Ma, H., Yu, L., Wang, X., and Zhao, J. (2012). Genome-wide survey and expression analysis of amino acid transporter gene family in rice (*Oryza sativa* L.). *PLoS One* 7:e49210. doi: 10.1371/journal.pone.0049210

Zhao, Y., Ma, Q., Jin, X., Peng, X., Liu, J., Deng, L., et al. (2014). A novel maize homeodomain-leucine zipper (*HD-Zip*) I gene, *Zmhdz10*, positively regulates drought and salt tolerance in both rice and Arabidopsis. *Plant Cell Physiol.* 55, 1142–1156. doi: 10.1093/pcp/pcu054

Zhao, Y., Yu, W., Hu, X., Shi, Y., Liu, Y., Zhong, Y., et al. (2018). Physiological and transcriptomic analysis revealed the involvement of crucial factors in heat stress response of *Rhododendron hainanense*. *Gene* 660, 109–119. doi: 10.1016/j.gene.2018.03.082

Zhou, Y. B., Liu, C., Tang, D. Y., Yan, L., Wang, D., Yang, Y. Z., et al. (2018). The receptor-like cytoplasmic kinase *STRK1* phosphorylates and activates CatC, thereby regulating H_2O_2 homeostasis and improving salt tolerance in rice. *Plant Cell* 30, 1100–1118. doi: 10.1105/tpc.17.01000

Zhu, D., Chu, W., Wang, Y., Yan, H., Chen, Z., and Xiang, Y. (2018). Genome-wide identification, classification and expression analysis of the serine carboxypeptidase-like protein family in poplar. *Physiol. Plant.* 162, 333–352. doi: 10.1111/ppl.12642

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 02 June 2022
doi: 10.3389/fpls.2022.902989

Rapeseed (*Brassica napus*) Mitogen-Activated Protein Kinase 1 Enhances Shading Tolerance by Regulating the Photosynthesis Capability of Photosystem II

Zhen Wang[1,2†], Miao Liu[3†], Mengnan Yao[4†], Xiaoli Zhang[1], Cunmin Qu[1,2], Hai Du[1,2], Kun Lu[1,2], Jiana Li[1,2], Lijuan Wei[1,2] and Ying Liang[1,2]**

[1]Engineering Research Center for Rapeseed, College of Agronomy and Biotechnology, Southwest University, Chongqing, China, [2]Engineering Research Center of South Upland Agriculture of Ministry of Education, Academy of Agricultural Sciences, Chongqing, China, [3]Key Laboratory of Plant Resource Conservation and Germplasm Innovation in Mountainous Region (Ministry of Education), Institute of Agro-Bioengineering College, Guizhou University, Guiyang, China, [4]Jiangsu Yanjiang Institute of Agricultural Sciences, Nantong, China

OPEN ACCESS

Edited by:
Mukesh Jain, Jawaharlal Nehru University, India

Reviewed by:
Xiaoxu Li, Chinese Academy of Agricultural Sciences, China
Xuehai Zhang, Henan Agricultural University, China

***Correspondence:**
Ying Liang
yliang@swu.edu.cn
Lijuan Wei
lijuan525888@163.com

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

Received: *23 March 2022*
Accepted: *16 May 2022*
Published: *02 June 2022*

Citation:
Wang Z, Liu M, Yao M, Zhang X, Qu C, Du H, Lu K, Li J, Wei L and Liang Y (2022) Rapeseed (Brassica napus) Mitogen-Activated Protein Kinase 1 Enhances Shading Tolerance by Regulating the Photosynthesis Capability of Photosystem II. Front. Plant Sci. 13:902989. doi: 10.3389/fpls.2022.902989

Rapeseed (*Brassica napus*) is the third-largest source of vegetable oil in the world with an edible, medicinal, and ornamental value. However, insufficient light or high planting density directly affects its growth, development, yield, and quality. Mitogen-activated protein kinases (MAPKs) are serine/threonine protein kinases that play key roles in regulating the responses to biotic and abiotic stresses in plants. In this study, we found that the promoter of *BnaMAPK1* contained several light-responsive elements (including the AT1-motif, G-Box, and TCT-motif), consistent with its shading stress-induced upregulation. Compared with the wild type under shading stress, *BnaMAPK1*-overexpressing plants showed higher light capture efficiency and carbon assimilation capacity, enhancing their shading tolerance. Using RNA sequencing, we systematically investigated the function of *BnaMAPK1* in shading stress on photosynthetic structure, Calvin cycle, and light-driven electron transport. Notably, numerous genes encoding light-harvesting chlorophyll a/b-binding proteins (BnaLHCBs) in photosystem II-light-harvesting complex (LHC) II supercomplex were significantly downregulated in the *BnaMAPK1*-overexpressing lines relative to the wild type under shading stress. Combining RNA sequencing and yeast library screening, a candidate interaction partner of BnaMAPK1 regulating in shading stress, BnaLHCB3, was obtained. Moreover, yeast two-hybrid and split-luciferase complementation assays confirmed the physical interaction relationship between BnaLHCB3 and BnaMAPK1, suggesting that BnaMAPK1 may involve in stabilizing the photosystem II–LHC II supercomplex. Taken together, our results demonstrate that *BnaMAPK1* positively regulates photosynthesis capability to respond to shading stress in rapeseed, possibly by controlling antenna proteins complex in photosystem II, and could provide valuable information for further breeding for rapeseed stress tolerance.

Keywords: *BnaMAPK1*, shading stress, photosynthesis, photosystem II, *Brassica napus*

INTRODUCTION

Rapeseed (*Brassica napus*) is grown across the globe in various climatic conditions, from boreal to subtropical climates. It is utilized as leafy vegetable for human consumption, and also used for flower ornamental plant, oil production, biofuel, and animal feed (Confortin et al., 2019). The production of rapeseed is highly influenced by biotic and abiotic stresses, such as *Sclerotinia sclerotiorum*, *Peronospora parasitica*, light, temperature, drought, and salinity, all of which can drastically decrease yields (Raza et al., 2020; He et al., 2021; Raza, 2021); therefore, identifying stress-tolerance genes and understanding the related genetic network will provide a theoretical basis for developing new cultivars in rapeseed.

Plants can adapt to the complex environments by regulating gene expression and signal transduction pathways to advantageously alter their physiology and development. Mitogen-activated protein kinase (MAPK) cascades are highly conserved signal transduction pathways present in all eukaryotes (Cristina et al., 2010). A canonical MAPK network is composed of MAPK kinase kinases (MAPKKKs), MAPK kinases (MKKs), and MAPKs (Ichimura et al., 2002). The phosphorylated MAPKs in turn phosphorylate various downstream substrates, which are involved in the regulation of a wide variety of stress responses (Ichimura et al., 2002; Jonak et al., 2002; Pitzschke et al., 2009; Cristina et al., 2010; Zhang et al., 2018). MAPKs can be classified into four subfamilies (A–D), with subfamily C containing four members (MAPK1, MAPK2, MAPK7, and MAPK14) (Ichimura et al., 2002), which mainly play a role in abiotic stress responses. *AtMAPK1*, *AtMAPK2*, and *PsMAPK2* in pea (*Pisum sativum*) are all transcriptionally activated in response to wounding, jasmonic acid, and hydrogen peroxide (Ortiz-Masia et al., 2007, 2008). The four subfamily C MAPKs in *Arabidopsis* are activated by AtMKK3 to regulate the ABA signaling pathway (Danquah et al., 2015). It is likely that the MAPKs play varying roles in different stress responses (Teige et al., 2004).

Among several stress factors, light plays a vital role as it provides energy for photosynthesis and determines the growth, development, and morphogenesis of plants (Ruberti et al., 2012; Zoratti et al., 2014; Vialet-Chabrand et al., 2017). Insufficient light or shade condition is a pervasive abiotic stress in plant breeding and cultivation due to light blockage from intercropping, high planting density, horticulture facilities, cloud, rain, and snow. In rapeseed, shading stress is not conducive to photosynthetic efficiency, resulting in a decreased accumulation of dry matter and yield losses (Liang and Li, 2004). Photosynthate produced in the leaves, stem, and siliques are the primary nutrient source for rapeseed growth (Li et al., 2013; Lu et al., 2017), with the leaves being the most important until the flowering stage, when nutrients from the stem are also used for reproductive growth in the inflorescence (Li et al., 2013; Wang et al., 2016; Lu et al., 2017). At maturity, photosynthesis in the siliques facilitates the biosynthesis of proteins and lipids to store in the seeds (Elahi et al., 2016; Wang et al., 2016; Zafar et al., 2019). The reduction of photosynthesis under shading stress is therefore a major problem for rapeseed leaf and flower development and biomass accumulation, and overcoming this issue would be of great importance for breeding rapeseed with high edible, medicinal, and ornamental value.

In order to adapt to a shady environment, plants have evolved many strategies to increase their photosynthetic rate, including enhancing the stability of the photosystem I (PS I) and PS II complexes, the transcription, translation, and post-translational modification of photosynthesis-related genes and proteins (Walters, 2005). In higher plants, the light-harvesting complexes (LHCs) are divided into LHC I and LHC II groups, serving as PS I and PS II antennae, respectively (Klimmek et al., 2005, 2006). LHC II proteins are divided into four types (LHC II a–d), of which LHC II b is the major type. The three LHC II b proteins are encoded by the highly similar genes *LHCB1*, *LHCB2*, and *LHCB3*, and probably form homo- or hetero-trimers (Crepin and Caffarri, 2018). The minor LHC II proteins associated with PS II, CP29 (LHC II a), CP26 (LHC II c), and CP24 (LHC II d), are encoded by *LHCB4*, *LHCB5*, and *LHCB6*, respectively (Caffarri et al., 2009; Kirilovsky and Büchel, 2019). In the PS II complex, the inner antennae CP43 (PsbC) and CP47 (PsbB) bind to the D1 (PsbA) and D2 (PsbD) subunits to form core polypeptides, which then associate to the LHC II complex to form the PS II–LHC II supercomplex, playing a crucial role in capturing light in photosynthesis (Standfuss and Kühlbrandt, 2004).

The shading response affects the dynamic balancing of light transmission and distribution between PS I and PS II, which modulates the stability of the LHC II by the phosphorylation/dephosphorylation of protein kinases and phosphatases, increasing the light-harvesting area, and enhancing the efficiency with which light energy is used (Tikkanen and Aro, 2012; Rantala et al., 2020). Reversible and differential phosphorylation are dependent on the serine/threonine protein kinases 7 (STN7), STN8, and the thylakoid-associated phosphatase of 38-kD/protein phosphatase 1 (TAP38/PPH1) playing important roles in the phosphorylation of LHC II (Pribil et al., 2010; Tikkanen and Aro, 2012; Rantala et al., 2016, 2020). In addition, Zhang's lab discovers that another type of serine/threonine protein kinases, AtMAPK3 and AtMAPK6, can rapidly downregulate various components in the PS II–LHC II supercomplex to participate in the hypersensitive response cell death in *Arabidopsis* (Su et al., 2018). In rapeseed, however, the effects of limited light and the molecular mechanisms of the shade response, especially the cross-link between MAPKs and LHCBs, remain undercharacterized and poorly understood.

Here, we isolated the promoter of *BnaMAPK1* (ProBnaMAPK1) and identified a series of light-responsive *cis*-acting elements, and tested the effect of shading on *BnaMAPK1* expression. Using wild-type (WT) and transgenic *BnaMAPK1*-overexpressing (*BnaMAPK1*-OE) rapeseed, the mechanisms of the *BnaMAPK1* response to shading stress were, respectively, investigated on photosynthetic structure (pigment and enzyme system and photosynthesis-related complex), Calvin cycle, and light-driven electron transport through the examination of a combination of photosynthetic traits, RNA sequencing (RNA-Seq), and protein–protein interactions. This research aims to characterize the function of BnaMAPK1 under shady conditions,

determine how it may regulate shading tolerance, and identify the main signaling pathways and downstream targets of this MAPK in rapeseed. This study may provide new insights into MAPK cascades and help us to better understand the underlying biological process in a limited-light environment.

MATERIALS AND METHODS

Plant Materials and Growth Conditions

The black-seed doubled haploid inbred *B. napus* cultivar Zhongyou821 (WT group), three *BnaMAPK1*-OE T_3 lines (OE group), and tobacco (*Nicotiana benthamiana*) used in this study were obtained from Southwest University in China. The complete coding sequence of *BnaMAPK1* (*BnaA06g06010D*) under the control the CaMV *35S* promoter was transformed into Zhongyou821 to create OE materials with Basta resistance, as described previously (Weng et al., 2014). The rapeseed and tobacco seeds were soaked in Petri dishes and stratified at 4°C for 2–3 days. The seeds were then sown in nutritious soil in a growth chamber (PGR15, Controlled Environments, Winnipeg, MB, Canada), with a 16-h light (25°C)/8-h dark (20°C) photoperiod, 75% humidity, and an 800 $\mu mol \cdot m^{-2} \cdot s^{-1}$ light intensity. The photosynthetic photon flux density (PPFD) was measured using a quantum radiometer/photometer (LI-185B; LI-COR Biosciences, Lincoln, NE, United States).

Photosynthetic Light-Response Curve Analysis and Shading Treatment

For the photosynthetic light-response curve measurements, three healthy and mature leaves from the center of each 3–4-week-old Zhongyou821 seedling at the six-leaf stage (with four expand true leaves and two bud leaves) were selected and marked as fixed measured leaves. Curves were measured 10 times for each marked leaf, so that 30 times were measured for each seedling, with three seedlings used for each light intensity. During the measurement, the CO_2 concentration was set to the atmospheric CO_2 concentration, the flow rate of air in the measuring chamber was about 500 $\mu mol \cdot s^{-1}$, the temperature of the leaf chamber was 25°C ± 1°C, and the relative humidity was 75% ± 5%. For every measurement, the photosynthetically active radiation (PAR) was set at 1,400, 1,200, 1,000, 800, 500, 400, 300, 200, 100, 50, 20, and 0 $\mu mol \cdot m^{-2} \cdot s^{-1}$ by portable photosynthetic system (LI-6400; LI-COR Biosciences) to measure the net photosynthetic rate (Pn) under different light intensities (Li et al., 2019). For each PAR, the measurement time was controlled to 4 min and Pn was stabilized and recorded automatically by the instrument. The light saturation point (LSP), light compensation point (LCP), dark respiration rate (Rd), maximum net photosynthetic rate (Pnmax), and apparent quantum efficiency (AQE) were estimated based on the trend of the light-response curve by the modified rectangular hyperbola model (Ye et al., 2013).

Based on the photosynthetic light-response curve, a light intensity of 300 $\mu mol \cdot m^{-2} \cdot s^{-1}$ was selected to perform the shading treatment on the WT and three independent *BnaMAPK1*-OE lines (OE-1, OE-2, and OE-3) at the six-leaf stage. These plants were exposed to the shading treatment for 0, 7, 14, 21, and 28 days to investigate their photosynthetic traits under this stress. For the analysis of *BnaMAPK1* expression, the leaves of the WT seedlings were, respectively, collected after 0, 1.5, 3, 6, 9, 12, and 15 h of the shading treatment and used for RNA extraction. Leaves were also collected after 0 and 12 h of treatment for RNA-Seq. These leaves samples were transferred into RNase-free microfuge tubes and immediately placed into liquid nitrogen. Three biological duplicates were performed for each experiment.

Photosynthetic Trait Measurements

For the measurements of the physiological and chlorophyll fluorescence traits, WT and *BnaMAPK1*-OE leaves after 0 and 28 days of shading stress were measured using a LI-6400 portable photosynthetic system, with the same settings as were used for the light-response curve (Zhao et al., 2015; Pan et al., 2018). Each leaf was tested 10 times, and the average values after 0 and 28 days of the shading treatment were calculated. The biochemical traits include the Pn, stomatal conductance (Gs), and intercellular CO_2 concentration (Ci), and transpiration rate (Tr). For the chlorophyll fluorescence measurements, all plants were dark-adapted in advance for at least 3 h. The minimal fluorescence (F0), photochemical quenching coefficient (qP), non-photochemical quenching coefficient (qN), and maximum photochemical efficiency of PS II (Fv/Fm) were measured in each leaf and the average values were calculated (Li et al., 2020).

For the measurements of the biochemical traits, the relative content of chlorophyll (SPAD value) was determined in the WT and *BnaMAPK1*-OE by means of a chlorophyll meter (SPAD-502; Konica Minolta, Tokyo, Japan). Each leaf was marked at 10 representative points, and the average SPAD values after 0, 7, 14, 21, and 28 days of the shading treatment were calculated (Li et al., 2020). The Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) activity was assayed as described in the protocol of the plant Rubisco enzyme-linked immunosorbent assay kit (Cusabio Biotech, Wuhan, China). Fresh leaf material (0.5 g) was obtained from the WT and *BnaMAPK1*-OE seedlings after 0, 7, 14, 21, and 28 days of the shading treatment. Using the slightly modified enzyme extraction method of Xu et al., (2020), the material was homogenized with 6 ml of cooled extraction buffer containing 100 mM Tris–HCl (pH 7.8), 10 mM $MgCl_2$, 1 mM EDTA, 20 mM β-hydroxy-1-ethanethiol, and 2% (m/v) polyvinyl pyrrolidone at 0°C–4°C. The product was centrifuged at 14,000×*g* for 20 min at 4°C, and the supernatant was used for Rubisco activity measurement.

RNA Isolation, cDNA Biosynthesis, and Real-Time Quantitative PCR Analyses

To identify the relative expression levels of *BnaMAPK1* during shading stress, Zhongyou821 seedling leaf samples were harvested after 0, 1.5, 3, 6, 9, 12, and 15 h of the shading treatment. Total RNA was extracted from non-stressed and shade-stressed rapeseed seedling leaves using the RNAprep pure plant kit

(Tiangen Biochemical Technology, Beijing, China), and reverse transcription reaction was performed by the PrimeScript RT reagent kit (Takara Bio, Tokyo, Japan). The real-time quantitative PCR (qRT-PCR) was performed on a qTOWER2.2 qRT-PCR thermal cycler system (Analytik Jena, Jena, Germany). All qRT-PCRs were conducted in triplicate biological replications, and the data were comparatively quantified and calibrated using the qPCRsoft 3.1 system (Analytik Jena). The expression level of each gene was determined by the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001), using rapeseed *BnaACT7* (*BnaC02g00690D*) expression as the internal control (Liu et al., 2010). At least three independent biological repeats were performed for each data set. All the primers used to detect the gene expression are listed in **Supplementary Table S1**.

RNA-Seq Analysis

The RNA-Seq analysis was performed on WT and *BnaMAPK1*-OE rapeseed plants after 0 and 12 h of shading stress. The leaves, stems, and roots of seedlings were collected as mixed samples for total RNA extraction. The RNAs of three OE lines were extracted, respectively, and mixed with equal quality by three biological replicates, named as OE-Rep1, OE-Rep2, and OE-Rep3. Libraries were constructed and sequenced at Beijing Biomarker Technology (Beijing, China). An Illumina Hiseq X 10 sequencing platform was used to generate paired-end reads. Sequencing data has been deposited in the Sequence Read Archive (SRA) of the National Center for Biotechnology Information (NCBI) database with the accession BioProject ID: PRJNA680826.

The raw data were first processed to perform quality control to obtain clean reads which were then mapped to the *B. napus* reference genome sequence with Hisat2 tools (Kim et al., 2015). The gene expression levels were estimated using the fragments per kilobase of transcript per million fragments mapped (FPKM) method (Mortazavi et al., 2008). A differential expression analysis of the control/shading-treated plants and the *BnaMAPK1*-OE/WT groups was performed using DESeq (Wang et al., 2009). The following pairwise comparisons were performed: *BnaMAPK1*-OE samples with WT samples under the control (0h_OE vs. WT) and shading treatments (12h_OE vs. WT), and WT plants (WT_12h vs. 0 h) or *BnaMAPK1*-OE plants (OE_12h vs. 0 h) subjected to the control and stress conditions. Differentially expressed genes (DEGs) were defined as having a FPKM value > 1 for each of the three biological replicates, an absolute fold change (FC) value ≥ 2, and false discovery rate (FDR) < 0.05. The enrichment analysis of Gene Ontology (GO) and Kyoto encyclopedia of genes and genomes (KEGG) pathways of the DEGs in the 12h_OE vs. WT group was implemented by GOseq R packages with a $p.adjust < 0.05$ (Young et al., 2010) and KOBAS software with a $q\text{-}value < 0.05$ (Mao et al., 2005), respectively. qRT-PCR was performed to verify the RNA-Seq data (primers are listed in **Supplementary Table S1**).

Yeast Two-Hybrid and Split-Luciferase Complementation Assays

For the Yeast two-hybrid (Y2H) assays, the open reading frames of *BnaLHCB3* (*BnaA10g07350D*) and *BnaMAPK1* were amplified from Zhongyou821 and ligated into the pGADT7 prey vector and pGBKT7 bait vector, respectively. The confirmed *BnaLHCB3* and *BnaMAPK1* constructs were cotransformed into Y2Hgold cells and then plated onto SD-Leu-Trp medium and SD–Ade-His-Leu-Trp medium to observe the protein interactions. The interaction between pGBKT7-53 and pGADT7-T was the positive control, while the pGBKT7-Lam and pGADT7-T pair was used as the negative control.

For the split-Luciferase complementation (split-LUC) assays, *BnaLHCB3* and *BnaMAPK1* were recombined into the pCAMBIA1300-nLUC and pCAMBIA1300-cLUC vectors, respectively, which were then independently transformed into *Agrobacterium tumefaciens* strain GV3101. Cells were suspended in infiltration buffer (10 mM $MgCl_2$, 100 μM acetosyringone, 10 mM MES; pH 5.8) to an OD_{600} of 0.4–0.6 and incubated in the dark for at least 3 h at 28°C. Corresponding cLUC and nLUC cells were equally mixed and infiltrated into tobacco leaves. After 2 days, leaves were collected to examine LUC activity with 1 mM of the LUC substrate D-luciferin (APExBIO Technology, Houston, TX, United States). Luminescence intensity was measured by a CCD imaging apparatus (IVIS Lumina XRMS Series III; PerkinElmer, Waltham, MA, United States) with an exposure time of 2 min and binning settings 3 × 3. pCAMBIA1300-nLUC and pCAMBIA1300-cLUC were used as the negative and reference controls to normalize the LUC activity and eliminate variation in the experiment. The relative LUC activity values were obtained for three independent samples and used to calculate the arithmetic mean, as previously reported (Hou et al., 2016; Zhu et al., 2017). Primers are listed in **Supplementary Table S1**.

Statistical Analysis

All experiments were assessed using at least three independent repeats, and all data are represented as the mean ± SD. The statistical analysis was performed using a Student's *t*-test, and the significance was indicated using asterisks: *($p\text{-}value < 0.05$) or **($p\text{-}value < 0.01$).

RESULTS

BnaMAPK1 Can Be Induced by Shading Stress in Rapeseed

To identify the putative *cis*-acting regulatory elements in the promoter of *BnaMAPK1* in the Zhongyou821 rapeseed line, a 1,389-bp sequence upstream of the start codon was obtained using a PCR amplification and cloned based on the promoter of *BnaA06g06010D* from the Genoscope database. A bioinformatics analysis of the promoter sequence using the PlantCARE online tools (Lescot et al., 2002) revealed the presence of *cis*-acting elements that were shown to regulate gene expression in response to stress in plants. Several light-responsive elements were also identified, including the AT1-motif (ATTAATTTTACA), G-box (CACGTC), and TCT-motif (TCTTAC), suggesting that *BnaMAPK1* may be involved in responding to light stress. The detected putative light-responsive

elements and their functions are listed in **Supplementary Table S2**.

Using a quantum radiometer/photometer, the Pn of Zhongyou821 rapeseed was measured to investigate the light-response curve. Before the PAR reached 800 μmol·m^{-2}·s^{-1}, the Pn increased markedly, although little change was observed when the PAR was between 300 and 400 μmol·m^{-2}·s^{-1} (**Supplementary Figure S1A**). Using the Ye Zi-Piao model to fit the light-response curve, the fitted result was similar to the actual measured value, with 800 μmol·m^{-2}·s^{-1} set as the light saturation point and 300 μmol·m^{-2}·s^{-1} as the shading stress light intensity (**Supplementary Table S3**). To detect the response patterns of *BnaMAPK1* expression, the plants were subjected to 15 h of shading stress at 300 μmol·m^{-2}·s^{-1}. A qRT-PCR analysis showed that *BnaMAPK1* was upregulated after 3 h, reached a peak at 12 h, then decreased, indicating that *BnaMAPK1* was induced by shading stress (**Supplementary Figure S1B**).

BnaMAPK1 Responds to Shading Stress and Improves Shading Tolerance

To evaluate the role of *BnaMAPK1* in protecting the photosynthetic apparatus from shading stress, we produced *BnaMAPK1*-OE lines of rapeseed. In the three overexpression independent lines (OE-1, OE-2, and OE-3), the relative expression levels of *BnaMAPK1* were significantly higher than in WT rapeseed (**Supplementary Figure S1C**). To investigate how the rapeseed plants were affected under shading stress, the physiological traits of WT and *BnaMAPK1*-OE rapeseed were determined. Under normal light conditions (Control, 0 days), the Pn, Gs, and Ci displayed no significant differences between *BnaMAPK1*-OE and WT rapeseed, although the Tr values of the *BnaMAPK1*-OE-1 and *BnaMAPK1*-OE-3 lines were significantly higher than that of the WT (p-*value* < 0.05). When the plants were exposed to the shading treatment for 28 days, the Pn of both the *BnaMAPK1*-OE and WT lines decreased markedly, although the decrease was greater in the WT than in the *BnaMAPK1*-OE plants (p-*value* < 0.01) (**Figure 1A**), suggesting that overexpressing *BnaMAPK1* could alleviate the decrease in photosynthetic rate caused by shading stress. In addition, the observed changes in Gs, Ci, and Tr were similar to Pn, with significantly greater decreases observed in the WT than the transgenic plants under shading stress (**Figures 1B–D**). These data indicated that overexpressing *BnaMAPK1* enhances the gas exchange ability of rapeseed to improve its photosynthetic rate under long-term low-light environments.

To observe the energy dissipation, the chlorophyll fluorescence characteristics of F0, qP, qN, and Fv/Fm were measured. When plants were exposed to shading stress for 28 days, the F0 and qN increased in both the WT and *BnaMAPK1*-OE rapeseed relative to the normal light condition, with these traits remaining significantly lower in the *BnaMAPK1*-OE lines than in the WT under the shaded condition (p-*value* < 0.01) (**Figures 2A,C**). This suggests that overexpressing *BnaMAPK1* enhances the light capture efficiency to improve plants' sensitivity to light. Additionally, after 28 days of shading, qP and Fv/Fm decreased in the WT and *BnaMAPK1*-OE plants, indicating a decrease in the rate of light energy conversion. Both qP and Fv/Fm remained higher in the *BnaMAPK1*-OE lines than in the WT plants under shading stress, demonstrating that the overexpression of *BnaMAPK1* could effectively maintain the transmission of light in PS II to regulate photosynthetic ability during shading stress (**Figures 2B,D**).

Additionally, the biochemical characteristics of WT and *BnaMAPK1*-OE rapeseed were also assessed under shading stress. As shown in **Figure 1E**, the relative chlorophyll contents of the *BnaMAPK1*-OE lines were all significantly higher (p-*value* < 0.01) than the WT after 14 days of shading stress (*BnaMAPK1*-OE plants: 61.52, 52.01, and 50.66; WT: 44.20, 30.35, and 29.33 after 14, 24, and 28 days of shading). This demonstrated that overexpressing *BnaMAPK1* led to the retention of more chlorophyll, increasing the light energy utilization under shading stress in rapeseed. Rubisco is a key enzyme for photosynthesis, and was therefore investigated next. The activities of Rubisco in the three *BnaMAPK1*-OE lines were significantly higher than in the WT plants under normal light condition, while the shading treatments decreased the Rubisco activity of both the *BnaMAPK1*-OE and WT lines, although the *BnaMAPK1*-OE lines remained markedly higher than that of the WT (**Figure 2E**). These findings suggest that overexpressing *BnaMAPK1* could improve Rubisco activity, which would affect the net photosynthetic rate. Our data therefore confirmed that *BnaMAPK1* positively regulates light capture efficiency and carbon assimilation capacity to improve shading tolerance in rapeseed, which was consistent with the results of the promoter *cis*-acting element analysis.

Identification of DEGs Under Shading Stress Using RNA-Seq

To gain insight into the regulatory mechanisms by which *BnaMAPK1* is involved in shading stress, an RNA-Seq analysis was performed. The RNA used for the RNA-Seq was extracted from the leaves of the WT and *BnaMAPK1*-OE plants under normal light conditions and after 12 h of the shading treatment. Through pairwise comparisons, a total of 3,225, 3,000, 7,163, and 5,558 DEGs were, respectively, identified between the *BnaMAPK1*-OE and WT plants under normal light conditions (0h_OE vs. WT), *BnaMAPK1*-OE and WT under shading stress (12h_OE vs. WT), WT under normal light conditions and shading stress (WT_12h vs. 0 h), and *BnaMAPK1*-OE under normal light conditions and shading stress (OE_12h vs. 0 h), with a FDR < 0.05 and a $|\mathrm{Log_2FC}|$ threshold > 1 (**Figures 3A,B**). In the 0h_OE vs. WT, 12h_OE vs. WT, WT_12h vs. 0 h, and OE_12h vs. 0 h groups, 1,904, 1,346, 3,213, and 2,302 DEGs were upregulated and 1,321, 1,654, 3,950, and 3,256 were downregulated in the shade treatment/overexpression line relative to the normal light condition/WT (**Figure 3B**). Moreover, using $|\mathrm{Log_2FC}| > 3$ as a more stringent cutoff, **Figure 3C** shows the 164 DEGs (accounting for 5.47% of all DEGs) in the 12h_OE vs. WT comparison, of which 85 were upregulated and 79 were downregulated.

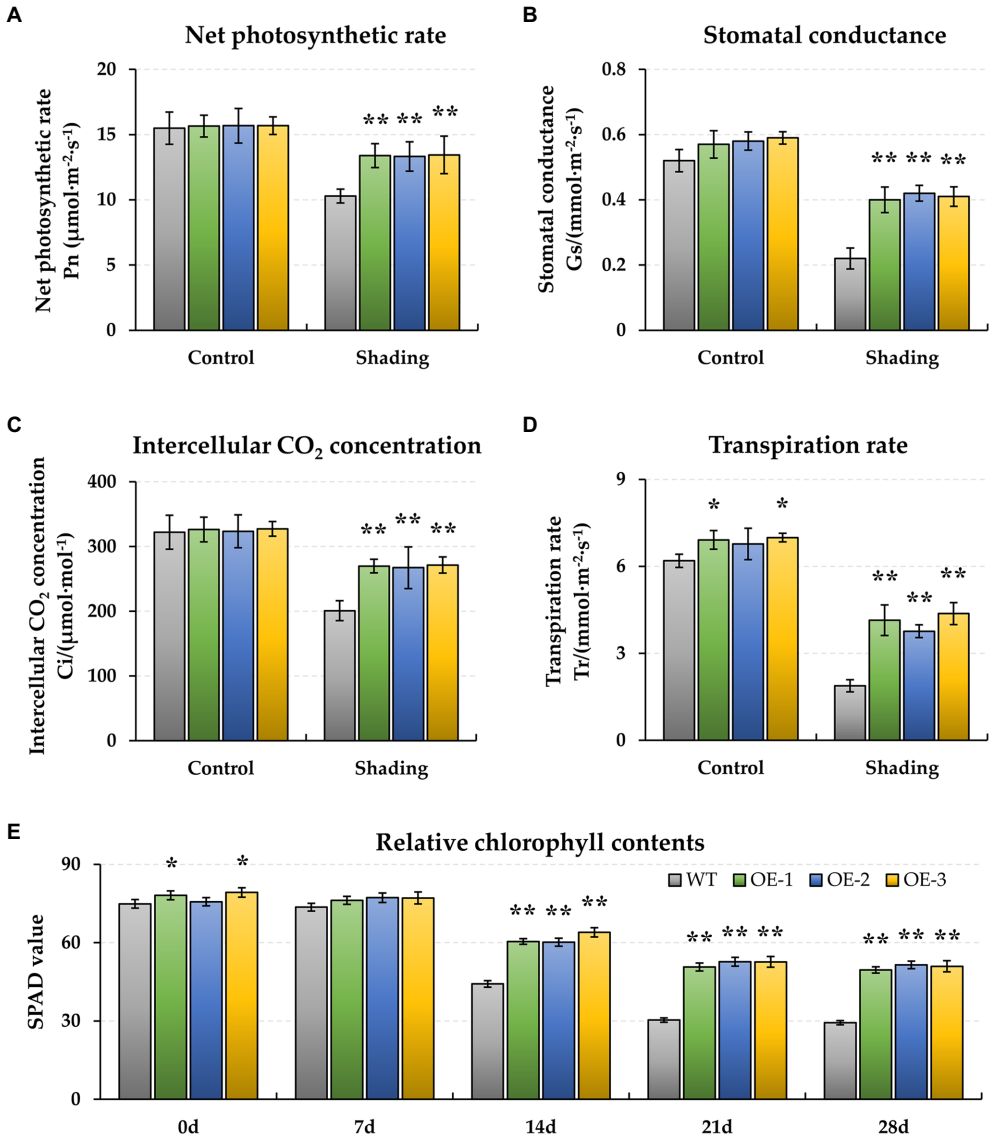


FIGURE 1 | Effect of shading stress on the physiological photosynthesis and pigment of wild type (WT) and *BnaMAPK1*-overexpressing (*BnaMAPK1*-OE) transgenic rapeseed. **(A)** Net photosynthetic rate (Pn). **(B)** Stomatal conductance (Gs). **(C)** Intercellular CO_2 concentration (Ci). **(D)** Transpiration rate (Tr). **(E)** Relative chlorophyll contents (SPAD). The physiological parameters (Pn, Gs, Ci, and Tr) were measured under normal light (control) and shading conditions at 0 and 28 days each, and the chlorophyll content (SPAD) at 0, 7, 14, 21, and 28 days each. Vertical bars are means of three replicates ± SD (three plants of each line, three leaves of each plant, and 10 times of each leaf were measured); asterisk and double asterisk indicate *p-value* < 0.05 and *p-value* < 0.01.

GO Enrichment Analysis of Photosynthesis-Related DEGs

A GO enrichment analysis was performed on the 3,000 DEGs in the 12h_OE vs. WT group to annotate their functions into three main categories: molecular function (MF), cellular component (CC), and biological process (BP). Using a *p.adjust* < 0.05, a total of 570 GO pathways were assigned to the 3,000 DEGs that responded to shading stress, with 19, 83, and 468 in the MF, CC, and BP categories, respectively. Within GO-MF, "oxidoreductase activity (GO:0016491)," "oxidoreductase activity, acting on single donors with incorporation of molecular oxygen (GO:0016701)," and "cobalt ion binding (GO:0050897)" were the most dominant enriched terms. Within GO-CC, the significantly enriched terms were related to "cytoplasmic part (GO:0044444)," "cytoplasm (GO:0005737)," and "plastid (GO:0009536)." Within GO-BP, the majority of the enriched GO terms were involved in "small molecule metabolic process (GO:0044281)," "oxoacid metabolic process (GO:0043436)," and "organic acid metabolic process (GO:0006082)." The eight most significantly overrepresented GO terms in each category are shown in **Figure 4A**.

Given that the overexpression of *BnaMAPK1* improves photosynthesis, we focused on the GO-BP category differences between the *BnaMAPK1*-OE and WT plants that were implicated

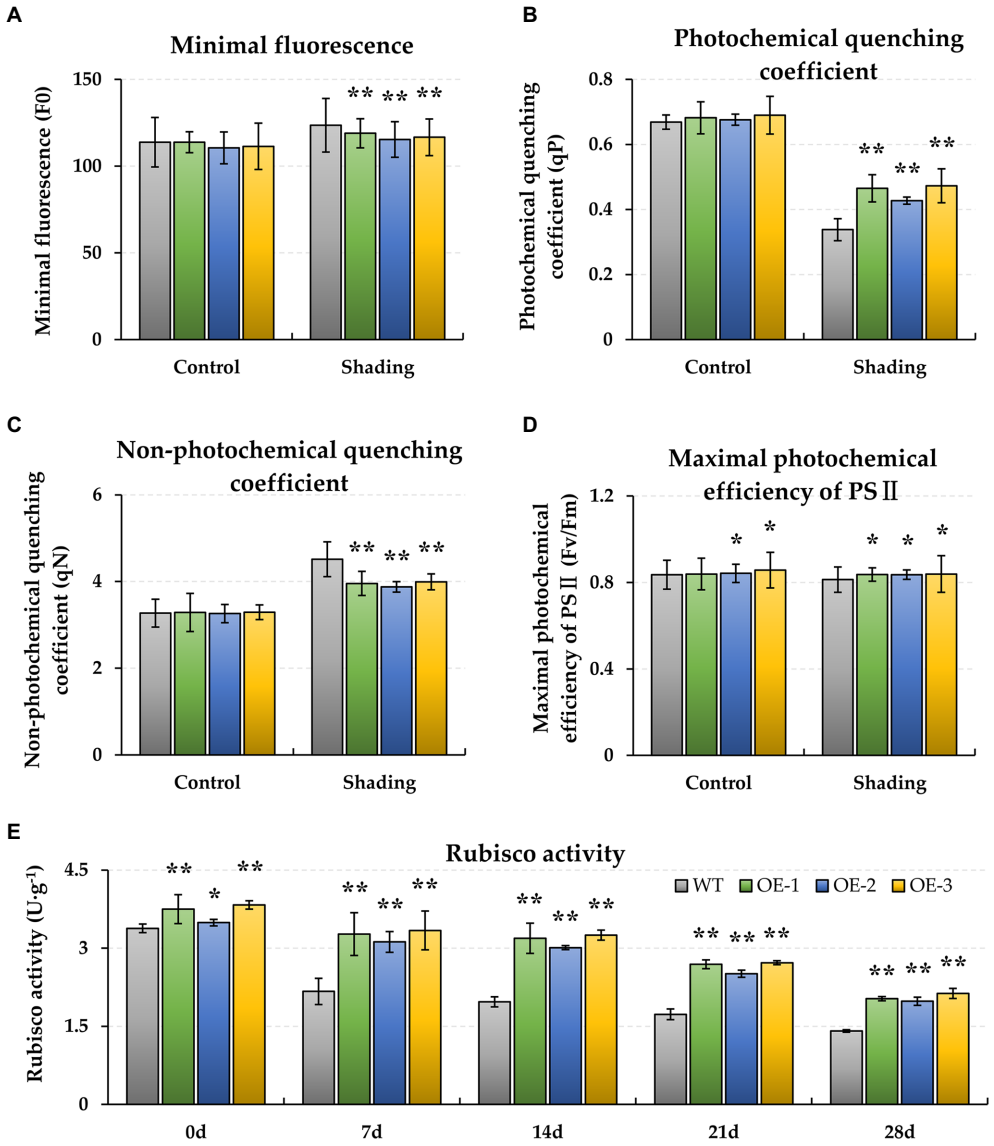


FIGURE 2 | Effect of shading stress on the chlorophyll fluorescence photosynthetic traits and Rubisco activity of WT and *BnaMAPK1*-OE transgenic rapeseed. **(A)** Minimal fluorescence (F0). **(B)** Photochemical quenching coefficient (qP). **(C)** Non-photochemical quenching coefficient (qN). **(D)** Maximum photochemical efficiency of PS II (Fv/Fm). **(E)** Rubisco activity. The chlorophyll fluorescence parameters (F0, qP, qN, and Fv/Fm) were measured under control and shading conditions at 0 and 28 days each, and the Rubisco activity at 0, 7, 14, 21, and 28 days each. Vertical bars are means of three replicates ± SD (three plants of each line, three leaves of each plant, and 10 times of each leaf were measured); asterisk and double asterisk indicate *p-value* < 0.05 and *p-value* < 0.01.

in the photosynthesis pathways, and identified 12 enriched GO pathways, including "response to light stimulus (GO:0009416)," "photosynthesis, light harvesting in PS II (GO:0009769)," and "response to light intensity (GO:0009642)" (**Figure 4B**; **Supplementary Table S4**). In these 12 photosynthesis-related GO-BP pathways, 276 DEGs were enriched in the 12h_OE vs. WT group, comprising 120 upregulated genes and 156 downregulated genes (**Supplementary Table S5**). According to the annotation, we found that the expression of *Rubisco activase* (*RCA*; *BnaA03g18710D*), *RAC-like 3* (*RAC*; *BnaA03g39270D*), *Early light-inducible protein* (*ELIP1*; *BnaCnng37300D*), and *Thioredoxin F2* (*TRXF2*; *BnaC02g06570D*) were significantly upregulated in *BnaMAPK1*-OE plants under shading stress, all of which encode the core subunits of PS I and PS II. We also observed that the key component of the light-harvesting antenna complex gene in PS II, *Light-harvesting complex photosystem II 4.2* (*LHCB4.2*; *BnaA05g29390D*), was significantly downregulated in the *BnaMAPK1*-OE plants compared with the WT under shaded conditions (**Supplementary Table S5**). Moreover, 19 DEGs were identified in this comparison using a $|Log_2FC| > 3$ cutoff, of which 12 were upregulated and seven were downregulated (**Figure 4C**; **Supplementary Table S5**). These data indicated that *BnaMAPK1* may participate in photosynthesis by regulating the transcription of the

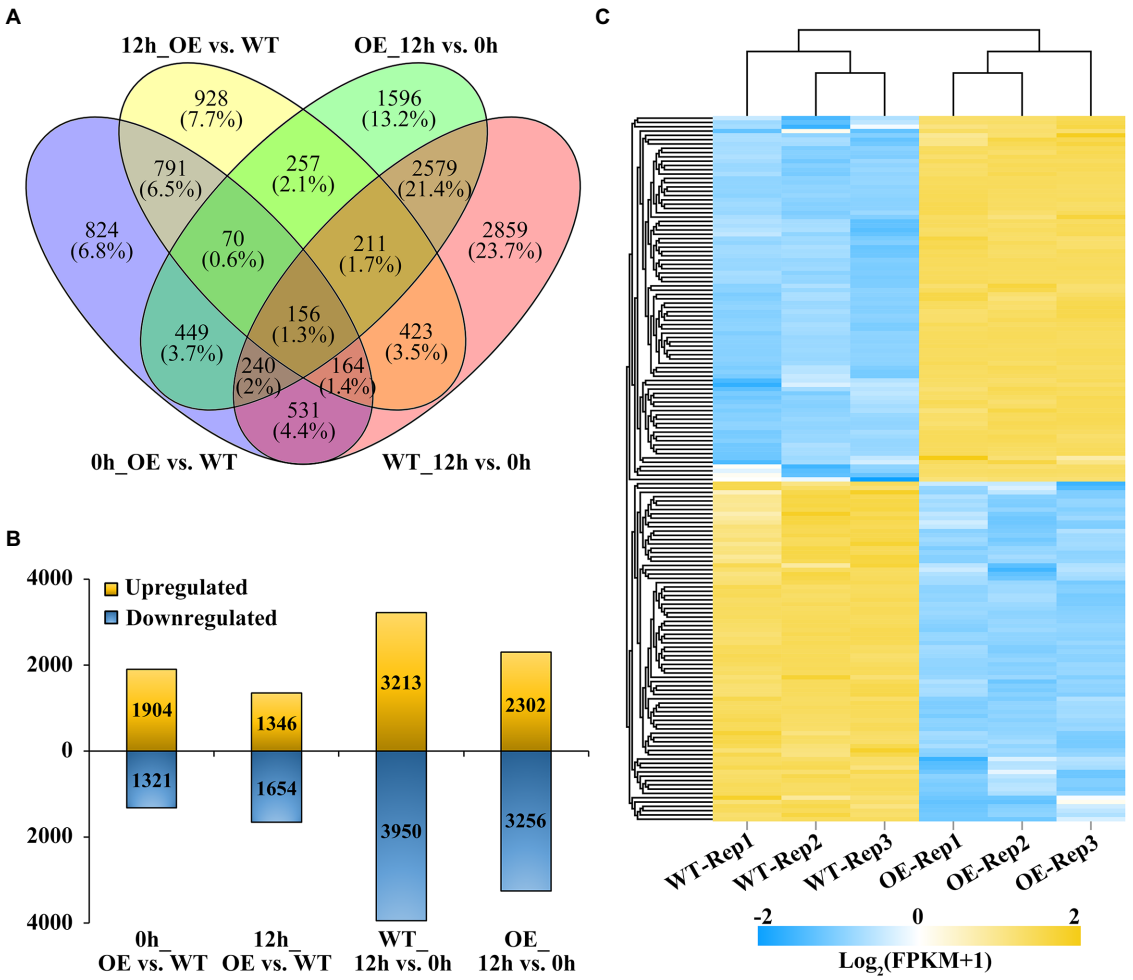


FIGURE 3 | Comparison of significant differentially expression genes (DEGs) identified in *BnaMAPK1*-OE and WT rapeseed under shading stress. **(A)** Venn diagram of DEGs in total by four comparisons of WT with *BnaMAPK1*-OE samples at control and shading stress conditions. DEGs screened with FPKM value of three replicates > 1, $|Log_2FC| > 1$, and FDR < 0.05 among each pairwise comparisons. **(B)** Statistics of the number of upregulated and downregulated DEGs in four comparisons of WT with *BnaMAPK1*-OE samples at control and shading stress. **(C)** Expression heatmap of significant DEGs in *BnaMAPK1*-OE and sWT rapeseed under shading stress. One hundred and sixty four DEGs screened with each FPKM value of three replicates > 1, $|Log_2FC| > 3$, and FDR < 0.05. 0h_OE vs. WT, *BnaMAPK1*-OE samples with WT at normal light; 12h_OE vs. WT, *BnaMAPK1*-OE samples with WT at shading stress; WT_12h vs. 0 h, normal light samples with shading treatment samples in WT; OE_12h vs. 0 h, normal light samples with shading treatment samples in *BnaMAPK1*-OE lines.

photosystem complex subunits to improve the shading tolerance of rapeseed.

KEGG Pathway Enrichment Analysis of the Photosynthesis-Related DEGs

For an exploration of the complex biological functions of the DEGs, the significantly enriched pathways were identified using KEGG analyses. The DEGs involved in these pathways are listed in **Supplementary Table S3**. The 3,000 DEGs from the 12h_OE vs. WT comparison were subjected to this KEGG characterization of their biological behaviors. In total, 18 KEGG pathways were obtained using a *q-value* < 0.05 (**Supplementary Table S6**). The top 10 KEGG pathways with the highest representation of DEGs are shown in **Figure 5A**, including "peroxisome (ko04146)," "valine, leucine and isoleucine degradation (ko00280)," "carbon metabolism (ko01200)," and "photosynthesis-antenna proteins (ko00196)." These data further confirmed the regulatory role of *BnaMAPK1* in the response to shading stress.

Based on the *BnaMAPK1*-mediated regulation of shading stress, we focused our analysis on 13 DEGs related to the "photosynthesis-antenna proteins" KEGG pathway, which includes genes related to PS II, such as *LHCB*. Based on their annotation, we found no obvious difference in the expression levels of genes encoding the PS I–LHC I supercomplex in the *BnaMAPK1*-OE plants compared with the WT under shading stress, while these 13 DEGs encoding LHCB subunits that function in the PS II–LHC II supercomplex were all significantly downregulated, especially *LHCB2.1* (*BnaC03g44110D*, *BnaA03g37990D*), *LHCB2.3* (*BnaC07g24660D*, *BnaC09g01520D*, *BnaA06g31910D*), *LHCB3* (*BnaA10g07350D*), *LHCB4.2* (*BnaA05g29390D*), and *LHCB6*

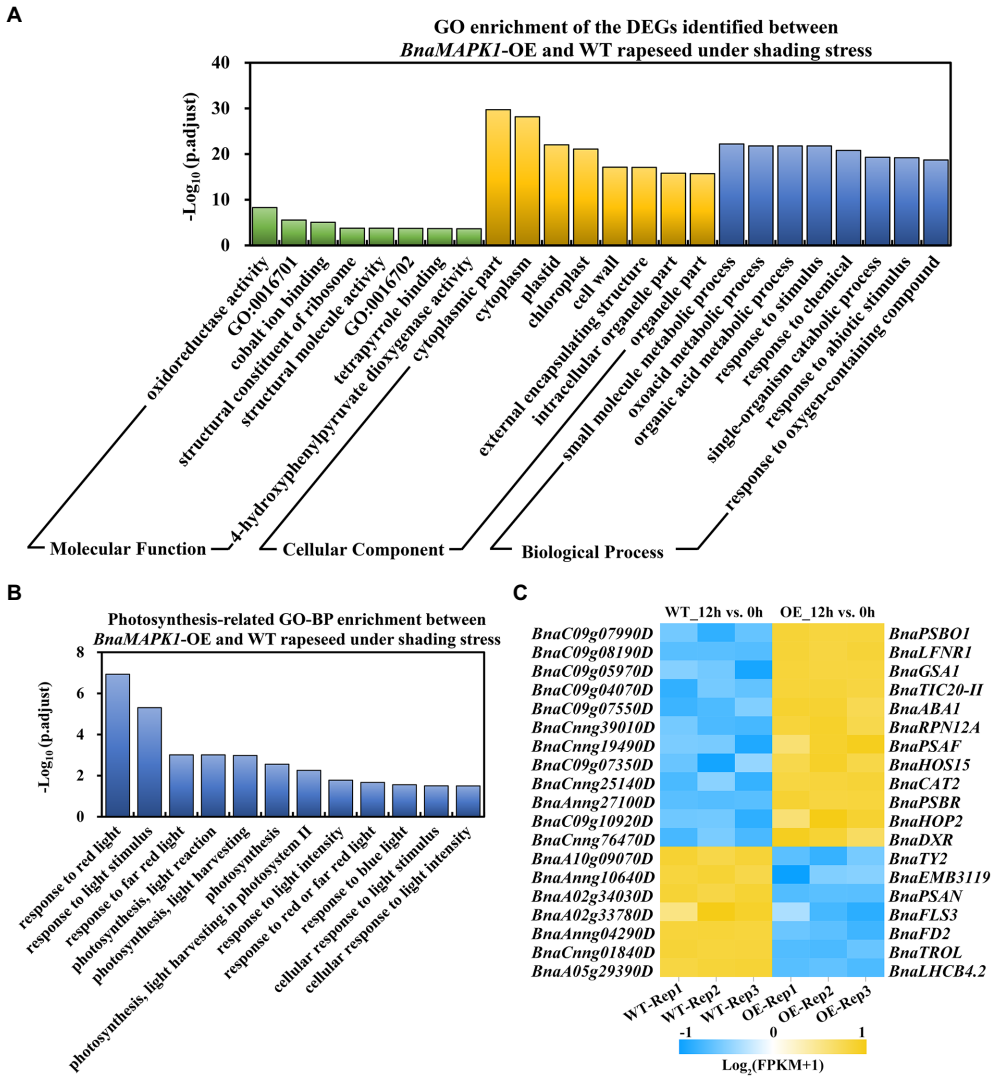


FIGURE 4 | Significantly enriched GO pathways of DEGs in *BnaMAPK1*-OE and WT rapeseed under shading stress. **(A)** Distribution of top eight significantly enrich GO pathways in the 3,000 DEGs in 12h_OE vs. WT group (*p.adjust* < 0.05). GO categories included molecular function, cellular component, and biological process. GO:0016701, oxidoreductase activity, acting on single donors with incorporation of molecular oxygen; GO:0016702, oxidoreductase activity, acting on single donors with incorporation of molecular oxygen, incorporation of two atoms of oxygen. **(B)** GO analysis of DEGs involved in photosynthesis-related biological process in 12h_OE vs. WT group under shading stress (*p.adjust* < 0.05). **(C)** Heatmap depiction of the *BnaMAPK1* regulated photosynthesis-related DEGs in 12h_OE vs. WT group under shading stress. FPKM value of three replicates > 1, $|Log_2FC| > 3$, and FDR < 0.05.

(*BnaC08g38660D*) (**Figures 5B,C**). These data indicated that the expression of genes encoding LHCB subunits of the PS II–LHC II supercomplex was regulated by *BnaMAPK1* during shading stress.

qRT-PCR Validation of the DEGs in the "Photosynthesis-Antenna Proteins" Pathway

To verify the authenticity of the expression levels obtained from the RNA-Seq data, the expression of nine selected *LHCB* genes were subjected to qRT-PCR analysis. *LHCB1.3* (*BnaA04g20150*), *LHCB1.4* (*BnaC03g18980*, *BnaA05g09410*, *BnaA03g15830*, *BnaA05g09380*), *LHCB2.3* (*BnaC09g01520*), *LHCB3* (*BnaA10g07350*), *LHCB4.2* (*BnaA05g29390*), and *LHCB6* (*BnaC08g38660*) were selected for validation, and *BnACT7* (*BnaC02g00690D*) was used as the internal control. The RNA templates for the qRT-PCR were obtained from 3 to 4-week-old mixed samples (leaf, stem, and root) of WT and *BnaMAPK1*-OE plants subjected to 0 or 12 h of the shading treatment. As shown in **Figure 5D** and **Supplementary Figure S2**, the changes in transcript expression revealed using RT-PCR were identical to those acquired using RNA-Seq. Collectively, the above expression results suggest that the RNA-Seq data were credible and that *BnaMAPK1* responds to shading stress by negatively regulating the *LHCB* genes in the PS II–LHC II supercomplex.

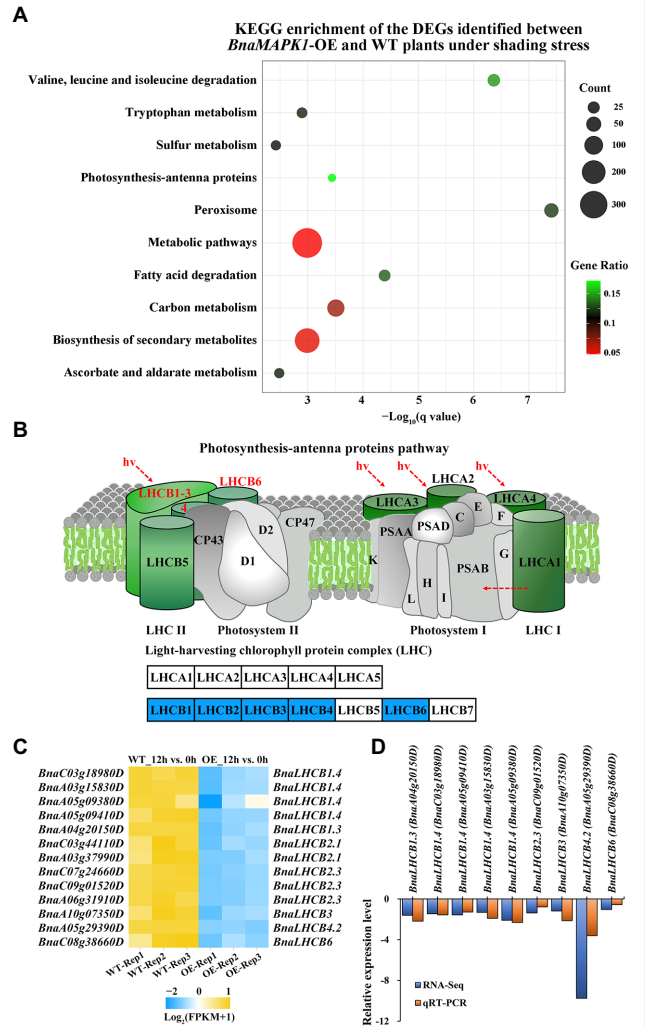


FIGURE 5 | Photosynthesis-antenna proteins downregulated in *BnaMAPK1*-OE rapeseed relative to WT plants under shading stress in RNA-Seq. **(A)** Top 10 significantly enriched KEGG pathways in the 3,000 DEGs in 12h_OE vs. WT pairwise comparisons. The advanced bubble chart shows the enrichment of the DEGs in the signaling pathways. The *x*-axis label represents the enrichment significance (*q-value* < 0.05), and the *y*-axis label indicates the pathway names. The size of the circle represents the number of genes enriched in a particular pathway, and the color of bubble presents the enrichment factor (gene ratio = amount of DEGs enriched in the pathway/total number of genes in background gene set). **(B)** Downregulated *BnaLHCBs* encoding LHCB subunits in KEGG map of PS II-LHC II supercomplex in 12h_OE vs. WT pairwise comparisons. **(C)** Heatmap clustering of 13 *BnaLHCBs* transcript abundances associated with photosynthesis-antenna proteins pathway. Each DEG was screened with FPKM value of three replicates > 1, $|Log_2FC| > 1$, and FDR < 0.05. **(D)** The relationship comparison between qRT-PCR and RNA-Seq fold change results.

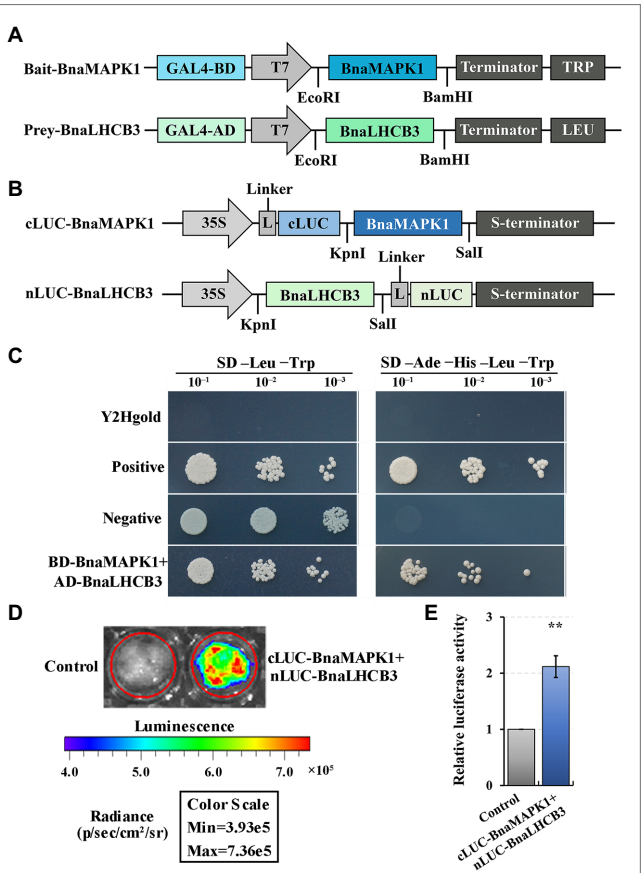


FIGURE 6 | Verifying the interaction relationship of BnaMAPK1 with BnaLHCB3. **(A)** Schematic diagram of the expression vectors used in Y2H assay. **(B)** Schematic diagram of the expression vectors used in Split-LUC assay. **(C)** BnaMAPK1 interacted with BnaLHCB3 in Y2H assay. Y2Hgold cells served as blank controls, Y2Hgold cells expressing pGBKT7-53 and pGADT7-T served as positive controls, and Y2Hgold cells expressing pGBKT7-Lam and pGADT7-T served as negative controls. **(D)** BnaMAPK1 interacted with BnaLHCB3 in Split-LUC assay. Empty pCAMBIA1300-nLUC and pCAMBIA1300-cLUC were used as controls for normalization. **(E)** Relative luciferase activities assays for determining whether BnaMAPK1 could interact with BnaLHCB3 in tobacco leaves. All the experiments were repeated three times with similar results. Double asterisk above bars indicate significant difference at *p-value* < 0.01.

Verification of the Physical Interaction of BnaLHC3 With BnaMAPK1 Using Y2H and Split-LUC System

Our previous study revealed that a subunit of the PS II–LHC II supercomplex, BnaLHCB3 (BnaA10g07350D), is a candidate interaction partner of BnaMAPK1, as determined using Y2H library screening assays (Wang et al., 2020). To further explore whether BnaLHCB3 was regulated by BnaMAPK1, the full-length coding sequences of *BnaLHCB3* and *BnaMAPK1* were inserted into the pGADT7 and pGBKT7 vectors, respectively (**Figure 6A**). The prey and bait constructs were cotransformed into yeast to explore the point-to-point protein interactions. The Y2Hgold strain expressing *BnaLHCB3* and *BnaMAPK1* grew well on the SD–Ade-His-Leu-Trp medium (**Figure 6C**), indicating that BnaLHCB3 could interact with BnaMAPK1 in yeast cells.

To further elucidate the physical interaction of BnaLHCB3 and BnaMAPK1, tobacco-based firefly luciferase complementation assays were performed. *BnaLHCB3* was inserted into the N-terminal of LUC, while *BnaMAPK1* was linked to the C-terminal of LUC (**Figure 6B**). The coexpression of the

nLUC-BnaLHCB3 and cLUC-BnaMAPK1 constructs showed a strong ability to rescue the luciferase enzyme activity (*p-value* < 0.01) (**Figures 6D,E**). Taken together, these results suggested that BnaLHCB3 and BnaMAPK1 interact *in vivo*, which might affect the stability of the PS II–LHC II supercomplex to further regulate shading tolerance in rapeseed.

DISCUSSION

Shade is widespread in nature, affecting many plants throughout their lifecycle. Plants respond on cellular and molecular levels to adapt to limited light. In our study, the AT1-motif, Box 4, G-box, GAG-motif, Sp1, TCT-motif, and As-2-box elements were identified in the promoter of *BnaMAPK1* (**Supplementary Table S2**), all of which are involved in the light-mediated transcriptional regulation of light-regulated genes in potato (*Solanum tuberosum*), maize (*Zea mays*), and *Arabidopsis thaliana* (Shariatipour and Heidari, 2020). Notably, a G-box in a promoter can bind tightly to phosphorylated proteins to negatively regulate the transcriptional expression of the *LHC II* genes to further respond to light in a variety of plants. This suggested that *BnaMAPK1* may play an important role in regulating photosynthesis in response to changing light levels.

BnaMAPK1 Regulates Light Capture Efficiency of the Photosynthetic Apparatus in Xanthophyll Cycle and Photosynthetic Pigment System

Previous studies show that the photosynthetic capability of shade-tolerant rapeseed is increased to extend the vegetative phase under a light-restricted environment (Brunel-Muguet et al., 2013), as well as optimized light capture and utilization through increased photosynthetic structures, primary reactions, and carbon assimilation in plants (Niinemets, 2010). Since the shading-induced expression pattern of *BnaMAPK1* was consistent with the *cis*-acting elements analysis of its promoter, we examined the photosynthetic structure of WT and *BnaMAPK1*-OE rapeseed under shade stress. In plants, Pn is the most important indicator for evaluating photosynthetic ability, especially under an artificial environment when it is provided with a constant light quality and intensity (Shengxin et al., 2016). Our data showed that the gas exchange values (Pn, Gs, Ci, and Tr) of both WT and *BnaMAPK1*-OE rapeseed decreased in the shading treatment, with the values in *BnaMAPK1*-OE higher than those in the WT (**Figures 1A–D**). Previously, however, silencing *Nicotiana attenuata NaMAPK4* (the homolog of *Arabidopsis AtMAPK4*) greatly enhances stomatal conductance, transpiration, and the photosynthetic rate in an ABA-dependent manner (Hettenhausen et al., 2012). Under shading stress, the decrease in Gs and Ci caused the decrease of Tr, which in turn caused the decrease of Pn, suggesting that *BnaMAPK1* plays a positive role in regulating the photosynthetic rate, which might be inconsistent with the mechanism by which *AtMAPK4* functions. In terms of energy dissipation, the trends of qP and Fv/Fm were similar to the gas exchange values, while the F0 and qN values of the WT and *BnaMAPK1*-OE plants increased after the shading treatment, with the WT displaying the higher values (**Figures 2A–D**). Photochemical damage is reflected in either an increase in F0 or decrease in the ratio of Fv/Fm under both shade stress and high-light conditions (Qiyuan and Ibrahim, 2016), implying that shading stress may limit the activity of the light reaction center and increase the energy dissipation in rapeseed. Our photosynthetic parameters demonstrated that *BnaMAPK1*-OE rapeseed maintained low heat dissipation to improve light capture efficiency to further enhance its photosynthetic rate under shading.

Xanthophylls play essential roles in both light absorption and photoprotection, acting as accessory pigments and structural elements of the antennae to stabilize the LHC II complex (Ballottari et al., 2012). In higher plants, the xanthophyll cycle is closely related to photosynthesis and photoinhibition under a shaded environment. This cycle mainly involves thermal energy dissipation, for which zeaxanthin epoxidase (ZEP) and violaxanthin de-epoxidase (VDE) are key catalyzing enzymes (Hieber et al., 2000). To gain a better understanding of how *BnaMAPK1* enhances photosynthetic capability under shading stress, GO pathway and related DEG analyses were performed in RNA-Seq. Our data showed that the expression of *Zeaxanthin epoxidase 1* (*ABA1/ZEP*; *BnaC09g07550D*) and *violaxanthin de-epoxidase-related protein* (*VDE*; *BnaCnng13890D*) were upregulated in transgenic rapeseed after the shading treatment (**Supplementary Table S5**), which was consistent with its photosynthetic traits. These data suggested that *BnaMAPK1* functions in the xanthophyll cycle of the photosynthetic apparatus to regulate energy dissipation and improve the light sensitivity of PS II.

Light affects the biosynthesis of the photosynthetic pigments, which in turn affects photosynthesis itself and plays an important role in plant growth and development. One such pigment, chlorophyll, absorbs and transmits light energy for photosynthesis (Sims and Gamon, 2002; Gao et al., 2020). In the present study, the relative chlorophyll contents of both WT and *BnaMAPK1*-OE were found to decrease under the shading treatment, but a long-term (14, 21, and 28-d) shading treatment resulted in significantly higher relative chlorophyll contents in *BnaMAPK1*-OE rapeseed compared with the WT (**Figure 1E**). Notably, three *Chlorophyllase 1* (*CLH1*; *BnaA02g18330D*, *BnaA06g13830D*, and *BnaC05g15260D*) genes, which are involved in chlorophyll degradation in PS II (Tian et al., 2021), were differentially expressed in the two genotypes, and were shown to be downregulated in the transgenic plants under shading stress. Here, the expression level of *BnaELIP1* (*BnaCnng37300D*), which belongs to the LHC-like protein family and is a target gene for controlling chlorophyll accumulation to enhance tolerance to high light and shading stresses (Heddad and Adamska, 2002), was upregulated in *BnaMAPK1*-OE rapeseed under the shading treatment (**Supplementary Table S5**). In *Arabidopsis* and *Chlamydomonas reinhardtii*, ELIP1 associates with LHCB antennae to function in the xanthophyll cycle, carbon assimilation, and chlorophyll accumulation (Stress et al., 2006; Lee et al., 2020). These findings indicated that the

overexpression of *BnaMAPK1* reduces the damage caused to the photosynthetic pigments by shading stress, maintaining photosynthetic capability.

BnaMAPK1 Regulates Carbon Assimilation in Photosynthetic Enzyme System and Calvin Cycle in PS II

In addition to the photoreaction and pigment systems, the enzyme system is also an important component of photosynthesis. Rubisco is the major enzyme involved in carbon assimilation and the limiting factor of photosynthetic efficiency and productivity (Mitra and Baldwin, 2008). Lower levels of Rubisco, ATP synthase, and PS II activities, as well as less electron transport and CO_2 consumption, are observed in shade-grown barley (*Hordeum vulgare*) leaves (Zivcak et al., 2014). By contrast, our photosynthetic measurements showed that the Rubisco activity in *BnaMAPK1*-OE rapeseed was significantly higher than the WT under both normal light and shade, which was positively correlated with the Pn values (**Figure 2E**). These findings are consistent with those of Liang et al., who report that shade-tolerant rapeseed has a higher photosynthetic net rate and chlorophyll content than the WT, as well as more stable Rubisco activity under shading stress (Liang and Li, 2004). Our results demonstrated that the photosynthetic enzyme system was less affected by shading in the *BnaMAPK1*-OE plants compared with WT rapeseed.

The Calvin cycle is responsible for photosynthetic carbon assimilation in plants. Relevant studies on shady environments reveal that the abundances of Rubisco, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), and F-type ATPase, which are involved in the Calvin cycle, electron transport, and carbon assimilation, respectively, are reduced, suppressing photosynthesis (Zivcak et al., 2014; Liu et al., 2020). Rubisco catalyzes the first step of photosynthetic carbon assimilation and is a rate-limiting enzyme for photosynthetic efficiency and productivity in the Calvin cycle, in which Rubisco activity is regulated by a second enzyme, RCA (Portis, 2003; Bracher et al., 2017). Compared with the WT under shading stress, our RNA-Seq analysis showed that the expression level of *BnaRCA* (*BnaA03g18710D*) was upregulated in the *BnaMAPK1*-OE lines, which was consistent with the finding that overexpressing *BnaMAPK1* improved the Rubisco activity of shaded rapeseed. Moreover, previous studies report that GAPDH is involved in carbon assimilation and limited the regeneration of ribulose-1,5-bisphosphate (RuBP) in the Calvin cycle (Suzuki et al., 2021). We also found that *Glyceraldehyde-3-phosphate dehydrogenase C1* (*BnaGAPC1*; *BnaA09g46260D* and *BnaC08g40330D*) and *BnaGAPC2* (*BnaA05g33200D*), encoding GAPDH, were all significantly downregulated in the *BnaMAPK1*-OE lines, indicating that *BnaMAPK1* may play key roles in the Calvin cycle by accelerating RuBP carboxylation and limiting carbon assimilation. We therefore propose that the overexpression of *BnaMAPK1* is beneficial to shade-stressed rapeseed because it improves the efficiency of light energy conversion and allows for decreased energy dissipation during the carbon assimilation in PS II in a shaded environment.

BnaMAPK1 Interacts With BnaLHCB3 to Participate in Photosynthesis-Antenna Proteins Pathway

Enhancing the photosynthetic rate is critical for increasing crop yields to meet the rising demands for food (Zhu et al., 2010; Long et al., 2015). The GO pathway analysis of the *BnaMAPK1*-OE lines in this study showed a significant enrichment in not only genes associated with the light reaction and harvesting in photosynthesis, but also those involved in plastoquinone biosynthesis (**Figure 4**; **Supplementary Table S4**). Of the DEGs, *Phytoene desaturation 1* (*BnaPDS1*; *BnaA09g49870D*, *BnaA10g04310D*, *BnaC05g04530D*, and *BnaC08g44820D*) and *Ferredoxin 3* (*BnaFD3*; *BnaA07g13260D*, *BnaC04g16810D*) were downregulated in *BnaMAPK1*-OE rapeseed compared with the WT under shading stress. *AtPDS1* was previously reported to function in plastoquinone biosynthesis, where it affects the availability of the electron carrier plastoquinone by controlling p-hydroxyphenylpyruvate dioxygenase activity during the light response in *Arabidopsis* (Qin et al., 2007). In higher plants, FD3 is a 2Fe2S plant-type iron–sulfur protein and associates with other ferredoxins to respond to a limitation in PS I acceptors (Voss et al., 2011). In the photoreaction process, ferredoxins act as electron acceptors of the photoreaction terminal, mediating electron transfer between PS I and carbon assimilation, and regulating electron reflow in photosynthetic phosphorylation (Fan et al., 2019). Electrons can be recycled from reduced ferredoxin at PS I to plastoquinone at PS II in the light-dependent reactions of photosynthesis (Munekage et al., 2004); therefore, based on the qP values presented here, it seems probable that *BnaMAPK1* influences the Calvin cycle and the xanthophyll cycle to further regulate electron transport to adapt to the shade, but is not highly likely to directly respond to an acceptor limitation at PS I or PS II in rapeseed.

Further mining our transcriptomic data, we conducted a KEGG pathway enrichment analysis to identify the potential genes and pathways involved in photosynthesis in the DEGs. We identified the enrichment of an important KEGG pathway, the photosynthesis-associated antenna proteins pathway (**Figure 5**; **Supplementary Table S6**). In higher plants and green algae, the chlorophyll-binding subunits of PS I and PS II are internal antennae; LHCs act as peripheral antennae that enable a more efficient absorption of light energy and rapidly transfer energy to the reaction center to improve the photosynthetic rate (Papadatos et al., 2017). Zhang's lab discovers that AtMAPK3/AtMAPK6 activation can rapidly downregulate the PS II–LHC II supercomplex to inhibit photosynthesis (Su et al., 2018); however, combining the RNA-Seq and RT-PCR data, we identified an interesting group of DEGs encoding photosynthesis-associated antenna proteins in PS II, the BnaLHCBs, which are essential for photoprotection in both high-light and low-light stresses (Ballottari et al., 2012). These genes were significantly downregulated in the *BnaMAPK1*-OE

rapeseed, including *BnaLHCB1*, *BnaLHCB2*, *BnaLHCB3*, *BnaLHCB4*, and *BnaLHCB6* (**Figure 5**; **Supplementary Figure S2**). In *Arabidopsis*, six *lhcb* single mutants, *lhcb1–6*, are investigated to show that each AtLHCB plays a specific role in photosynthesis and ABA signaling (Xu et al., 2012). Each of the five *BnaLHCB*s showed similar shading-sensitive expression patterns in our study, suggesting that all *BnaLHCB*s were necessary for building the antenna complex and keeping it intact. It is tempting to speculate that *BnaMAPK1* plays an important role in regulating the stability of the core molecular complex of the PS II antenna machinery, and that these *BnaLHCB*s might be the interacting partners of *BnaMAPK1* in regulating photosynthesis to improve the shade tolerance of rapeseed.

A previous proteomics analysis reports that differentially expressed proteins were also mapped to the photosynthesis-associated antenna proteins pathway; GmLHCB1–6 (except GmLHCB3) are significantly differentially abundant in soybean plants subjected to shading stress (Fan et al., 2019). Shading stress also induces PS state transitions (state 1–state 2 transitions), which maximize the efficiency of light harvesting at low light intensities (Mikko et al., 2006). However, LHCB3 affects the excitation energy transfer, the macrostructure of PS II, and the state transitions in *Arabidopsis* (Adamiec et al., 2015), and confers continuous-light tolerance and enhances yields in tomato (*Solanum lycopersicum*) and oil palm (Velez-Ramirez et al., 2014; Neoh et al., 2019). We previously use Y2H library screening assays to show that BnaLHCB3 is a candidate interacting partner of BnaMAPK1 (Wang et al., 2020). In the present study, we used point-to-point Y2H and split-LUC results to validate this interaction relationship between BnaMAPK1 and BnaLHCB3 (**Figure 6**), suggesting the function of BnaMAPK1 in stabilizing the PS II–LHC II supercomplex under shading stress in rapeseed.

Taken together, our findings provide useful insights into the mechanisms underlying the rapeseed shading response through the exploration of *BnaMAPK1*, which is mainly involved in PS II-associated processes. This gene is particularly important in regulating the efficiency of the xanthophyll cycle and chlorophyll-mediated light capture and energy dissipation, controlling the Calvin cycle-mediated carbon assimilation, and stabilizing the PS II–LHC II supercomplex to improve photosynthesis capacity (**Figure 7**). This study provides important evidence for the molecular mechanisms underlying the shading response of the MAPK cascades, and is potentially relevant for use in molecular breeding strategies.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA680826/, PRJNA680826.

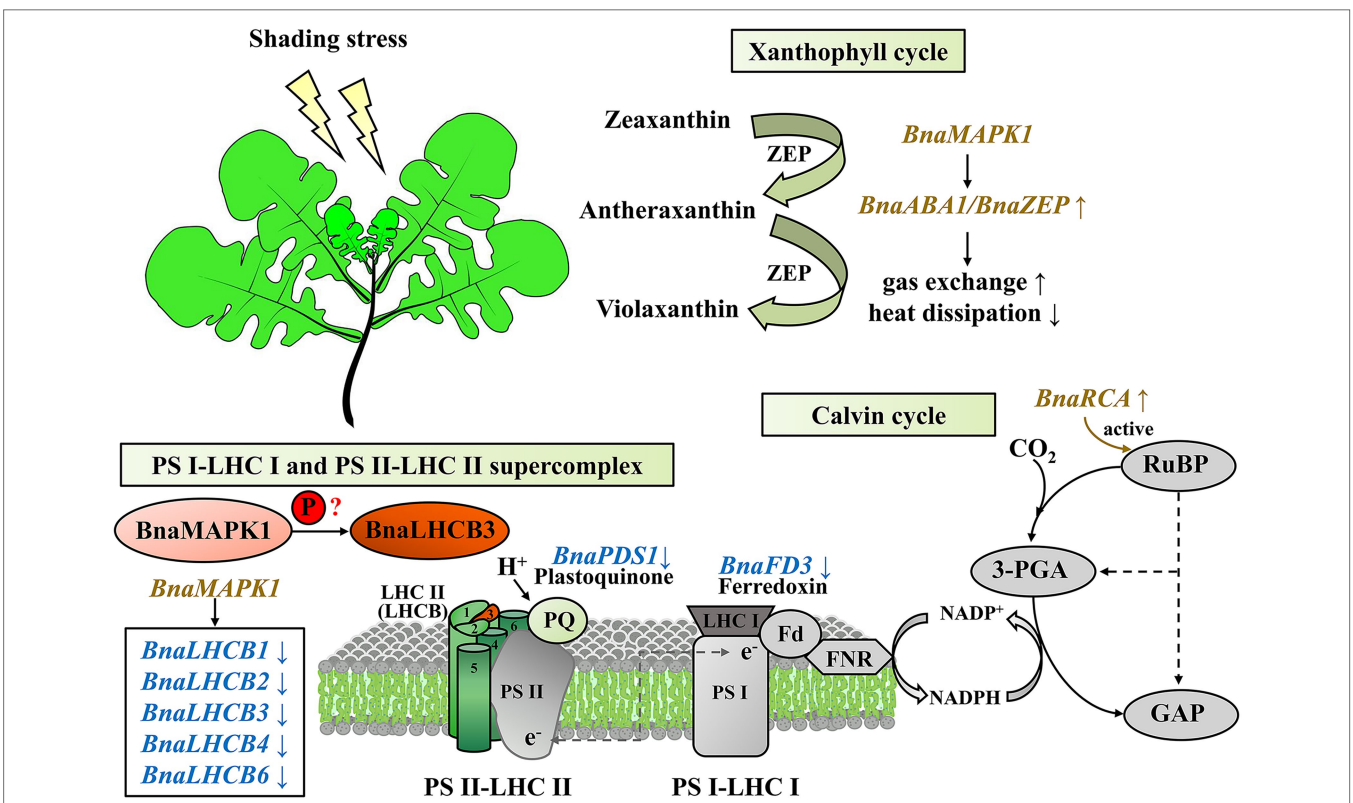


FIGURE 7 | A proposed working model of *BnaMAPK1* functions in photosynthesis under shading stress in *Brassica napus*.

AUTHOR CONTRIBUTIONS

YL, LW, KL, and JL designed the research and supervised the project. ZW, ML, MY, and XZ performed the experiments. ZW, CQ, and HD analyzed the data. ZW and ML drafted the manuscript. ZW, LW, KL, and YL revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was funded by the National Natural Science Foundation of China (31872876 and 32101663), the China Postdoctoral Science Foundation (2021M692683), the Chongqing Postdoctoral Natural Science Foundation (cstc2021jcyj-bshX0234), the Chongqing Special Funding in Postdoctoral Scientific Research (2010010006157688), the 111 Project of China (B12006), and the Agriculture Research System of Ministry of Finance and Minister of Agriculture and Rural Affairs of China.

ACKNOWLEDGMENTS

We thank Xiupeng Mei for providing the Y2H vectors and strains, Junxing Lu for preparing the transgenic rapeseed, Xiaogai Wang for measuring light response curve, and Jiaming Yin for providing the Zhongyou821 doubled haploid line seeds.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.902989/full#supplementary-material

REFERENCES

Adamiec, M., Gibasiewicz, K., Luciński, R., Giera, W., Chełminiak, P., Szewczyk, S., et al. (2015). Excitation energy transfer and charge separation are affected in *Arabidopsis thaliana* mutants lacking light-harvesting chlorophyll a/b binding protein Lhcb3. *J. Photochem. Photobiol. B Biol.* 153, 423–428. doi: 10.1016/j.jphotobiol.2015.11.002

Ballottari, M., Girardon, J., Dall'Osto, L., and Bassi, R. (2012). Evolution and functional properties of photosystem II light harvesting complexes in eukaryotes. *Biochim. Biophys. Acta Bioenerg.* 1817, 143–157. doi: 10.1016/j.bbabio.2011.06.005

Bracher, A., Whitney, S. M., Hartl, F. U., and Hayer-Hartl, M. (2017). Biogenesis and metabolic maintenance of Rubisco. *Annu. Rev. Plant Biol.* 68, 29–60. doi: 10.1146/annurev-arplant-043015-111633

Brunel-Muguet, S., Beauclair, P., Bataillé, M. P., Avice, J. C., Trouverie, J., Etienne, P., et al. (2013). Light restriction delays leaf senescence in winter oilseed rape (*Brassica napus* L.). *J. Plant Growth Regul.* 32, 506–518. doi: 10.1007/s00344-013-9317-9

Caffarri, S., Kouřil, R., Kereïche, S., Boekema, E. J., and Croce, R. (2009). Functional architecture of higher plant photosystem II supercomplexes. *EMBO J.* 28, 3052–3063. doi: 10.1038/emboj.2009.232

Confortin, T. C., Todero, I., Luft, L., Ugalde, G. A., Mazutti, M. A., Oliveira, Z. B., et al. (2019). Oil yields, protein contents, and cost of manufacturing of oil obtained from different hybrids and sowing dates of canola. *J. Environ. Chem. Eng.* 7:102972. doi: 10.1016/j.jece.2019.102972

Crepin, A., and Caffarri, S. (2018). Functions and evolution of Lhcb isoforms composing LHCII, the major light harvesting complex of photosystem II of green eukaryotic organisms. *Curr. Protein Pept. Sci.* 19, 699–713. doi: 10.2174/1389203719666180222101534

Cristina, M., Petersen, M., and Mundy, J. (2010). Mitogen-activated protein kinase signaling in plants. *Annu. Rev. Plant Biol.* 61, 621–649. doi: 10.1146/annurev-arplant-042809-112252

Danquah, A., De Zélicourt, A., Boudsocq, M., Neubauer, J., Frei Dit Frey, N., Leonhardt, N., et al. (2015). Identification and characterization of an ABA-activated MAP kinase cascade in *Arabidopsis thaliana*. *Plant J.* 82, 232–244. doi: 10.1111/tpj.12808

Elahi, N., Duncan, R. W., and Stasolla, C. (2016). Modification of oil and glucosinolate content in canola seeds with altered expression of *Brassica napus LEAFY COTYLEDON1*. *Plant Physiol. Biochem.* 100, 52–63. doi: 10.1016/j.plaphy.2015.12.022

Fan, Y., Chen, J., Wang, Z., Tan, T., Li, S., Li, J., et al. (2019). Soybean (*Glycine max* L. Merr.) seedlings response to shading: leaf structure, photosynthesis and proteomic analysis. *BMC Plant Biol.* 19:34. doi: 10.1186/s12870-019-1633-1

Gao, S., Liu, X., Liu, Y., Cao, B., Chen, Z., and Xu, K. (2020). Photosynthetic characteristics and chloroplast ultrastructure of welsh onion (Allium fistulosum L.) grown under different LED wavelengths. *BMC Plant Biol.* 20:78. doi: 10.1186/s12870-020-2282-0

He, H., Lei, Y., Yi, Z., Raza, A., Zeng, L., Yan, L., et al. (2021). Study on the mechanism of exogenous serotonin improving cold tolerance of rapeseed (*Brassica napus* L.) seedlings. *Plant Growth Regul.* 94, 161–170. doi: 10.1007/s10725-021-00700-0

Heddad, M., and Adamska, I. (2002). The evolution of light stress proteins in photosynthetic organisms. *Comp. Funct. Genomics* 3, 504–510. doi: 10.1002/cfg.221

Hettenhausen, C., Baldwin, I. T., and Wu, J. (2012). Silencing *MPK4* in *Nicotiana attenuata* enhances photosynthesis and seed production but compromises abscisic acid-induced stomatal closure and guard cell-mediated resistance to *Pseudomonas syringae* pv tomato DC3000. *Plant Physiol.* 158, 759–776. doi: 10.1104/pp.111.190074

Hieber, A. D., Bugos, R. C., and Yamamoto, H. Y. (2000). Plant lipocalins: Violaxanthin de-epoxidase and zeaxanthin epoxidase. *Biochim. Biophys. Acta Protein Struct. Mol. Enzymol.* 1482, 84–91. doi: 10.1016/S0167-4838(00)00141-2

Hou, Y., Zhu, Y., Wang, P., Zhao, Y., Xie, S., Batelli, G., et al. (2016). Type one protein phosphatase 1 (TOPP1) and its regulatory protein inhibitor 2 (AtI-2) negatively regulate ABA signaling. *PLoS Genet.* 12:e1005835. doi: 10.1371/journal.pgen.1005835

Ichimura, K., Shinozaki, K., Tena, G., Sheen, J., Henry, Y., Champion, A., et al. (2002). Mitogen-activated protein kinase cascades in plants: A new nomenclature. *Trends Plant Sci.* 7, 301–308. doi: 10.1016/S1360-1385(02)02302-6

Jonak, C., Ökrész, L., Bögre, L., and Hirt, H. (2002). Complexity, cross talk and integration of plant MAP kinase signalling. *Curr. Opin. Plant Biol.* 5, 415–424. doi: 10.1016/S1369-5266(02)00285-6

Kim, D., Langmead, B., and Salzbergl, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements Daehwan HHS public access. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317.HISAT

Kirilovsky, D., and Büchel, C. (2019). Evolution and function of light-harvesting antenna in oxygenic photosynthesis. *Adv. Bot. Res.* 91, 247–293. doi: 10.1016/bs.abr.2019.01.002

Klimmek, F., Ganeteg, U., Ihalainen, J. A., Van Roon, H., Jensen, P. E., Scheller, H. V., et al. (2005). Structure of the higher plant light harvesting complex I: In *vivo* characterization and structural interdependence of the Lhca proteins. *Biochemistry* 44, 3065–3073. doi: 10.1021/bi047873g

Klimmek, F., Sjödin, A., Noutsos, C., Leister, D., and Jansson, S. (2006). Abundantly and rarely expressed *Lhc* protein genes exhibit distinct regulation patterns in plants. *Plant Physiol.* 140, 793–804. doi: 10.1104/pp.105.073304.1

Lee, J. W., Lee, S. H., Han, J. W., and Kim, G. H. (2020). Early light-inducible protein (ELIP) can enhance resistance to cold-induced photooxidative stress in *Chlamydomonas reinhardtii*. *Front. Physiol.* 11:1083. doi: 10.3389/fphys.2020.01083

Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van De Peer, Y., et al. (2002). PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences. *Nucleic Acids Res.* 30, 325–327. doi: 10.1093/nar/30.1.325

Li, Y., Li, W., Hu, D., Shen, P., Zhang, G., and Zhu, Y. (2020). Comparative analysis of the metabolome and transcriptome between green and albino zones of variegated leaves from *Hydrangea macrophylla* "Maculata" infected by hydrangea ringspot virus. *Plant Physiol. Biochem.* 157, 195–210. doi: 10.1016/j.plaphy.2020.10.012

Li, Y. L., Liu, X. G., Hao, K., Yang, Q. L., Yang, X. Q., Zhang, W. H., et al. (2019). Light-response curve of photosynthesis and model fitting in leavesof mangifera indica under different soil water conditions. *Photosynthetica* 57, 796–803. doi: 10.32615/ps.2019.095

Li, H., Tang, C., and Xu, Z. (2013). The effects of different light qualities on rapeseed (*Brassica napus* L.) plantlet growth and morphogenesis in *vitro*. *Sci. Hortic.* 150, 117–124. doi: 10.1016/j.scienta.2012.10.009

Liang, Y., and Li, J. (2004). The varietal difference of tolerance to low light intensity in rape (*Brassica napus* L.) plants. *Acta Agron. Sin.* 30, 360–364.

Liu, J., Hua, W., Zhan, G., Wei, F., Wang, X., Liu, G., et al. (2010). Increasing seed mass and oil content in transgenic *Arabidopsis* by the overexpression of *wri1*-like gene from *Brassica napus*. *Plant Physiol. Biochem.* 48, 9–15. doi: 10.1016/j.plaphy.2009.09.007

Liu, Y., Pan, T., Tang, Y., Zhuang, Y., Liu, Z., Li, P., et al. (2020). Proteomic analysis of rice subjected to low light stress and overexpression of *OsGAPB* increases the stress tolerance. *Rice* 13:30. doi: 10.1186/s12284-020-00390-8

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Long, S. P., Marshall-Colon, A., and Zhu, X. G. (2015). Meeting the global food demand of the future by engineering crop photosynthesis and yield potential. *Cell* 161, 56–66. doi: 10.1016/j.cell.2015.03.019

Lu, Z., Pan, Y., Hu, W., Cong, R., Ren, T., Guo, S., et al. (2017). The photosynthetic and structural differences between leaves and siliques of *Brassica napus* exposed to potassium deficiency. *BMC Plant Biol.* 17:240. doi: 10.1186/s12870-017-1201-5

Mao, X., Cai, T., Olyarchuk, J. G., and Wei, L. (2005). Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics* 21, 3787–3793. doi: 10.1093/bioinformatics/bti430

Mikko, T., Mirva, P., Marjaana, S., Sari, S., Paula, M., Julia, V., et al. (2006). State transitions revisited - a buffering system for dynamic low light acclimation of *Arabidopsis*. *Plant Mol. Biol.* 62, 779–793. doi: 10.1007/s11103-006-9044-8

Mitra, S., and Baldwin, I. T. (2008). Independently silencing two photosynthetic proteins in *Nicotiana attenuata* has different effects on herbivore resistance. *Plant Physiol.* 148, 1128–1138. doi: 10.1104/pp.108.124354

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* 5, 621–628. doi: 10.1038/nmeth.1226

Munekage, Y., Hashimoto, M., Miyake, C., Tomizawa, K. I., Endo, T., Tasaka, M., et al. (2004). Cyclic electron flow around photosystem I is essential for photosynthesis. *Nature* 429, 579–582. doi: 10.1038/nature02598

Neoh, B. K., Wong, Y. C., Teh, H. F., Mei Ng, T. L., Tiong, S. H., Ooi, T. E. K., et al. (2019). Diurnal biomarkers reveal key photosynthetic genes associated with increased oil palm yield. *PLoS One* 14:e0213591. doi: 10.1371/journal.pone.0213591

Niinemets, Ü. (2010). A review of light interception in plant stands from leaf to canopy in different plant functional types and in species with varying shade tolerance. *Ecol. Res.* 25, 693–714. doi: 10.1007/s11284-010-0712-4

Ortiz-Masia, D., Perez-Amador, M. A., Carbonell, P., Aniento, F., Carbonell, J., and Marcote, M. J. (2008). Characterization of PsMPK2, the first C1 subgroup MAP kinase from pea (*Pisum sativum* L.). *Planta* 227, 1333–1342. doi: 10.1007/s00425-008-0705-5

Ortiz-Masia, D., Perez-Amador, M. A., Carbonell, J., and Marcote, M. J. (2007). Diverse stress signals activate the C1 subgroup MAP kinases of *Arabidopsis*. *FEBS Lett.* 581, 1834–1840. doi: 10.1016/j.febslet.2007.03.075

Pan, G., Liu, W., Zhang, H., and Liu, P. (2018). Morphophysiological responses and tolerance mechanisms of *Xanthium strumarium* to manganese stress. *Ecotoxicol. Environ. Saf.* 165, 654–661. doi: 10.1016/j.ecoenv.2018.08.107

Papadatos, S., Charalambous, A. C., and Daskalakis, V. (2017). A pathway for protective quenching in antenna proteins of photosystem II. *Sci. Rep.* 7:2523. doi: 10.1038/s41598-017-02892-w

Pitzschke, A., Schikora, A., and Hirt, H. (2009). MAPK cascade signalling networks in plant defence. *Curr. Opin. Plant Biol.* 12, 421–426. doi: 10.1016/j.pbi.2009.06.008

Portis, A. R. (2003). Rubisco activase - Rubisco's catalytic chaperone. *Photosynth. Res.* 75, 11–27. doi: 10.1023/A:1022458108678

Pribil, M., Pesaresi, P., Hertle, A., Barbato, R., and Leister, D. (2010). Role of plastid protein phosphatase TAP38 in LHCII dephosphorylation and thylakoid electron flow. *PLoS Biol.* 8:e1000288. doi: 10.1371/journal.pbio.1000288

Qin, G., Gu, H., Ma, L., Peng, Y., Deng, X. W., Chen, Z., et al. (2007). Disruption of phytoene desaturase gene results in albino and dwarf phenotypes in *Arabidopsis* by impairing chlorophyll, carotenoid, and gibberellin biosynthesis. *Cell Res.* 17, 471–482. doi: 10.1038/cr.2007.40

Qiyuan, T., and Ibrahim, B. (2016). Effect of low light stress at different phases of grain filling on rice seed germination and seed vigour. *J. Biol. Agric. Healthc.* 6, 72–82.

Rantala, M., Lehtimäki, N., Aro, E. M., and Suorsa, M. (2016). Downregulation of TAP38/PPH1 enables LHCII hyperphosphorylation in *Arabidopsis* mutant lacking LHCII docking site in PSI. *FEBS Lett.* 590, 787–794. doi: 10.1002/1873-3468.12117

Rantala, M., Rantala, S., and Aro, E. M. (2020). Composition, phosphorylation and dynamic organization of photosynthetic protein complexes in plant thylakoid membrane. *Photochem. Photobiol. Sci.* 19, 604–619. doi: 10.1039/d0pp00025f

Raza, A. (2021). Eco-physiological and biochemical responses of rapeseed (*Brassica napus* L.) to abiotic stresses: consequences and mitigation strategies. *J. Plant Growth Regul.* 40, 1368–1388. doi: 10.1007/s00344-020-10231-z

Raza, A., Hafeez, M. B., Zahra, N., Shaukat, K., Umbreen, S., Tabassum, J., et al. (2020). "The plant family *Brassicaceae*: Introduction, biology, and importance," in *The Plant Family Brassicaceae*. ed. M. Hasanuzzaman (Singapore: Springer, Singapore), 1–43.

Ruberti, I., Sessa, G., Ciolfi, A., Possenti, M., Carabelli, M., and Morelli, G. (2012). Plant adaptation to dynamically changing environment: The shade avoidance response. *Biotechnol. Adv.* 30, 1047–1058. doi: 10.1016/j.biotechadv.2011.08.014

Shariatipour, N., and Heidari, B. (2020). Meta-analysis of expression of the stress tolerance associated genes and uncover their cis-regulatory elements in rice (*Oryza sativa* L.). *Open Bioinform. J.* 13, 39–49. doi: 10.2174/1875036202013010039

Shengxin, C., Chunxia, L., Xuyang, Y., Song, C., Xuelei, J., Xiaoying, L., et al. (2016). Morphological, photosynthetic, and physiological responses of rapeseed leaf to different combinations of red and blue lights at the rosette stage. *Front. Plant Sci.* 7:1144. doi: 10.3389/fpls.2016.01144

Sims, D. A., and Gamon, J. A. (2002). Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sens. Environ.* 81, 337–354. doi: 10.1016/S0034-4257(02)00010-X

Standfuss, J., and Kühlbrandt, W. (2004). The three isoforms of the light-harvesting complex II: spectroscopic features, trimer formation, and functional roles. *J. Biol. Chem.* 279, 36884–36891. doi: 10.1074/jbc.M402348200

Stress, P., Rossini, S., Casazza, A. P., Engelmann, E. C. M., Havaux, M., Jennings, R. C., et al. (2006). Suppression of both ELIP1 and ELIP2 in *Arabidopsis* does not affect tolerance to photoinhibition and photooxidative stress. *Plant Physiol.* 141, 1264–1273. doi: 10.1104/pp.106.083055.the

Su, J., Yang, L., Zhu, Q., Wu, H., He, Y., Liu, Y., et al. (2018). Active photosynthetic inhibition mediated by MPK3/MPK6 is critical to effector-triggered immunity. *PLoS Biol.* 16:e2004122. doi: 10.1371/journal.pbio.2004122

Suzuki, Y., Ishiyama, K., Sugawara, M., Suzuki, Y., Kondo, E., Takegahara-Tamakawa, Y., et al. (2021). Overproduction of chloroplast Glyceraldehyde-3-phosphate dehydrogenase improves photosynthesis slightly under elevated [CO_2] conditions in rice. *Plant Cell Physiol.* 62, 156–165. doi: 10.1093/pcp/pcaa149

Teige, M., Scheikl, E., Eulgem, T., Dóczi, R., Ichimura, K., Shinozaki, K., et al. (2004). The MKK2 pathway mediates cold and salt stress signaling in *Arabidopsis*. *Mol. Cell* 15, 141–152. doi: 10.1016/j.molcel.2004.06.023

Tian, Y. N., Zhong, R. H., Wei, J. B., Luo, H. H., Eyal, Y., and Jin, H. L., et al. (2021). *Arabidopsis* CHLOROPHYLLASE 1 protects young leaves from long-term photodamage by facilitating FtsH-mediated D1 degradation in photosystem II repair. *Mol. Plant*. 14, 1149–1167. doi:10.1016/j.molp.2021.04.006.

Tikkanen, M., and Aro, E. M. (2012). Thylakoid protein phosphorylation in dynamic regulation of photosystem II in higher plants. *Biochim. Biophys. Acta Bioenerg.* 1817, 232–238. doi: 10.1016/j.bbabio.2011.05.005

Velez-Ramirez, A. I., Van Ieperen, W., Vreugdenhil, D., Van Poppel, P. M. J. A., Heuvelink, E., and Millenaar, F. F. (2014). A single locus confers tolerance to continuous light and allows substantial yield increase in tomato. *Nat. Commun.* 5:4549. doi: 10.1038/ncomms5549

Vialet-Chabrand, S., Matthews, J. S. A., Simkin, A. J., Raines, C. A., and Lawson, T. (2017). Importance of fluctuations in light on plant photosynthetic acclimation. *Plant Physiol.* 173, 2163–2179. doi: 10.1104/pp.16.01767

Voss, I., Goss, T., Murozuka, E., Altmann, B., McLean, K. J., Rigby, S. E. J., et al. (2011). FdC1, a novel ferredoxin protein capable of alternative electron partitioning, increases in conditions of acceptor limitation at photosystem I. *J. Biol. Chem.* 286, 50–59. doi: 10.1074/jbc.M110.161562

Walters, R. G. (2005). Towards an understanding of photosynthetic acclimation. *J. Exp. Bot.* 56, 435–447. doi: 10.1093/jxb/eri060

Wang, L., Feng, Z., Wang, X., Wang, X., and Zhang, X. (2009). DEGseq: An R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* 26, 136–138. doi: 10.1093/bioinformatics/btp612

Wang, C., Hai, J., Yang, J., Tian, J., Chen, W., Chen, T., et al. (2016). Influence of leaf and silique photosynthesis on seeds yield and seeds oil quality of oilseed rape (*Brassica napus* L.). *Eur. J. Agron.* 74, 112–118. doi: 10.1016/j.eja.2015.12.008

Wang, Z., Yao, M. N., Zhang, X. L., Qu, C. M., Lu, K., Li, J. N., et al. (2020). Prokaryotic expression, subcellular localization and yeast two-hybrid library screening of BnMAPK1 in *Brassica napus*. *Acta Agron. Sin.* 46, 1312–1321. doi: 10.3724/SP.J.1006.2020.04019

Weng, C. M., Lu, J. X., Wan, H. F., Wang, S. W., Wang, Z., Lu, K., et al. (2014). Over-expression of *BnMAPK1* in *Brassica napus* enhances tolerance to drought stress. *J. Integr. Agric.* 13, 2407–2415. doi: 10.1016/S2095-3119(13)60696-6

Xu, C., He, C. G., Wang, Y. J., Bi, Y. F., and Jiang, H. (2020). Effect of drought and heat stresses on photosynthesis, pigments, and xanthophyll cycle in alfalfa (*Medicago sativa* L.). *Photosynthetica* 58, 1226–1236. doi: 10.32615/ps.2020.073

Xu, Y. H., Liu, R., Yan, L., Liu, Z. Q., Jiang, S. C., Shen, Y. Y., et al. (2012). Light-harvesting chlorophyll a/b-binding proteins are required for stomatal response to abscisic acid in *Arabidopsis*. *J. Exp. Bot.* 63, 1095–1106. doi: 10.1093/jxb/err315

Ye, Z. P., Suggett, D. J., Robakowski, P., and Kang, H. J. (2013). A mechanistic model for the photosynthesis-light response based on the photosynthetic electron transport of photosystem II in C3 and C4 species. *New Phytol.* 199, 110–120. doi: 10.1111/nph.12242

Young, M. D., Wakefield, M. J., Smyth, G. K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 11:R14. doi: 10.1186/gb-2010-11-2-r14

Zafar, S., Li, Y. L., Li, N. N., Zhu, K. M., and Tan, X. L. (2019). Recent advances in enhancement of oil content in oilseed crops. *J. Biotechnol.* 301, 35–44. doi: 10.1016/j.jbiotec.2019.05.307

Zhang, M., Su, J., Zhang, Y., Xu, J., and Zhang, S. (2018). Conveying endogenous and exogenous signals: MAPK cascades in plant growth and defense. *Curr. Opin. Plant Biol.* 45, 1–10. doi: 10.1016/j.pbi.2018.04.012

Zhao, X., Li, Y., Zheng, M., Bian, X., Liu, M., Sun, Y., et al. (2015). Comparative analysis of growth and photosynthetic characteristics of (*Populus simonii* × *P. nigra*) × (*P. nigra* × *P. simonii*) hybrid clones of different ploidides. *PLoS One* 10:e0119259. doi: 10.1371/journal.pone.0119259

Zhu, X. G., Long, S. P., and Ort, D. R. (2010). Improving photosynthetic efficiency for greater yield. *Annu. Rev. Plant Biol.* 61, 235–261. doi: 10.1146/annurev-arplant-042809-112206

Zhu, Y., Wang, B., Tang, K., Hsu, C. C., Xie, S., Du, H., et al. (2017). An *Arabidopsis* Nucleoporin NUP85 modulates plant responses to ABA and salt stress. *PLoS Genet.* 13:e1007124. doi: 10.1371/journal.pgen.1007124

Zivcak, M., Brestic, M., and Kalaji, H. M.Govindjee (2014). Photosynthetic responses of sun- and shade-grown barley leaves to high light: is the lower PSII connectivity in shade leaves associated with protection against excess of light? *Photosynth. Res.* 119, 339–354. doi: 10.1007/s11120-014-9969-8

Zoratti, L., Karppinen, K., Escobar, A. L., Häggman, H., and Jaakola, L. (2014). Light-controlled flavonoid biosynthesis in fruits. *Front. Plant Sci.* 5:534. doi: 10.3389/fpls.2014.00534

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 02 June 2022
doi: 10.3389/fpls.2022.893896

HbMYB44, a Rubber Tree MYB Transcription Factor With Versatile Functions in Modulating Multiple Phytohormone Signaling and Abiotic Stress Responses

Bi Qin[1†], Song-Le Fan[2†], Hai-Yang Yu[2†], Yan-Xi Lu[1] and Li-Feng Wang[1]*

[1] *Key Laboratory of Biology and Genetic Resources of Rubber Tree, Ministry of Agriculture and Rural Affairs, Rubber Research Institute, Chinese Academy of Tropical Agricultural Sciences, Haikou, China,* [2] *Institute of Tropical Crops, Hainan University, Haikou, China*

OPEN ACCESS

Edited by:
Prasanta Kumar Subudhi, Louisiana State University, United States

Reviewed by:
Deepanker Yadav, Guru Ghasidas Vishwavidyalaya, India
Peng Zhao, Northwest University, China

***Correspondence:**
Li-Feng Wang
lfwang@catas.cn

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

Received: *11 March 2022*
Accepted: *09 May 2022*
Published: *02 June 2022*

Citation:
Qin B, Fan S-L, Yu H-Y, Lu Y-X and Wang L-F (2022) HbMYB44, a Rubber Tree MYB Transcription Factor With Versatile Functions in Modulating Multiple Phytohormone Signaling and Abiotic Stress Responses. Front. Plant Sci. 13:893896. doi: 10.3389/fpls.2022.893896

The vital roles of R2R3-MYB transcription factors (TFs) in regulating stress response and phytohormone signaling have been thoroughly studied in numerous plant species, but the functions of these TFs in rubber tree are poorly understood. Rubber tree is the most important source of natural rubber but often suffers from various abiotic and biotic stresses that cause severe yield losses each year. In this study, we reported a novel *MYB44* gene in rubber tree (named *HbMYB44*) and revealed its biological function. HbMYB44 was highly similar to AtMYB44 and clustered into subgroup 22. Transient expression indicated that HbMYB44 is a nuclear localized protein and displays transactivation activity at the C-terminus. *HbMYB44* was ubiquitously expressed in rubber tree, and its expression was strongly induced by multiple phytohormones, drought stress, wounding, and H_2O_2 treatments. Furthermore, overexpression of *HbMYB44* in *Arabidopsis* (OE) demonstrated that OE plants significantly enhanced stress tolerance, i.e., salt stress, osmotic stress, and drought stress. Additionally, *HbMYB44* promoted recovery from root growth inhibition of OE plants caused by exogenous phytohormones (including abscisic acid, methyl jasmonic acid, gibberellic acid 3, and salicylic acid), but the opposite effect was present in response to ethephon. Interestingly, *HbMYB44* increased the expression of its homologous genes and interacting protein-encoding genes in OE plants. Overall, *HbMYB44* plays versatile functions in modulating multiple phytohormone signaling pathways and stress tolerance.

Keywords: rubber tree, MYB transcription factor, HbMYB44, phytohormone signaling, stress tolerance

Abbreviations: aa, amino acids; ABA, abscisic acid; DAPI, 4',6-diamidino-2-phenylindole; DNA-BD, DNA-binding domain; ET, ethephon; GA_3, gibberellic acid 3; GFP, green fluorescent protein; MeJA, methyl jasmonic acid; NR, natural rubber; PEG, polyethylene glycol; SA, salicylic acid; TDO, SD/-Trp/-His/-Ade; TF, transcription factor; TSA, transcriptome shotgun assembly; X-α-Gal, 5-bromo-4-chloro-3-indolyl-α-d-galactopyranoside.

INTRODUCTION

The first MYB transcription factor (TF) was identified as a protein encoded by a viral oncogene (*v-MYB*) in birds (Wang X. et al., 2018). In plants, MYB TFs compose one of the most important families of TFs and regulate various plant development processes and multiple stress responses (Ampomah-Dwamena et al., 2019). MYB TFs can be classified into four groups according to their repeats (R) with 50–53 amino acid residues: 1R-, R2R3-, 3R-, and 4R-MYBs (Zhang T. et al., 2018). R2R3-MYB TFs play important roles in regulating plant development, abiotic and biotic stress responses, etc. For example, the kiwifruit *MYB7* modulates chlorophyll and carotenoid accumulation by activating the promoter of *lycopene beta-cyclase* (*AdLCY*-β) (Ampomah-Dwamena et al., 2019). Overexpression of wild soybean *GsMYB1* enhances resistance to salt stress and *Helicoverpa armigera* in transgenic *Arabidopsis* (Shen et al., 2018). Recently, advances in genome and transcriptome sequencing techniques have facilitated the identification of the MYB family in non-model plant species, such as peach (Zhang C. et al., 2018) and *Pleurotus ostreatus* (Wang L. et al., 2018).

The *Arabidopsis* R2R3-MYB family can divide into 25 subgroups based on the conserved motifs present in the C-terminus to the MYB domain. AtMYB44, together with AtMYB70, AtMYB73, and AtMYB77, belongs to subgroup 22, which contains two conserved motifs, namely, 22.1 (TGLYMSPxSP) and 22.2 (GxFMxVVQEMIxxEVRSYM) (Stracke et al., 2001). The MYBs of subgroup 22 play essential roles in regulating plant stress response and phytohormone signaling. For instance, *AtMYB73* expression was found to be upregulated under salt stress conditions in *Arabidopsis*, and it negatively regulated *SOS1* and *SOS3* transcripts (Kim et al., 2013). AtMYB44 confers abiotic stress tolerance and non-specific phytohormonal induction in *Arabidopsis* (Jung et al., 2008, 2010). Moreover, by activating *EIN2* expression, AtMYB44 functions in various biotic stress responses, such as resistance to *Myzus persicae* Sulzer and *Plutella xylostella* L. (Lu et al., 2013). On the one hand, AtMYB44 regulates salicylic acid (SA)- and jasmonic acid (JA)-mediated defense responses by directly regulating *WRKY70* expression; thus, this TF acts as an integrator of cross-talk between SA and JA in plant defense responses (Shim and Choi, 2013; Shim et al., 2013). On the other hand, AtMYB44 interacts with the abscisic acid (ABA) receptors PYL9 and PYL8 and regulates ABA signaling (Shim et al., 2013; Li et al., 2014; Zhao et al., 2014). Additionally, AtMYB44 is regulated by MPK3-mediated phosphorylation (Persak and Pitzschke, 2013). Recently, chromatin immunoprecipitation showed that several features present in the promoter region of *AtMYB44* facilitate the binding of different TFs involving various signals, including a low density of nucleosomes, no DNA methylation, and multiple TF-binding elements (Nguyen et al., 2012; Nguyen and Cheong, 2018a,b). Although the function of *MYB44* has been reported in the model plant species *Arabidopsis*, its function in non-model plant species is poorly understood. To improve our knowledge of its transcriptional regulation mechanism, the functions of *MYB44* in different species need to be characterized.

Rubber tree (*Hevea brasiliensis* Müll. Arg) is the most important plant species for the production of natural rubber (NR) (Puskas et al., 2006). Latex within NR is produced in the cytoplasm of laticifer cells and harvested by regularly tapping the bark of rubber trees. Ethephon (ET) is widely used in NR production to increase the latex yield. Rubber trees often suffer from abiotic stress, wounding (caused by frequent tapping), drought, low temperature, and biotic stress, especially the destructive disease South American leaf blight (SALB), which is still a serious threat to plantations. From a biological point of view, NR is important for disease and insect defense and high-temperature stress responses in rubber trees (Sharkey and Yeh, 2001). Although important roles of MYBs have been revealed in plants, few MYB TFs in rubber tree have been characterized. *HbMyb1* can suppress cell death induced by stress conditions in tobacco (Peng et al., 2011). The expression of *HbSM1* (a 1R-MYB) was shown to be induced by phytohormones and wounding (Qin et al., 2014). Recently, 44 *MYBs* (*HblMYB1-HblMYB44*) were identified from the laticifer transcriptome database and the genome database of rubber tree. Among them, HblMYB44 (belonging to a new group G3) and HblMYB19 (belonging to a new group G2) were shown to bind to the promoters of *HRT1*, *FDPS1*, and *SRPP* in yeast and increase their expression in transient expression assays of tobacco (Wang et al., 2017). Until now, the function of MYBs of subgroup 22 has remained unknown in rubber tree. Due to the important functions of MYB44 in the plant stress response and phytohormone signaling, we assumed that MYB44 homologs may play a vital role in the stress response in rubber tree. In this study, *HbMYB44* was cloned, and its function was characterized in rubber tree and transgenic *Arabidopsis*. These results will lay a good foundation for elucidating the transcriptional regulatory mechanism underlying the stress response in rubber tree.

MATERIALS AND METHODS

Plant Materials and Treatments

The rubber tree clone CATAS7-33-97 was planted at the experimental farm of Danzhou, Hainan Province, China. At the flowering stage, samples of leaves, bark, latex, and flowers were collected from the same plants simultaneously. Tissue culture-generated seedlings were used for different stress and phytohormone treatments according to our reported method (Qin et al., 2015). Briefly, plants were subjected to drought stress treatment by continuously withholding water for 10 days, and the control plants were well-watered, with samples collected at the same time points. Wounds were made on the leaves by pinching with forceps, and untreated plants were used as controls. For chemical treatments, whole plants were sprayed with 2% (v/v) H_2O_2, 1.0% (v/v) ET, 200 mM methyl jasmonate (MeJA), 3 mM gibberellic acid 3 (GA_3), 200 μM ABA, and 5 mM SA, respectively. All the chemicals were diluted in distilled water that included 0.05% (v/v) ethanol, and the control plants were sprayed with distilled water that included 0.05% (v/v) ethanol. Six plants were included in each treatment, and

leaves from the six plants were harvested at each time point. Three different biological replicates were analyzed. The harvested samples were immediately frozen in liquid nitrogen and stored at −80°C until analysis.

Cloning of *HbMYB44* and Determining Its Structural Features and Phylogenetic Analysis

The AtMYB44 (AT5G67300) sequence was subjected to tBLASTn searches of the transcriptome shotgun assembly (TSA) and genome database of *H. brasiliensis*. According to the obtained sequence, primers (MYB44-F1 and MYB44-R1, **Table 1**) were designed to amplify the full-length open reading frame (ORF) of *HbMYB44* from the rubber tree.

The National Center for Biotechnology Information (NCBI) conserved domain searches[1] and SMART[2] were used to analyze the structure of HbMYB44. Sequences of *HblMYBs* identified from the laticifer transcriptome database and the genome database of *H. brasiliensis* (Wang et al., 2017), and *Arabidopsis MYBs* obtained from the phytozome,[3] were used for the phylogenetic analysis. Information on *H. brasiliensis* and *Arabidopsis MYBs* is listed in **Supplementary Table 1**. Sequence alignments were conducted using the DNAMAN version 8.0 and ClustalX (Thompson et al., 2002). The phylogenetic analysis was performed using the MEGA version X based on the neighbor-joining method with a Poisson correction model and a bootstrap test of 1,000 replicates (Tamura et al., 2013).

RNA Isolation and Analysis of Gene Expression

Total RNA was extracted from different tissues using a Total RNA Isolation kit (OMEGA Bio-tek, Norcross, GA, United States). cDNA was synthesized with ReverTra Ace qPCR RT Master Mix with gDNA Remover (TOYOBO, Shanghai, China) according to the manufacturer's instructions. Quantitative real-time PCR (qRT-PCR) was performed on a CFX96^{TM} Real-Time System (Bio-Rad Laboratories, Hercules, CA, United States) using the AceQ Universal SYBR qPCR Master Mix (Vazyme, Nanjing, China) according to the manufacturers' instructions. Specific primers MYB44-F2 and MYB44-R2 were used for amplifying *HbMYB44* (**Table 1**). *HbActin* (GenBank accession number: HO004792) and *AtActin2* (GenBank accession number: NM_001338358.1) were used as housekeeping genes in the rubber tree and *Arabidopsis*, respectively. The relative expression levels of target genes were calculated using the $2^{-\Delta\Delta CT}$ method (Nolan et al., 2006; Schmittgen and Livak, 2008). All the data were analyzed by one-way ANOVA, and multiple comparisons were performed by the Tukey's test at the $P < 0.01$ level. Figures were constructed by OriginPro 2018 (OriginLab Corporation, Northampton, MA, United States).

[1]http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi?

[2]http://smart.embl-heidelberg.de/

[3]https://phytozome.jgi.doe.gov/pz/portal.html

Subcellular Localization of HbMYB44

To determine the subcellular localization of HbMYB44, a fluorescent GFP tag was fused to the C-terminus of HbMYB44. The ORF of *HbMYB44* was amplified with primers (MYB44-EcoR I-F1 and MYB44-BamH I-R1, **Table 1**) and subsequently inserted in the EcoR I/BamH I restriction sites of the vector 35S: XXGFP (provided by Dr. Xingliang Hou). The recombinant plasmid *35S:HbMYB44:mGFP* was then transformed into *Agrobacterium tumefaciens* strain GV3101, and then, the GV3101 strain harboring the recombinant plasmid was transiently expressed in tobacco (*Nicotiana benthamiana*) leaf epidermal cells according to the methods of our previous study (Qin et al., 2019). The fluorescent protein expression was examined by fluorescence imaging together with confocal microscopy (Leica, Wetzlar, Germany) after 1 day of dark/light cultivation post infiltration. The nuclei were stained with 4′,6-diamidino-2-phenylindole (DAPI) as previously reported (Kapuscinski, 1995) and visualized by confocal laser scanning microscopy.

Transactivation Analysis in Yeast

The complete ORF of *HbMYB44* (amplified by MYBY44-EcoR I-F1 and MYBY44-BamH I-R2) and fragments of *HbMYB44* encoding the N-terminus [amino acids (aa) 1-200, amplified by MYBY44-EcoR I-F1 and MYBY44-▲N-BamH I-R3] or the C-terminus (aa 201-345, amplified by MYBY44-▲C-EcoR I-F2 and MYBY44-BamH I-R2), were amplified and cloned into the EcoR I/BamH II restriction sites of the vector pGBKT7,[4] respectively. The target fragment was fused to the 3′ end of the GAL4 DNA-binding domain (DNA-BD) of the vector pGBKT7 between the EcoR I and BamH I restriction sites of MCS, such that a GAL4 DNA-BD-*HbMYB44* fusion protein would be produced when the yeast cells were transformed. The pGBKT7-HbMYB44, pGBKT7-▲N-HbMYB44, and pGBKT7-▲ C-HbMYB44 constructs; the negative control pGBKT7; and the positive control pGBKT7-HbNAC24 were used (HbNAC24 is an NAC TF of the rubber tree and has been shown to have transcriptional activation activity) (Kang et al., 2014). These constructs were individually transformed into yeast strain Y2H Gold using a Yeastmaker^{TM} Yeast Transformation System 2 (Clontech, Mountain View, CA, United States) according to the manufacturer's instructions. If the fusion protein had transactivation activity, the colonies containing the pGBKT7+fusion protein construct in Y2H Gold would grow on SD/-Trp/-His/-Ade (TDO) media and become blue on TDO supplemented with X-α-Gal (TDO/X) media. All the yeast transformants were grown at 30°C for 2–3 days on SD media lacking tryptophan (SD/-Trp). Then, the positive clones were transferred to TDO media or TDO/X media and cultured at 30°C for 2–3 days (Fujita et al., 2004; Zhang et al., 2014). The experiment was repeated three times.

Plasmid Construction and Transformation of *Arabidopsis*

GV3101 cells harboring the recombinant plasmid 35S:HbMYB44:mGFP were used in *Arabidopsis* transformation

[4]www.clontech.com

TABLE 1 | Sequences of the primers used in this study.

Name	Sequence (5′ to 3′)	Name	Sequence (5′ to 3′)
MYB44-F1	GCCAAGCTTCCATTTTATGATTACT	MYB44-R1	CTCTGTACAATGCCCAAAATTATTC
MYB44-F2	TTAAAGAGATCGGTGAGTGC	MYB44-R2	CTGTACACATGCGAAGAAGA
MYB44-EcoRI-F1	CGGAATTCATGGCTGTTACTGGGAA	MYB44-BamHI-R1	CGGGATCCCTCAATCTTACTAATC
MYB44-▲C-EcoRI-F2	CGGAATTCTCTAATAACAACGACC	MYB44-BamHI-R2	CGGGATCCCTACTCAATCTTACTAA
MYB44-▲N-BamHI-R3	CGGGATCCAGACGACGTCGTTTCTA		
HbActin-F	GATGTGGATATCAGGAAGG	HbActin-R	CATACTGCTTGGAGCAAGA
AtMYB44-F	AAGCCCAACTGGATCTGATG	AtMYB44-R	GGTGGATCATCGGAAGAAGA
AtMYB70-F	GTAACAAGTGGGCGACGATT	AtMYB70-R	ACTCTGCCCTTCCTCTCCTC
AtMYB73-F	TGGGCTACGATCTCTCGTCT	AtMYB73-R	TCAACGGTTGCTCTTCTCCT
AtMYB77-F	GCGTTGATGTTTCCGAGATT	AtMYB77-R	TTTCCGCCATGTAACTCCTC
AtPYL8-F	CATCTCTCTTCACCCCGAGA	AtPYL8-R	ACCGCAAGACGTTCAGAGAT
AtPYL9-F	TACGGACGCATCATCAACAT	AtPYL9-R	ACTGCCGATTTCAGGATCAC
AtWRKY70-F	GAGGACGCATTTTCTTGGAG	AtWRKY70-R	GCTCAACCTTCTGGACTTGC
AtActin2-F	GATCTCCAAGGCCGAGTATGAT	AtActin2-R	CCCATTCATAAAACCCCAGC

The restriction endonuclease site was underlined.

for the overexpression of *HbMYB44* (OE). *Arabidopsis thaliana* Col-0 wild type (WT) was transformed *via* the floral-dip method as previously described (Guan et al., 2014). All the transgenic plants were selected on Murashige and Skoog (MS) media that included 10 mg L^{-1} Basta. Then, the resistant plants of the T2 generation were tested with the primers MYB44-EcoR I-F1 and MYB44-BamH I-R1 of *HbMYB44* and used for the following studies.

The expression of *HbMYB44*, its homologous genes (including *AtMYB44*, *AtMYB73*, *AtMYB70*, and *AtMYB77*), and genes (including *AtWRKY70*, *AtPYL8*, and *AtPYL9*) encoding the interacting proteins of AtMYB44 in OE and WT plants were analyzed using qRT-PCR according to the method mentioned above. The primers used are listed in **Table 1**.

Phytohormone Response and Stress Tolerance Assays in Transgenic *Arabidopsis* Lines

The seeds of the T3 generation and WT seeds were sterilized with 10% bleach and 0.001% Triton X-100 for 10 min and washed five times with sterile water. The sterilized seeds were sown on MS media, stratified at 4°C in darkness for 2 days, and then transferred to a plant growth chamber (MMM/Climacell 707, Germany) set at 22°C with a 16 h light/8 h dark photoperiod (light intensity of 13,000 lux) for germination for 1–2 days. Then, the germinated seeds were transferred to MS media supplemented with 100 mM NaCl, 250 mM mannitol, 15% (w/v) PEG4000, 0.3 μM ET, 10 μM ABA, 20 μM MeJA, 50 μM GA_3, and 10 μM SA. The plates were placed vertically in the illumination incubator under the same conditions used for germination, and the root length of the seedlings was measured after 10–13 days of treatments. For each treatment, at least six plants were tested in triplicate, and three different biological replicates were performed. The final value was the average of three batches of measurement data.

RESULTS

HbMYB44 Is an R2R3-MYB Transcription Factor and Is Homologous to AtMYB44

The rubber tree *MYB44* homolog *HbMYB44* (GenBank accession number: AZZ09361.1) was cloned using RT-PCR. The ORF of *HbMYB44* is 1,038 bp in length and encodes a protein of 345 aa. No introns were found in the genomic sequence of *HbMYB44*. The deduced HbMYB44 protein contains 2 SANT domains (one is from the 11th to 60th aa, and the other is from the 63rd to 111th aa; **Figures 1A,B**), which indicated that HbMYB44 is an R2R3-MYB TF. A BLASTp search of the NCBI non-redundant (NR) protein database revealed that HbMYB44 was highly similar to MYB44 homologs from *Ricinus communis*, *Populus euphratica*, *Glycine max, and A. thaliana* with 72, 71, 61, and 55% identities, respectively (**Figure 1B**).

Multiple sequence alignments between rubber tree and *A. thaliana* MYB family members indicated that HbMYB44 shared the highest similarity with rubber tree HblMYB7 (with 82% identity) and *Arabidopsis* AtMYB44 and AtMYB73 (with 55% identity). The phylogenetic analysis demonstrated that HbMYB44 clustered into subgroup 22 together with AtMYB44, AtMYB70, AtMYB73, AtMYB77, rubber tree HblMYB7, HblMYB9, HblMYB10, and HblMYB35 (**Figure 2**). Moreover, two specific motifs of subgroup 22, namely, motif 22.1 (TGLYMSPxSP) and motif 22.2 (GxFMxVVQEMIxxEVRSYM) were found to be present in HbMYB44 (**Figure 1B**, indicated by red boxes).

HbMYB44 Is a Nuclear Protein With Transactivation Activity

To determine the subcellular localization of HbMYB44, a fusion protein HbMYB44-GFP driven by the 35S promoter was generated and transformed into tobacco epidermal cells. The nucleus dyed with DAPI was blue, and HbMYB44-GFP fusion protein was green, both of which were cyan after being merged.

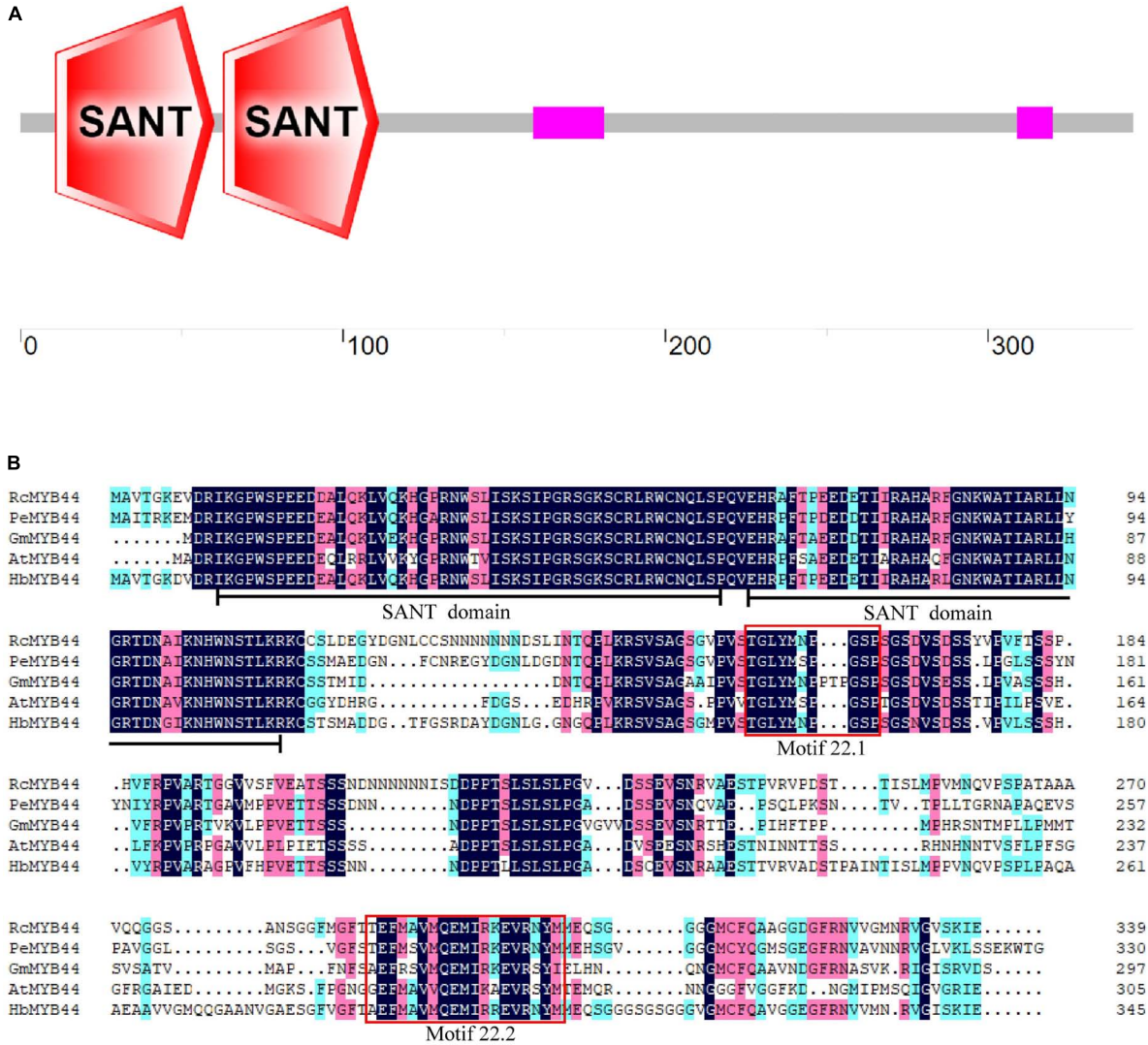


FIGURE 1 | Analysis of HbMYB44 protein structure and multiple sequence alignment. **(A)** Protein architecture of HbMYB44. **(B)** Multiple sequence alignment of HbMYB44 and its homologous proteins. The proteins and their accession numbers are as follows, HbMYB44 (*Hevea brasiliensis*, JN038190), RcMYB44 (*Ricinus communis*, XP_002524926.1), PeMYB44 (*Populus euphratica*, XP_011006987.1), GmMYB44 (*Glycine max*, NP_001238087.1), and AtMYB44 (*Arabidopsis thaliana*, AT5G67300). Identical amino acids are marked with a dark background and well-conserved residues are marked with a pink background. Two SANT domains are underlined and two conserved motifs 22.1 and 22.2 are shown in red boxes.

The results demonstrated that HbMYB44-GFP was exclusively localized in the nucleus of tobacco epidermal cells, while the control empty vector free-GFP was distributed throughout the cells (**Figure 3A**). These results demonstrated that HbMYB44 is a nuclear protein.

Since MYB proteins function as TFs (Li et al., 2006), we tested the transactivation activity of HbMYB44. GAL4 DNA-BD-HbMYB44 fusion proteins were expressed in yeast, and their ability to activate transcription from the GAL4 upstream activation sequence and to promote yeast growth in defective media was analyzed. The full-length ORF of *HbMYB44* (HbMYB44), the N-terminal (encoding the 1st–200th aa, HbMYB44-▲N) and the C-terminal (encoding the 201st–345th aa, HbMYB44-▲C) fragments of *HbMYB44* were individually fused to the GAL4 DNA-BD within the pGBKT7 vector (**Figure 3B**) and transformed into yeast strain Y2H Gold for transactivation activity analysis as previously reported (Zhao T. et al., 2016). As shown in **Figure 3B**, all constructs (including pGBKT7-HbMYB44, pGBKT7-HbMYB44-▲N, pGBKT7-HbMYB44-▲C, negative control pGBKT7, and positive control pGBKT7-HbNAC24) grew on the SD/-Trp media, but only pGBKT7-HbNAC24, pGBKT7-HbMYB44, and pGBKT7-HbMYB44-▲C grew on the TDO media and appeared blue on TDO/X media. The full-length ORF of *HbMYB44* fused with the DNA-BD of GAL4 can self-activate the reporter genes in yeast cells. There was no effect on its transactivation activity when 200 aa within the N-terminus was deleted. However, truncation of 145 aa within the C-terminus resulted in the loss

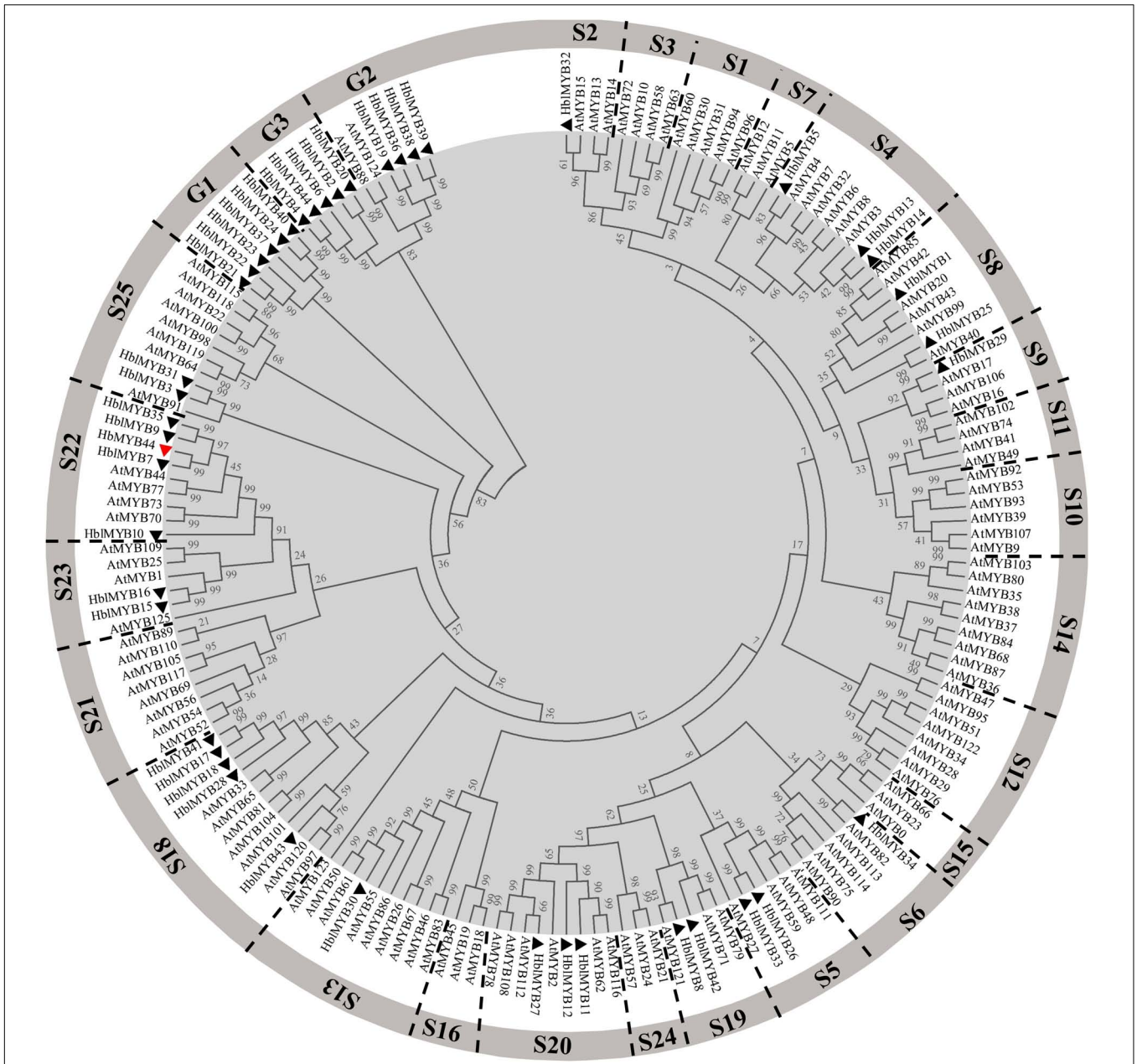


FIGURE 2 | Phylogenetic tree analysis of HbMYB44 and the MYB family members of *Arabidopsis thaliana* and *Hevea brasiliensis*. *HblMYBs* identified from the laticifer transcriptome database and the genome database of *Hevea brasiliensis* (Wang et al., 2017) are indicated by triangles. Information on *Hevea brasiliensis* and *Arabidopsis* MYBs are listed in **Supplementary Table 1**.

of transactivation activity. These results showed that HbMYB44 is a nuclear protein whose last 145 aa of the C-terminus have transactivation activity.

HbMYB44 Expression Was Induced in Response to Multiple Phytohormones and Stresses

The expression pattern of *HbMYB44* in rubber tree was measured using qRT-PCR. *HbMYB44* was expressed in all the tested tissues, but its expression in flowers was higher than that in other tissues; the expression in the flowers was nearly 6.5-fold higher than that in the bark (**Figure 4A**). Since MYBs of subgroup 22 play important roles in regulating the stress response and phytohormone signaling in plants, *HbMYB44* expression following different phytohormones and stress treatments was analyzed systematically. As shown in **Figure 4B**, *HbMYB44* expression was rapidly induced by wounding, and it was upregulated by 45.9-fold at 2 h after treatment. Under drought stress, *HbMYB44* expression was dramatically increased and

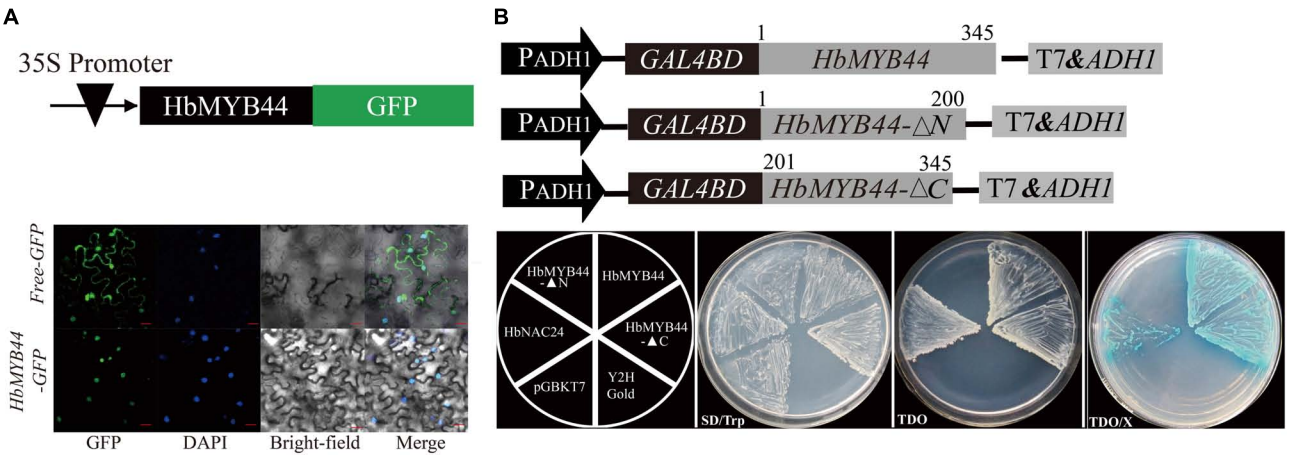


FIGURE 3 | Subcellular localization and transactivation activity of HbMYB44. **(A)** Subcellular localization of HbMYB44 in the leaf epidermal cells of tobacco. Bars = 30 μm. **(B)** Analysis of the transactivation activity of HbMYB44 in yeast. The full length ORF of *HbMYB44*, and fragments encoding the N-terminus (1st–200th aa, *HbMYB44-▲N*) or the C-terminus (201st–345th aa, *HbMYB44-▲C*) of HbMYB44 were fused to the GAL4 DNA-binding domain of the pGBKT7 vector for the assay. TDO, SD/-Trp/-His/-Ade media; TDO/X, TDO media supplemented with X-α-Gal.

peaked at 10 days after treatment, the level of which was 34.1-fold greater than that at 0 day (before treatment) (**Figure 4C**). After H_2O_2 treatment, *HbMYB44* expression was increased by 12.7-fold at 6 h (**Figure 4D**). Furthermore, *HbMYB44* expression was strongly induced in response to ET, MeJA, and ABA and peaked at 6 h, the levels of which were increased to 31.2-fold, 9.9-fold, and 28.4-fold, respectively (**Figure 4E**). In response to SA treatment, the *HbMYB44* expression level was upregulated by 2.8-fold at 2 h (**Figure 4E**). Interestingly, *HbMYB44* expression was rapidly triggered by GA_3 application. *HbMYB44* expression peaked at 0.5 h, which was an increase of approximately 83-fold, but then quickly fell back at 2 h (**Figure 4E**). These results indicated that the *HbMYB44* expression was induced by multiple stress conditions and exogenous phytohormones.

HbMYB44 Overexpression Increases the Expression of Homologous Genes and Interacting Protein Genes in *Arabidopsis*

To verify the function of *HbMYB44* in response to stress and phytohormones, transgenic *Arabidopsis* lines overexpressing *HbMYB44* (OE) were generated. The OE plants of the T2 generation were tested using PCR with the specific primers of *HbMYB44*, the results of which indicated that all 10 OE lines contained the target gene (**Figure 5A**). Then, the expression level of *HbMYB44* in OE lines was confirmed using qRT-PCR. There were differences in the expression abundance of *HbMYB44* among different transgenic lines, and OE2 with the highest expression was used for further analysis (**Figure 5B**).

Since HbMYB44 functions as a TF with transactivation activity (**Figure 3B**), and its sequence is highly homologous to those of *Arabidopsis* MYBs of subgroup 22 (**Figure 2**), we analyzed whether HbMYB44 altered the expression of its homologous genes and their interacting proteins in transgenic *Arabidopsis* plants by qRT-PCR. As shown in **Figure 6**, except for that of *AtMYB77*, the expression of genes of subgroup S22, including *AtMYB44*, *AtMYB70,* and *AtMYB73*, was obviously increased in the *HbMYB44*-overexpressing plants compared with the WT plants. Moreover, the expression of *AtWRKY70* and *AtPYL8* but not *AtPYL9* was also obviously increased in OE plants. These results indicated that HbMYB44 enhanced the expression of its homologous genes and interacting protein-encoding genes in transgenic *Arabidopsis*.

Overexpression of *HbMYB44* in *Arabidopsis* Enhances Tolerance to Abiotic Stress and Promotes Recovery From Root Growth Inhibition by Exogenous Phytohormones

Under normal conditions, there was no marked difference in the morphology of WT and OE plants cultivated in MS media, but different phenotypes, especially in terms of root growth, were found under stress and treatments involving exogenous phytohormones (**Figure 7A**). As shown in **Figure 7**, the growth of leaves and roots of both in WT and OE2 plants was severely suppressed under stress and exogenous phytohormone conditions. In the media supplemented with exogenous ABA, MeJA, GA_3, and SA, the root length of OE plants significantly increased compared with that of the WT plants under the same conditions. However, the root length of OE plants was reduced compared with that of the WT plants in the media supplemented with ET (**Figures 7A,B**). These results demonstrated that the root growth of OE plants is less sensitive to ABA, MeJA, GA_3, and SA, but more sensitive to ET. Under NaCl-salt stress, mannitol-mediated osmotic stress, and PEG-simulated drought stress conditions, the root lengths were markedly increased in the OE plants compared with the WT; the roots

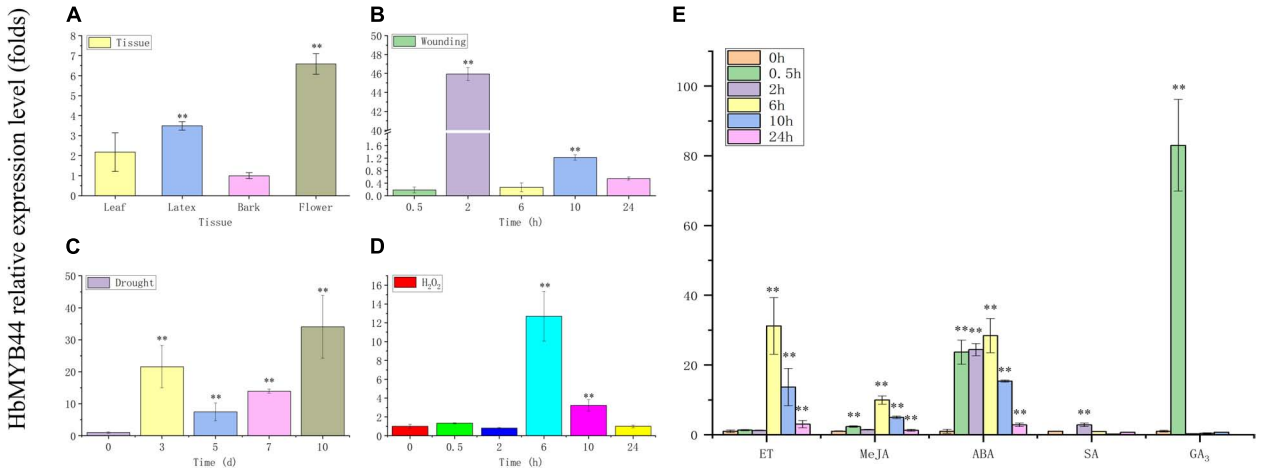


FIGURE 4 | Expression of *HbMYB44* in different tissues and in response to stress and phytohormone treatments in rubber tree. The expression of *HbMYB44* was determined in different tissues of tapping trees **(A)** and in the leaves of tissue culture-generated seedlings subjected to wounding **(B)**, drought stress **(C)**, H_2O_2 **(D)**, and phytohormones **(E)**. The relative expression levels were compared with those at 0 h or 0 day. **Indicates a significant differences at the $P < 0.01$ level.

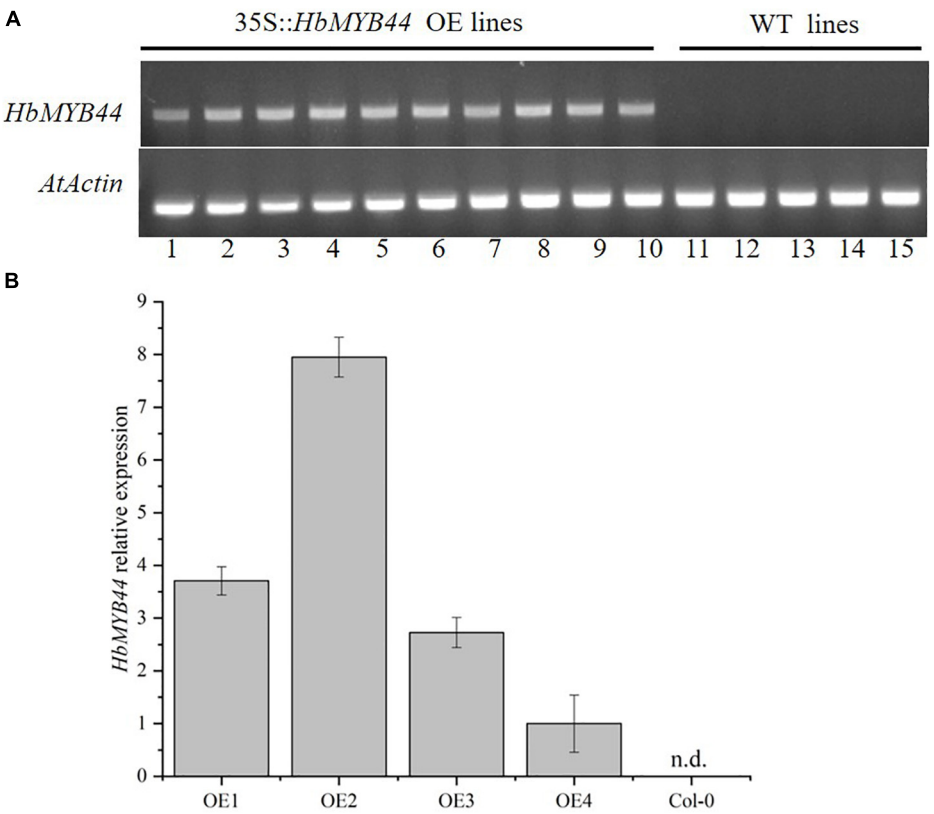


FIGURE 5 | Verification of *HbMYB44* expression in transgenic *Arabidopsis*. **(A)** Validation of *HbMYB44* integration into the genome of transgenic plants. *Arabidopsis AtActin2* was used as a control. **(B)** Analysis of *HbMYB44* expression levels in transgenic plants by qRT-PCR. *AtActin2* was used as a housekeeping gene. Each value is the average of six plants tested in triplicate. n.d. means no detection.

especially increased more than 5-fold under NaCl conditions (**Figures 7A,B**). These results indicated that *HbMYB44* enhanced tolerance to abiotic stress and promoted recovery from root growth inhibition in response to exogenous ABA, MeJA, GA3, and SA, suggesting that *HbMYB44* is involved in a complex phytohormone signaling network.

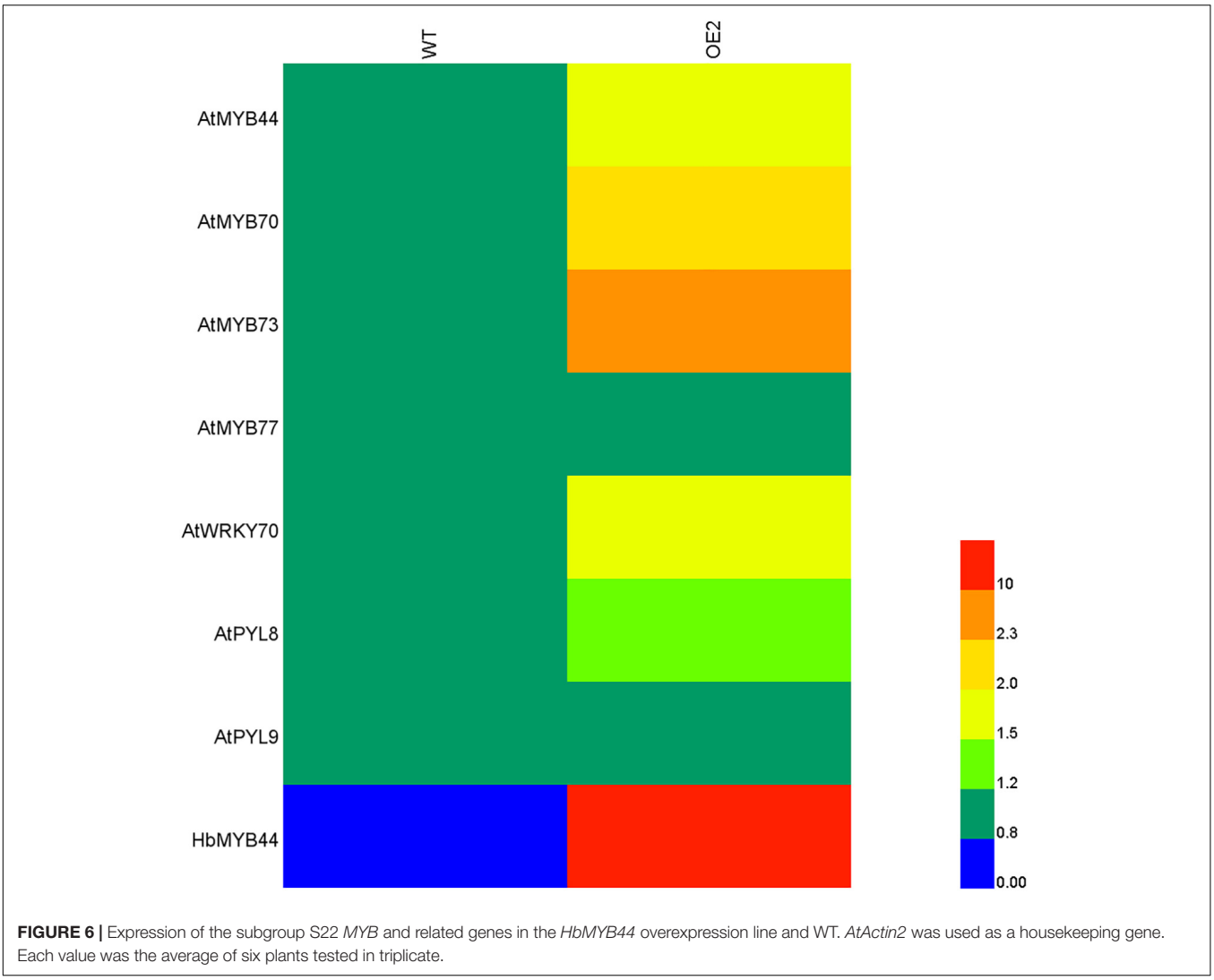


FIGURE 6 | Expression of the subgroup S22 *MYB* and related genes in the *HbMYB44* overexpression line and WT. *AtActin2* was used as a housekeeping gene. Each value was the average of six plants tested in triplicate.

DISCUSSION

The DNA-BD of MYB TFs is essential for DNA recognition and transactivation (Ogata et al., 1996). However, significant differences in plant and animal MYB TFs are present in the domains of these TFs (Williams and Grotewold, 1997). Most plant MYB functions are related to their DNA-BD (Millard et al., 2019). The phylogenetic analysis showed that 125 R2R3-MYBs identified from the genome of *A. thaliana* can be divided into 25 subgroups (Stracke et al., 2001). There are two typical SANT domains present in subgroup 22 MYB TFs, enabling DNA-binding specificities and DNA interactions (Millard et al., 2019). Additionally, the subgroup 22 MYB TFs contain two conserved motifs, namely, motif 22.1 and motif 22.2, which are the binding domains of interacting proteins (Stracke et al., 2001; Li et al., 2014). This study showed that HbMYB44 is a member of subgroup 22, as supported by the phylogenetic analysis in which HbMYB44 is categorized into subgroup 22 together with AtMYB44, AtMYB70, AtMYB73, AtMYB77 and with rubber tree HblMYB7, HblMYB9, HblMYB10, and HblMYB35 (**Figure 2**). The two conserved motifs (22.1 and 22.2) are also present in HbMYB44 (**Figure 1B**). Moreover, HbMYB44 is located to the nucleus and has transactivation activity at the C-terminus. The region encompassing aa 201-345 is indispensable for transcriptional activation (**Figure 3**). These results indicated that HbMYB44 is a TF with transcriptional activation activity. Based on the conserved function of the genes in subgroup 22, it is speculated that HbMYB44, HblMYB7, HblMYB9, HblMYB10, and HblMYB35 in this subgroup may play an important role in phytohormone signaling and stress responses. Thus, further study should be carried out to verify the functions of these genes, which would help elucidate phytohormone signaling networks involved in plant defense and NR biosynthesis in rubber tree.

In plants, MYB44 is a key factor functioning as a bHLH-MYB-WRKY70 transcriptional module in plant non-specific phytohormone induction (Jung et al., 2010), defense responses (Shim and Choi, 2013; Shim et al., 2013), root elongation (Zhao Q. et al., 2016), stress resistance (Park et al., 2012),

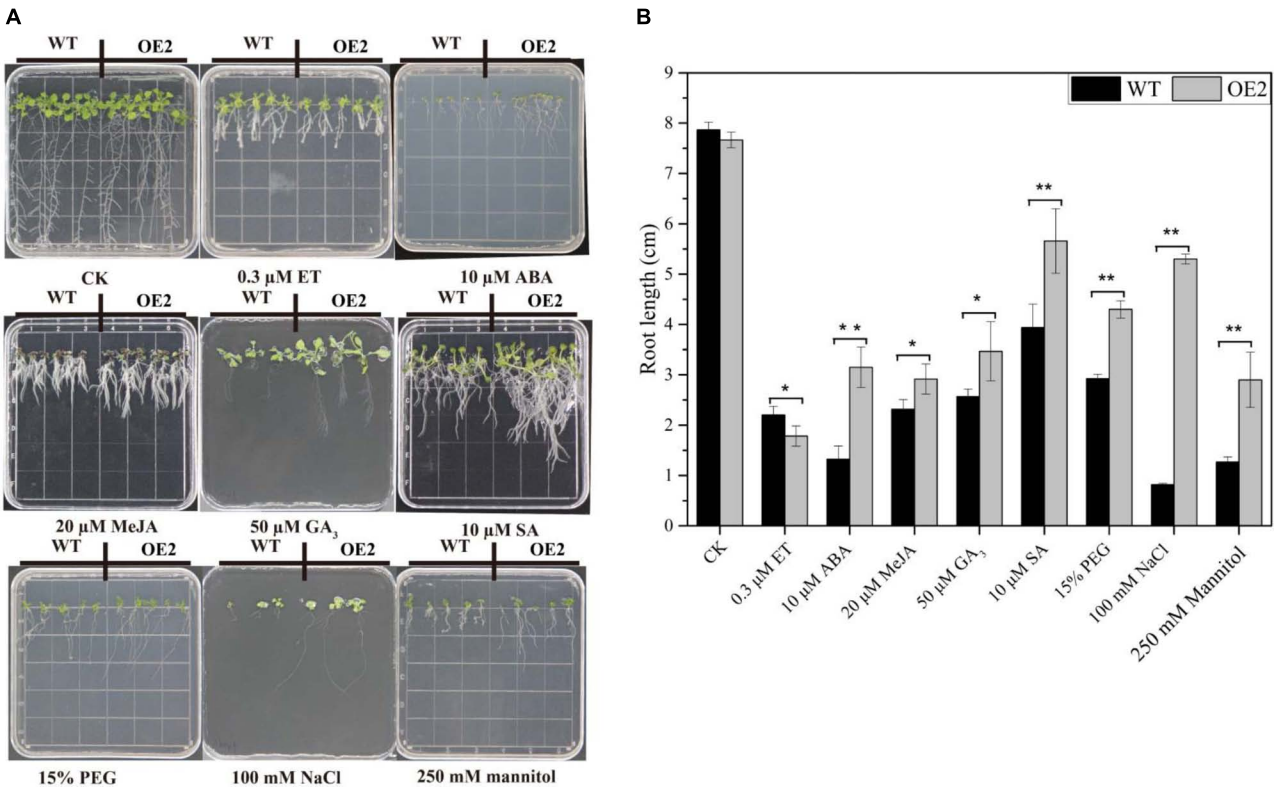


FIGURE 7 | Phytohormone response and stress tolerance assays in the *HbMYB44* overexpression line and wild type (WT). **(A)** Phenotypes of the *HbMYB44* overexpression line and WT upon application of different phytohormones and stress conditions. **(B)** The root length of the seedlings under different phytohormone and stress treatments was measured after 10–13 days of treatments. Each value is the average of three batches of measurement data. The error bars represent ± SDs of three biological replicates, and the asterisks indicate significant differences between OE2 and WT plants (* and **Indicate $P < 0.05$ and $P < 0.01$, respectively).

etc. Genes encoding the subgroup 22 MYB TFs have similar expression patterns and are associated with resistance to disease and stress. *AtMYB44* is expressed in different tissues throughout plants, with the highest abundance in the vasculature and stomata (Kirik et al., 1998). Moreover, *AtMYB44* expression was shown to be inducted in response to drought, cold, high salinity, $CdCl_2$, sugars, wounding, SA, MeJA, ET, GA_3, ABA, brassinosteroid, and cytokinin (Chen et al., 2006; Jung et al., 2008, 2010). The expression patterns of *AtMYB44*, *AtMYB73*, and *AtMYB77* were also shown to increase after wounding (Cheong et al., 2002), white-light (Ma et al., 2005) and cold stress treatments (Fowler and Thomashow, 2002). In this study, *HbMYB44* was expressed in all tissues, and it was strongly induced by multiple stresses (including wounding, drought stress, and H_2O_2 treatments) in rubber tree (**Figure 4**). Moreover, its transcript rapidly accumulated in response to phytohormones, and a significant increase in *HbMYB44* transcription was observed within 0.5 h after treatment with MeJA, ABA, and GA_3 (**Figure 4E**). Therefore, *HbMYB44* transcripts accumulated non-specifically under diverse stress conditions and in response to various phytohormones, the results of which are similar to those of *AtMYB44* in *A. thaliana*, suggesting *HbMYB44* has functions similar to those of *AtMYB44*.

The stress response and growth regulation of MYB TFs are related to the phytohormone signaling pathway. For instance, AtMYB44 regulates *EIN2* expression, and *EIN2* is the key gene in ethylene signaling and insect resistance. Moreover, AtMYB44 regulates plant insect resistance by activating *EIN2* and its related defense responses in *Arabidopsis* (Lu et al., 2013). The ABA receptor RCAR3/PYL8 promotes lateral root growth by directly interacting with MYB77, MYB44, and MYB73, then enhancing MYB77-dependent transcription of auxin-responsive genes, to augment auxin signaling (Zhao et al., 2014). In the ABA signaling pathway, the ABA receptor RCAR1/PYL9 interacts with all the members of the subgroup 22 MYBs, i.e., AtMYB44, AtMYB70, AtMYB73, and AtMYB77 (Li et al., 2014). In *Solanum melongena*, the AtMYB44 homolog SmMYB44 enhances resistance to *Ralstonia solanacearum* by directly binding to the promoter of spermidine synthase (SPDS) gene *SmSPDS* and activating its expression (Qiu et al., 2019). These studies suggested that the members of subgroup 22 MYBs share similar functions and are functionally redundant. Through hormone crosstalk, AtMYB44 directly binds to the promoter of *AtWRKY70*, regulates *AtWRKY70* expression, and modulates antagonistic interactions between SA and JA signaling (Shim et al., 2013). Phytohormones play important roles in the NR production of rubber trees.

The application of ET to the bark of rubber tree increases NR yield per tapping (Kush, 1994), and JA treatment of the bark can promote secondary laticifer differentiation and subsequent NR biosynthesis (Haoji-Lin, 2000; Deng et al., 2018). Rubber tree plantations often suffer from various biotic stresses and abiotic stresses, especially wounding (caused by tapping in the bark for NR harvesting). Thus, the identification and functional characterization of stress response and phytohormone signaling genes are of great significance for improving plant stress resistance and NR yield in rubber trees. In this study, overexpression of *HbMYB44* in *Arabidopsis* demonstrated that OE plants enhanced resistance to mannitol-mediated osmotic stress, NaCl-salt stress, and PEG-simulated drought stress (**Figure 7**). Moreover, *HbMYB44* promoted recovery from root growth inhibition caused by ABA, MeJA, GA_3, and SA (**Figure 7**). In addition, the leaf biomass of OE plants was higher than those of the WT plants in the media supplemented with GA_3, and the root biomass of OE plants was higher than those of WT plants in the media supplemented with SA (**Figure 7A**). These results demonstrated that *HbMYB44* is involved in a complex phytohormone signaling network. Interestingly, expression of the genes of subgroup S22 (including *AtMYB44*, *AtMYB70,* and *AtMYB73*) and interacting protein-encoding genes (*AtWRKY70* and *AtPYL8*) were significantly increased in OE plants (**Figure 6**). These results suggested that *HbMYB44* enhances stress tolerance and modulates phytohormone induction by increasing the expression of its homologous genes and interacting protein-encoding genes in *Arabidopsis*. Therefore, overexpression of *HbMYB44* in the rubber tree will be further conducted to determine its function in NR biosynthesis, stress resistance, and phytohormone signaling.

CONCLUSION

The rubber tree HbMYB44 is a member of the subgroup 22 MYB TFs, which functions in the nucleus and has transactivation activity *via* its C-terminus. *HbMYB44* expression was strongly induced in response to multiple phytohormones and stresses. The overexpression of *HbMYB44* in *Arabidopsis* enhanced tolerance to abiotic stress and promoted recovery from root growth inhibition by exogenous phytohormones. This study suggested that *HbMYB44* plays an important role in modulating multiple phytohormone signaling pathways and stress resistance of rubber tree.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

BQ was the experimental designer and the executor of the experimental research, completed the data analysis, and wrote the first draft of the manuscript. BQ, S-LF, H-YY, and Y-XL participated in experimental design, performed experimental research, and experimental result analysis. BQ and L-FW directed the design of the experiment, data analysis, and the modification of the manuscript. All authors read and approved the final manuscript.

FUNDING

This study was funded by the Major Science and Technology Plan of Hainan Province (ZDKJ2021004) and National Natural Science Foundation of China (Grant Nos. 31570591, 31270643, and 31970364).

ACKNOWLEDGMENTS

We are grateful to Xing-Liang Hou (South China Botanical Garden, Chinese Academy of Sciences), providing the vector 35S: XXGFP, and Gui-Juan Kang (Rubber Research Institute, Chinese Academy of Tropical Agricultural Sciences) for the assistance in the transcriptional activation assay.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.893896/full#supplementary-material

REFERENCES

Ampomah-Dwamena, C., Thrimawithana, A. H., Dejnoprat, S., Lewis, D., Espley, R. V., and Allan, A. C. (2019). A kiwifruit (actinidia deliciosa) R2R3-MYB transcription factor modulates chlorophyll and carotenoid accumulation. *New Phytol.* 221, 309–325. doi: 10.1111/nph.15362

Chen, Y. H., Yang, X. Y., He, K., Liu, M. H., Li, J. G., Gao, Z. F., et al. (2006). The MYB transcription factor superfamily of *Arabidopsis*: expression analysis and phylogenetic comparison with the rice MYB family. *Plant Mol. Biol.* 60, 107–124. doi: 10.1007/s11103-005-2910-y

Cheong, Y. H., Chang, H.-S., Gupta, R., Wang, X., Zhu, T., and Luan, S. (2002). Transcriptional profiling reveals novel interactions between wounding, pathogen, abiotic stress, and hormonal responses in *Arabidopsis*. *Plant Physiol.* 129, 661–677. doi: 10.1104/pp.002857

Deng, X., Guo, D., Yang, S., Shi, M., Chao, J., Li, H., et al. (2018). Jasmonate signalling in the regulation of rubber biosynthesis in laticifer cells of rubber tree, *Hevea brasiliensis*. *J. Exp. Bot.* 69, 3559–3571. doi: 10.1093/jxb/ery169

Fowler, S., and Thomashow, M. F. (2002). *Arabidopsis transcriptome* profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway. *Plant Cell* 14, 1675–1690. doi: 10.1105/tpc.003483

Fujita, M., Fujita, Y., Maruyama, K., Seki, M., Hiratsu, K., Ohme-Takagi, M., et al. (2004). A dehydration-induced NAC protein, RD26, is involved in a novel ABA-dependent stress-signaling pathway. *Plant J.* 39, 863–876. doi: 10.1111/j.1365-313X.2004.02171.x

Guan, Q. J., Wang, L. F., Bu, Q. Y., and Wang, Z. Y. (2014). The rice gene OsZFP6 functions in multiple stress tolerance responses in yeast and *Arabidopsis*. *Plant Physiol. Biochem.* 82, 1–8. doi: 10.1016/j.plaphy.2014.04.021

Haoji-Lin, B. Z. (2000). Laticifer differentiation in *Hevea brasiliensis*: induction by exogenous jasmonic acid and linolenic acid. *Ann. Bot.* 85, 37–43. doi: 10.1006/anbo.1999.0995

Jung, C., Seo, J. S., Han, S. W., Koo, Y. J., Kim, C. H., Song, S. I., et al. (2008). Overexpression of AtMYB44 enhances stomatal closure to confer abiotic stress tolerance in transgenic *Arabidopsis*. *Plant. Physiol.* 146, 623–635. doi: 10.1104/pp.107.110981

Jung, C., Shim, J. S., Seo, J. S., Lee, H. Y., Kim, C. H., Choi, Y. D., et al. (2010). Non-specific phytohormonal induction of AtMYB44 and suppression of jasmonate-responsive gene activation in *Arabidopsis thaliana*. *Mol. Cells* 29, 71–76. doi: 10.1007/s10059-010-0009-z

Kang, G., Yu, L. I., and Zeng, R. (2014). Clone and expression of HbNAC24 gene from *Hevea brasiliensis*. *Acta Bot. Boreali Occidentalia Sin.* 34, 2374–2381.

Kapuscinski, J. (1995). DAPI: a DNA-specific fluorescent probe. *Biotech Histochem.* 70, 220–233. doi: 10.3109/10520299509108199

Kim, J. H., Nguyen, N. H., Jeong, C. Y., Nguyen, N. T., Hong, S. W., and Lee, H. (2013). Loss of the R2R3 MYB, AtMyb73, causes hyper-induction of the SOS1 and SOS3 genes in response to high salinity in *Arabidopsis*. *J. Plant. Physiol.* 170, 1461–1465. doi: 10.1016/j.jplph.2013.05.011

Kirik, V., Kolle, K., Misera, S., and Baumlein, H. (1998). Two novel MYB homologues with changed expression in late embryogenesis-defective *Arabidopsis mutants*. *Plant Mol. Biol.* 37, 819–827. doi: 10.1023/a:1006011002499

Kush, A. (1994). Isoprenoid biosynthesis: the hevea factory. *Plant Physiol. Biochem.* 32, 761–767.

Li, D., Li, Y., Zhang, L., Wang, X., Zhao, Z., Tao, Z., et al. (2014). *Arabidopsis* ABA receptor RCAR1/PYL9 interacts with an R2R3-type MYB transcription factor, AtMYB44. *Int. J. Mol. Sci.* 15, 8473–8490. doi: 10.3390/ijms15058473

Li, J., Yang, X., Yan, W., Li, X., Gao, Z., Meng, P., et al. (2006). Two groups of MYB transcription factors share a motif which enhances trans-activation activity. *Biochem. Biophys. Res. Commun.* 341, 1155–1163. doi: 10.1016/j.bbrc.2006.01.077

Lu, B. B., Li, X. J., Sun, W. W., Li, L., Gao, R., Zhu, Q., et al. (2013). AtMYB44 regulates resistance to the green peach aphid and diamondback moth by activating EIN2-affected defences in *Arabidopsis*. *Plant. Biol. (Stuttg)* 15, 841–850. doi: 10.1111/j.1438-8677.2012.00675.x

Ma, L., Sun, N., Liu, X., Jiao, Y., Zhao, H., and Deng, X. W. (2005). Organ-specific expression of *Arabidopsis genome* during development. *Plant Physiol.* 138, 80–91. doi: 10.1104/pp.104.054783

Millard, P. S., Kragelund, B. B., and Burow, M. (2019). R2R3 MYB transcription factors - functions outside the DNA-binding domain. *Trends Plant Sci.* 24, 934–946. doi: 10.1016/j.tplants.2019.07.003

Nguyen, N. H., and Cheong, J. J. (2018a). H2A.Z-containing nucleosomes are evicted to activate AtMYB44 transcription in response to salt stress. *Biochem. Biophys. Res. Commun.* 499, 1039–1043. doi: 10.1016/j.bbrc.2018.04.048

Nguyen, N. H., and Cheong, J. J. (2018b). The AtMYB44 promoter is accessible to signals that induce different chromatin modifications for gene transcription. *Plant Physiol. Biochem.* 130, 14–19. doi: 10.1016/j.plaphy.2018.06.030

Nguyen, X. C., Hoang, M. H., Kim, H. S., Lee, K., Liu, X. M., Kim, S. H., et al. (2012). Phosphorylation of the transcriptional regulator MYB44 by mitogen activated protein kinase regulates *Arabidopsis* seed germination. *Biochem. Biophys. Res. Commun.* 423, 703–708. doi: 10.1016/j.bbrc.2012.06.019

Nolan, T., Hands, R. E., and Bustin, S. A. (2006). Quantification of mRNA using real-time RT-PCR. *Nat. Protoc.* 1, 1559–1582. doi: 10.1038/nprot.2006.236

Ogata, K., Kanei-Ishii, C., Sasaki, M., Hatanaka, H., Nagadoi, A., Enari, M., et al. (1996). The cavity in the hydrophobic core of Myb DNA binding domain is reserved for DNA recognition and trans-activation. *Nat. Struct. Mol. Biol.* 3, 178–187. doi: 10.1038/nsb0296-178

Park, J.-B., Sendon, P. M., Kwon, S. H., Seo, H. S., Park, S.-K., Kim, J. H., et al. (2012). Overexpression of stress-related genes, BrERF4 and AtMYB44, in *Arabidopsis thaliana* alters cell expansion but not cell proliferation during leaf growth. *J. Plant Biol.* 55, 406–412. doi: 10.1007/s12374-012-0114-y

Peng, S. Q., Wu, K. X., Huang, G. X., and Chen, S. C. (2011). HbMyb1, a Myb transcription factor from *Hevea brasiliensis*, suppresses stress induced cell death in transgenic tobacco. *Plant Physiol. Biochem.* 49, 1429–1435. doi: 10.1016/j.plaphy.2011.09.007

Persak, H., and Pitzschke, A. (2013). Tight interconnection and multi-level control of *Arabidopsis* MYB44 in MAPK cascade signalling. *PLoS One* 8:e57547. doi: 10.1371/journal.pone.0057547

Puskas, J. E., Gautriaud, E., Deffieux, A., and Kennedy, J. P. (2006). Natural rubber biosynthesis—a living carbocationic polymerization? *Progr. Polymer Sci.* 31, 533–548.

Qin, B., Wang, M., He, H. X., Xiao, H. X., Zhang, Y., and Wang, L. F. (2019). Identification and characterization of a potential candidate Mlo gene conferring susceptibility to powdery mildew in rubber tree. *Phytopathology* 109, 1236–1245. doi: 10.1094/PHYTO-05-18-0171-R

Qin, B., Zhang, Y., and Wang, M. (2014). Molecular cloning and expression of a novel MYB transcription factor gene in rubber tree. *Mol. Biol. Rep.* 41, 8169–8176. doi: 10.1007/s11033-014-3717-1

Qin, B., Zheng, F., and Zhang, Y. (2015). Molecular cloning and characterization of a Mlo gene in rubber tree (*Hevea brasiliensis*). *J. Plant Physiol.* 175, 78–85. doi: 10.1016/j.jplph.2014.10.019

Qiu, Z., Yan, S., Xia, B., Jiang, J., Yu, B., Lei, J., et al. (2019). The eggplant transcription factor MYB44 enhances resistance to bacterial wilt by activating the expression of spermidine synthase. *J. Exp. Bot.* 70, 5343–5354. doi: 10.1093/jxb/erz259

Schmittgen, T. D., and Livak, K. J. (2008). Analyzing real-time PCR data by the comparative C(T) method. *Nat. Protoc.* 3, 1101–1108. doi: 10.1038/nprot.2008.73

Sharkey, T. D., and Yeh, S. (2001). Isoprene emission from plants. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* 52, 407–436. doi: 10.1146/annurev.arplant.52.1.407

Shen, X. J., Wang, Y. Y., Zhang, Y. X., Guo, W., Jiao, Y. Q., and Zhou, X. A. (2018). Overexpression of the wild soybean R2R3-MYB transcription factor GsMYB15 enhances resistance to salt stress and *Helicoverpa armigera* in transgenic *Arabidopsis*. *Int. J. Mol. Sci.* 19:3958. doi: 10.3390/ijms19123958

Shim, J. S., and Choi, Y. D. (2013). Direct regulation of WRKY70 by AtMYB44 in plant defense responses. *Plant Signal Behav.* 8:e20783. doi: 10.4161/psb.24509

Shim, J. S., Jung, C., Lee, S., Min, K., Lee, Y. W., Choi, Y., et al. (2013). AtMYB44 regulates WRKY70 expression and modulates antagonistic interaction between salicylic acid and jasmonic acid signaling. *Plant J.* 73, 483–495. doi: 10.1111/tpj.12051

Stracke, R., Werber, M., and Weisshaar, B. (2001). The R2R3-MYB gene family in *Arabidopsis thaliana*. *Curr. Opin. Plant. Biol.* 4, 447–456. doi: 10.1016/s1369-5266(00)00199-0

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197

Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2002). Multiple sequence alignment using ClustalW and ClustalX. *Curr. Protoc. Bioinform.* Chapter 2:Unit23. doi: 10.1002/0471250953.bi0203s00

Wang, L., Gao, W., Wu, X., Zhao, M., Qu, J., Huang, C., et al. (2018). Genome-wide characterization and expression analyses of *Pleurotus ostreatus* MYB transcription factors during developmental stages and under heat stress based on de novo sequenced genome. *Int. J. Mol. Sci.* 19:2052. doi: 10.3390/ijms19072052

Wang, X., Angelis, N., and Thein, S. L. (2018). MYB - a regulatory factor in hematopoiesis. *Gene* 665, 6–17. doi: 10.1016/j.gene.2018.04.065

Wang, Y., Zhan, D. F., Li, H. L., Guo, D., Zhu, J. H., and Peng, S. Q. (2017). Transcriptome-wide identification and characterization of MYB transcription factor genes in the laticifer cells of *Hevea brasiliensis*. *Front. Plant Sci.* 8:1974. doi: 10.3389/fpls.2017.01974

Williams, C. E., and Grotewold, E. (1997). Differences between plant and animal Myb domains are fundamental for DNA binding activity, and chimeric Myb

domains have novel DNA binding specificities. *J. Biol. Chem.* 272, 563–571. doi: 10.1074/jbc.272.1.563

Zhang, C., Ma, R., Xu, J., Yan, J., Guo, L., Song, J., et al. (2018). Genome-wide identification and classification of MYB superfamily genes in peach. *PLoS One* 13:e0199192. doi: 10.1371/journal.pone.0199192

Zhang, J., Chen, J., Yi, Q., Hu, Y., Liu, H., Liu, Y., et al. (2014). Novel role of ZmaNAC36 in co-expression of starch synthetic genes in maize endosperm. *Plant Mol. Biol.* 84, 359–369. doi: 10.1007/s11103-013-0153-x

Zhang, T., Zhao, Y., Wang, Y., Liu, Z., and Gao, C. (2018). Comprehensive analysis of MYB gene family and their expressions under abiotic stresses and hormone treatments in *Tamarix hispida*. *Front. Plant. Sci.* 9:1303. doi: 10.3389/fpls.2018.01303

Zhao, Q., Li, M., Jia, Z., Liu, F., Ma, H., Huang, Y., et al. (2016). AtMYB44 positively regulates the enhanced elongation of primary roots induced by N-3-oxo-hexanoyl-homoserine lactone in *Arabidopsis thaliana*. *Mol. Plant Microbe Int.* 29, 774–785. doi: 10.1094/MPMI-03-16-0063-R

Zhao, T., Huang, X., and Xia, Y. (2016). Human heart cell proteins interacting with a C-terminally truncated 2A protein of coxsackie B3 virus: identification by the yeast two-hybrid system. *Virus Genes* 52, 172–178. doi: 10.1007/s11262-015-1270-1

Zhao, Y., Xing, L., Wang, X., Hou, Y. J., Gao, J., Wang, P., et al. (2014). The ABA receptor PYL8 promotes lateral root growth by enhancing MYB77-dependent transcription of auxin-responsive genes. *Sci. Signal* 7:ra53. doi: 10.1126/scisignal.2005051

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 14 June 2022
doi: 10.3389/fpls.2022.803400

Transcriptome of Endophyte-Positive and Endophyte-Free Tall Fescue Under Field Stresses

Md. Shofiqul Islam[1,2†], Nick Krom[3†], Taegun Kwon[4,5†], Guifen Li[4] and Malay C. Saha[1†]*

[1]Grass Genomics, Noble Research Institute LLC, Ardmore, OK, United States, [2]Genetics Laboratory, Indiana Crop Improvement Association, Lafayette, IN, United States, [3]Scientific Computing, Noble Research Institute LLC, Ardmore, OK, United States, [4]Genomics Core Facility, Noble Research Institute LLC, Ardmore, OK, United States, [5]Genomics Center, BioDiscovery Institute, University of North Texas, Denton, TX, United States

OPEN ACCESS

Edited by:
Mukesh Jain,
Jawaharlal Nehru University, India

Reviewed by:
Mohammad R. Sabzalian,
Isfahan University of Technology, Iran
Satendra Kumar Mangrauthia,
Indian Institute of Rice Research (ICAR), India

***Correspondence:**
Malay C. Saha
msaha06@yahoo.com

†ORCID:
Md. Shofiqul Islam
orcid.org/0000-0003-3018-4739
Nick Krom
orcid.org/0000-0002-7973-8037
Taegun Kwon
orcid.org/0000-0002-0205-8535
Malay C. Saha
orcid.org/0000-0003-4442-2320

Specialty section:
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

Received: 27 October 2021
Accepted: 05 May 2022
Published: 14 June 2022

Citation:
Islam MS, Krom N, Kwon T, Li G and Saha MC (2022) Transcriptome of Endophyte-Positive and Endophyte-Free Tall Fescue Under Field Stresses.
Front. Plant Sci. 13:803400.
doi: 10.3389/fpls.2022.803400

Tall fescue is one of the primary sources of forage for livestock. It grows well in the marginal soils of the temperate zones. It hosts a fungal endophyte (*Epichloë coenophiala*), which helps the plants to tolerate abiotic and biotic stresses. The genomic and transcriptomic resources of tall fescue are very limited, due to a complex genetic background and outbreeding modes of pollination. The aim of this study was to identify differentially expressed genes (DEGs) in two tissues (pseudostem and leaf blade) between novel endophyte positive (E+) and endophyte-free (E–) Texoma MaxQ II tall fescue genotypes. Samples were collected at three diurnal time points: morning (7:40–9:00 am), afternoon (1:15–2:15 pm), and evening (4:45–5:45 pm) in the field environment. By exploring the transcriptional landscape *via* RNA-seq, for the first time, we generated 226,054 and 224,376 transcripts from E+ and E– tall fescue, respectively through *de novo* assembly. The upregulated transcripts were detected fewer than the downregulated ones in both tissues (S: 803 up and 878 down; L: 783 up and 846 down) under the freezing temperatures (–3.0–0.5°C) in the morning. Gene Ontology enrichment analysis identified 3 out of top 10 significant GO terms only in the morning samples. Metabolic pathway and biosynthesis of secondary metabolite genes showed lowest number of DEGs under morning freezing stress and highest number in evening cold condition. The 1,085 DEGs were only expressed under morning stress condition and, more importantly, the eight candidate orthologous genes of rice identified under morning freezing temperatures, including orthologs of rice phytochrome A, phytochrome C, and ethylene receptor genes, might be the possible route underlying cold tolerance in tall fescue.

Keywords: abiotic stress, cold, endophyte, RNA-seq, tall fescue, transcriptome

INTRODUCTION

Tall fescue (*Festuca arundinacea* Schreb.) is a cool-season perennial grass species cultivated in the temperate zone worldwide. It grows well in the transition zone of the United States (Buckner et al., 1979), where cool- and warm-season grasses are cultivated successfully. Tall fescue is highly productive and provides quality forage. It is grown for pasture, hay, and silage and is

used as a primary source of herbage protein of livestock feed. Tall fescue can grow in a wide range of temperatures between 4 and 35°C with an optimum 20–25°C for higher biomass production (Rogers and Locke, 2013). It can tolerate cold stress, which includes chilling (0–12°C) and/or freezing (<0°C) temperatures for a short period during fall (Interrante et al., 2020). The forage production is reduced during the cold months, January and February, and growth resumes when temperature rebound to ≥12°C in spring. To maintain reasonable production level throughout the cooler portions of the growing season (Casler and van Santen, 2008), understanding cold stress tolerance mechanism of the cool-season grasses is a prime interest for the grass breeder to create freezing tolerance pasture grass varieties for adaptation under the extreme cold climates.

Freezing tolerance is a complex process that requires cumulative small effect of multiple genes (Zhang, 2009). The major cold signaling pathway is the C-repeat binding factors (CBFs)/dehydration-responsive element binding factor (DREB)-mediated transcriptional cascade, which is essential for the activation of *cold responsive* (*COR*) genes reviewed by Shi et al. (2015). In addition to the *COR* genes, the classical phytohormones, such as auxin, abscisic acid, ethylene, cytokinins, gibberellins, jasmonic acid, and brassinosteroids, are involved in the regulation of plant growth and abiotic stress response (Peleg and Blumwald, 2011). In a non-model plant species, where the genomic sequence information is not available, next-generation sequencing of mRNA (RNA-seq) is a robust method to evaluate transcriptional responses under environmental stress condition. For instances, to improve the understanding of cold/freezing responsive genes, comparisons of transcripts under different temperatures have been conducted in Kentucky bluegrass (*Poa pratensis* L.; Zhang et al., 2016), zoysiagrass (*Zoysia* spp. Willd.), sheepgrass (*Leymus chinensis*; Chen et al., 2013). Transcriptomic analyses have been used to identify differentially expressed genes (DEGs) responsible for the expression of traits within contrasting plant materials.

Tall fescue is a natural host of fungal endophyte (*Epichloë coenophiala*), which produces alkaloids that are harmful to the grazing animals (Waller, 2009), but helps the host to fight against biotic and abiotic stresses (Bacon and Siegel, 1988). To understand the cold stress tolerance in field-grown tall fescue, we analyzed transcripts from the pseudostem (S) and leaf blade (L) tissues of endophyte positive (E+) and endophyte-free (E−) tall fescue genotypes at three different time points in a day from a natural field environment using RNA-seq. Tissue type, cellular conditions, and environmental factors all guided transcript profiles that may influence regulatory events, such as splicing and the expression of genes or their isoforms (Marco-Puche et al., 2019). We thus used two tissue types utilizing E+ and E− tall fescue for transcriptome analysis in this study to monitor the changes in plant gene expressions under cold stress in the natural field environment by considering genetic and environmental interaction, evaluating plant responses, and endophyte's influence on the host responses. The aims of this study were to: (i) investigate genome-wide transcriptomic profile of E+ and E− Texoma MaxQ II tall fescue, (ii) identify DEGs in two tissues under three time points, and (iii) identify candidate gene(s) responsible for cold tolerance under freezing condition in the natural field environment. Moreover, the present study would provide us useful information whether an endophyte influences host's genes and their regulatory pathways associated with cold/freezing response in tall fescue.

MATERIALS AND METHODS

Plant Materials

Novel endophyte (AR584) positive (E+) and endophyte-free (E−) tall fescue genotypes of cv. Texoma MaxQ II (referred as "Texoma"; Pennington, United States)[1] were developed at the Noble Research Institute headquartered in Ardmore, Oklahoma, United States. Texoma is a commercial cultivar freely available for cultivation in the United States. The E+ and E− plants were transplanted in the field for seed production *via* open pollination among them. Since the endophyte does not transmit through pollen, seeds were harvested from the E+ and E− mother plants separately. The seeds obtained from the E+ and E− Texoma genotypes were sown in rows in the experimental farm located at Dupy (Latitude: 34°17′12.106″N, Longitude: 96°59′36.608″W), Gene Autry, Oklahoma. Before collecting samples for transcriptome analyses, we collected S tissues separately in ice-cold 15 ml falcon tubes from 15 each E+ and E− genotypes from three random rows of the plot to perform PCR test to identify their endophyte status, as the endophyte is residing only in the S tissues, not in L (Takach et al., 2012). The S samples were freeze-dried and ground separately in the presence of liquid N_2 using mortar and pestle. Genomic DNA was extracted using MagAttract 96 DNA Plant Core Kit (QIAGEN Cat. No. 67163, Hilden, Germany) according to the manufacturer's recommendation. PCR amplifications were performed using primers described in (Charlton et al., 2014) to confirm the E+ and E− status of the Texoma tall fescue genotypes (data not shown).

After confirming E+ and E− status, equal length of S and L tissues were collected on December 10, 2018 under freezing/cold temperatures from the field at: morning between 7:40 am (−3°C) and 9:00 am (0.5°C), afternoon between 1:15 pm (11°C) and 2:15 pm (12°C), and evening between 4:45 pm (12°C) and 5:45 pm (10°C) (**Supplementary Figure 1**). The 12 samples with three replicates that were collected from the E+ and E− tall fescue at morning, afternoon, and evening were referred as E+MS (endophyte positive, morning, pseudostem), E+ML (endophyte positive, morning, leaf blade), E+NS (endophyte positive, afternoon, pseudostem), E+NL (endophyte positive, afternoon, leaf blade), E+ES (endophyte positive, evening, pseudostem), E+EL (endophyte positive, evening, leaf blade), E−MS (endophyte-free, morning, pseudostem), E−ML (endophyte-free, morning, leaf blade), E−NS (endophyte-free, afternoon, pseudostem), E−NL (endophyte-free, afternoon, leaf blade), E−ES (endophyte-free, evening, pseudostem), and E−EL (endophyte-free, evening, leaf blade; **Supplementary Figure 2**). In each sample, L/S tissue from 10 genotypes were pooled in a 15 ml tube and immediately frozen in liquid N_2. After arrival to the laboratory, the samples were stored at −80°C until processing.

[1]https://www.pennington.com/

RNA Extraction and Sequencing

The total RNA of each of 36 samples (**Supplementary Figure 2**) was isolated from approximately 100 mg ground tissues using a Spectrum™ Plant Total RNA Kit (Sigma, Cat. No. STRN250, St. Louis, United States) according to the manufacturer recommendation. The RNA quality was measured in Agilent 2,100 Bioanalyzer (Agilent, Santa Clara, United States) using Agilent RNA 6000 Nano Kit (Agilent, 5,067–1,511), and RNA was quantified using Qubit® RNA BR (Broad-Range) Assay Kit (Life Technologies, Cat. No. Q10211, Carlsbad, United States). The RNA samples were then treated with TURBO DNA-free Kit (Invitrogen, Cat. No. AM1907, Carlsbad, United States) following their protocol.[2] RNA samples were then cleaned using the RNeasy MinElute Cleanup Kit (Qiagen, Cat. No. 74204, Hilden, Germany) according to the manufacturer protocol.[3]

RNA-seq libraries were prepared using TruSeq Stranded mRNA Sample Preparation Kit (Illumina, Cat. No. 20020594). Briefly, mRNA was purified from one microgram of total RNA, fragmented, and converted to double-stranded DNA for sequencing. Individual libraries were uniquely indexed using TruSeq RNA CD Indexes (Illumina, Cat. No. 20019792), and pooled in equimolar ratio. The pooled libraries were sequenced on an Illumina NovaSeq 6000 150PE Sequencing system.

Quality Assessment and Assembly of the RNA-Seq Reads

The raw reads of 36 samples (**Supplementary Figure 2**) were quality trimmed to remove any low-quality bases and primer/adapter sequences before performing the assembly using the Trimmomatic (v. 0.36) using default settings (Bolger et al., 2014). Reads less than 30 bases long after trimming were discarded, along with their mate pair. Endophyte-derived reads were identified by mapping the trimmed reads to the *E. coenophiala* transcriptome (Schardl et al., 2013)[4] and successfully mapped reads were excluded from further analysis. The trimmed and filtered reads from each sample were independently *de novo* assembled using the software Trinity (v. 2.8.5) with default parameters (Grabherr et al., 2011). These assemblies were then combined by randomly selecting one as a starting transcriptome and then iteratively aligning the transcriptome with each assembly, identifying that assembly's novel transcripts, and adding those transcripts to the combined transcriptome. Each sample was then mapped to the combined transcriptome using HISAT2 (v. 2.0.5)[5] with 24 threads and the default mapping parameters (Kim et al., 2019). The expressed transcripts in each sample were quantified using the StringTie (v. 1.2.4) with the default assembly parameters to produce more complete and accurate reconstructions of transcripts and better estimates of their expression levels (Pertea et al., 2015).

Identification of Differentially Expressed Genes

To identify genes which expressed under different temperature condition during morning, afternoon, and evening time with or without the presence of endophyte, pairwise differential gene expression testing was performed using DESeq2 with default parameters setting (Love et al., 2014). DESeq2 method was used for differential read counts per gene in RNA-seq, using shrinkage estimation for dispersions and fold changes (FC) to improve stability of estimates across experimental conditions. A log2 FC $\leq$ -5 and $\geq$ 5 and adjusted value of $p \leq 0.05$ were used to determine the significant differences in differential gene expression between two samples. The DEGs with log2 FC with "–" and "+" sign indicates downregulated and upregulated genes, respectively.

Hierarchical Clustering and Visualization of Differentially Expressed Genes

Hierarchical clustering analysis of DEGs from the six comparisons, E+MS vs. E−MS, E+ML vs. E−ML, E+NS vs. E−NS, E+NL vs. E-NL, E+ES vs. E−ES, and E+EL vs. E−EL, was constructed using the function heatmap.2 in the R package gplots (Warnes et al., 2015) in R Studio. The DEGs that were biologically significant were visualized using the web-based software DiVenn (Sun et al., 2019). The red and blue nodes represent up- and downregulated genes, respectively. The yellow nodes represent upregulated in one dataset but downregulated in the other dataset.

Identification of Orthologous Genes Using Tall Fescue Transcripts

As annotation of tall fescue genome is not available till to October 25, 2021, the complete and accurate tall fescue transcripts were aligned against the switchgrass non-redundant protein sequences in Phytozome v13 database[6] using BLASTX searches to identify best-matched switchgrass orthologues. Using switchgrass orthologues, we also obtained rice and *Arabidopsis* orthologues of tall fescue transcripts from the Phytozome database.

Gene Ontology Analysis of Differentially Expressed Genes

We performed GO enrichment analysis using orthologue genes of rice to identify their involvement in biological process (BP), molecular function (MF), and cellular component (CC) categories. To study the influence of endophyte, the DEGs between E+MS vs. E−MS, E+ML vs. E−ML, E+NS vs. E−NS, E+NL vs. E−NL, E+ES vs. E−ES, and E+EL vs. E−EL were used for GO enrichment analysis. The rice orthologues of the tall fescue DEGs were used as input data to perform GO analysis using singular enrichment analysis (SEA) tool against *Oryza sativa japonica* annotation of the web-based AgriGO v2.0 (Tian et al., 2017) with modified statistical parameter settings: statistical

[2]https://assets.fishersci.com/TFS-Assets/LSG/manuals/1907M_turbodnafree_UG.pdf
[3]https://www.qiagen.com
[4]http://csbio-l.csr.uky.edu/ec/
[5]https://daehwankimlab.github.io/hisat2/
[6]https://phytozome.jgi.doe.gov/pz/portal.html

test method, Fisher; multi_test adjustment method, Yekutieli (FDR under dependency); significance level, 0.01; and minimum number of mapping entries,10. MSU7.0 gene ID (TIGR) of rice orthologues was used as reference during SEA analysis.

KEGG Pathway Enrichment Analysis

The KEGG pathway enrichment analysis was performed using KOBAS 3.0 (Xie et al., 2011) on the basis of Fisher's exact test with Benjamini and Hochberg (1995) FDR-corrected value of p <0.05. The top most significant pathways based on FDR-corrected p-values in all six comparisons, E+MS vs. E−MS, E+ML vs. E−ML, E+NS vs. E−NS, E+NL vs. E−NL, E+ES vs. E−ES, and E+EL vs. E−EL were presented.

Prediction of Candidate Genes Responsible for Cold Tolerance in Tall Fescue

To identify candidate genes associated with cold tolerance in tall fescue, we searched genes of each GO term whose function related to stress, abiotic stimulus, ethylene stimulus, DNA damage stimulus, light stimulus, and radiation etc.

Validation of DEGs Using Quantitative Real-Time Reverse Transcription-PCR

Eight DEGs obtained from E+MS vs. E−MS were selected for validation using quantitative real-time reverse transcription-PCR (qRT-PCR) analysis. Total RNA treated with TURBO DNA-free Kit (Invitrogen, Cat. No. AM1907, Carlsbad, United States) was used to synthesize the first-strand cDNA using SuperScript III Reverse Transcriptase (RT) kit (Invitrogen, Carlsbad, CA, United States, Cat. no.: 18080044) following the manufacturer protocol.[7] Briefly, for each RNA sample, the following components were combined in a PCR tube on ice to a volume of 12 µl containing 5 µl DNase-free RNA (200 ng/µL), 1 µl 50 mM oligo $(dT)_{20}$, 1 µl 10 mM dNTPs and 5 µl RNase/DNase-free water. The reaction mixtures were incubated at 65°C for 5 min and then placed on ice for 2 min. The first-strand cDNA synthesis master mix was prepared on ice by adding 4 µl 5x First-Strand buffer, 1 µl 0.1 M DTT, 1 µl RNaseOUT Recombinant RNase Inhibitor (40 units/µL), and 1 µl SuperScript III RT (200 units/µL). The first-strand cDNA synthesis master mix was mixed properly by gentle vortex and was added into the pre-incubated RNA and oligo tube. The reaction mixture was mixed by pipetting up and down and incubated at 50°C for 60 min. The reaction was terminated at 70°C for 15 min and cooled on ice. The cDNA was stored at −20°C.

qRT-PCR reactions were prepared in an optical 384-well plate in a volume of 10 µl containing 2 µl of the forward and reverse primer (1 µM/µL), 5 µl of 2x Sigma KiCqStart SYBR Green qPCR Ready-Mix (Cat no.: KCQS01), 1 µl molecular biology grade water, and 2 µl cDNA (1:20). qRT-PCR amplifications were performed on QuantStudio 7 Flex Real-Time PCR system (Thermo Fisher Scientific, Singapore) using a protocol of 2-step PCR cycle (an initial denaturation of cDNA at 95°C for 3 min, followed by 40 cycles of denaturation at 95°C for 15 s and annealing at 60°C for 45 s) and a 3-step of melting curve analysis (95°C for 15 s, 60°C for 1 min and 95°C for 15 s). Experiments were performed with three technical replicates of each S tissues of E+ and E- tall fescue collected under freezing temperature in the morning. Gene expression was quantified using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001). The tall fescue *Actin* gene was used as an internal reference gene. Primers used to amplify 53–63 bp of the genes were designed using Primer Express software (v3.0.1; Thermo Fisher Scientific; **Supplementary Table 1**) and synthesized by Sigma-Aldrich, MO, United States.

[7]https://www.thermofisher.com

RESULTS

Sequencing and *de novo* Assembly of the Texoma Tall Fescue Transcriptome

After the quality assessment and data filtered, a total of 553.8 million high-quality paired-end reads were identified in the E+ (18) samples and 484.8 million in the E− (18) samples (**Table 1**). The filtered reads were *de novo* assembled into 5,520,386 and 5,133,272 contigs in the E+ and E− samples, respectively for downstream analysis. From these contigs, we identified unique transcripts varied from 186,653 to 200,380 in the E+ samples and from 188,468 to 194,606 in the E− samples. Overall, a total of 226,054 transcripts were identified in the E+ samples and 224,376 transcripts in the E− samples. The length of these transcripts varied from 177 to 27,968 bp and the N50 varied from 1,288-1,326. Finally, we found 234,883 transcripts from all the samples collected in this study (**Table 1**). In addition, the result showed about 5.18 and 0.95% more transcripts were expressed in endophyte-infected S and L, respectively over endophyte-free S and L under freezing temperatures in the morning.

Identification of Orthologue Genes From Other Plant Species

Among the combined transcripts (234,883), about 13.5% got hit to switchgrass (31,780), rice (31,622), and *Arabidopsis* (31,604) genomes. A total of 5,919, 4,476, and 4,002 orthologue genes were identified in switchgrass, rice, and *Arabidopsis*, respectively (**Supplementary Figure 3**). Due to lack of well-annotated tall fescue genome, the function of the majority fescue transcripts remains unknown when compared to the reference genomes of the related species.

Analysis of DEGs Between E+ and E− Texoma Tall Fescue in Different Time Points

Differential gene expression in E+ tall fescue was analyzed relative to E− tissues under three different time points under natural field environment. In total, 5,757 significant DEGs (value of $p \leq 0.05$) with log2 FC ≤ -5 and ≥ 5 were identified at least in one of the six following comparisons: E+MS vs. E−MS, E+ML vs. E−ML, E+NS vs. E−NS, E+NL vs. E−NL E+ES vs. E−ES, and E+EL vs. E−EL (**Figure 1**; **Supplementary Table 2**). Analysis of DEGs showed that higher

TABLE 1 | Assembly and annotation of the transcriptome data obtained from short read Illumina sequencing of Texoma endophyte positive (E+) and endophyte negative (E–) tissues under cold stress in the natural field environment.

Parameters		Texoma E+			Texoma E–			Combined
	Time	Pseudostem	Leaf blade	Pooled	Pseudostem	Leaf blade	Pooled	
Filtered paired-end reads*	M	137,172,033	62,692,906	553,802,405	71,139,077	78,140,815	484,814,092	
	N	87,040,508	71,783,216		92,460,683	73,609,080		
	E	86,632,321	108,481,421		95,417,751	74,046,686		
No of contigs after Trinity assembly*	M	1,125,143	780,760	5,520,386	817,647	841,846	5,133,272	
	N	929,259	769,068		906,092	799,442		
	E	941,726	974,430		930,388	837,857		
No of unique transcript	M	200,380	190,568	226,054	190,514	188,760	224,376	234,883
	N	193,981	186,653		193,370	188,468		
	E	194,588	193,295		194,606	190,710		
Transcript length (bp)	M	177–16,616	179–15,501	177–16,616	177–16,644	178–15,236	177–27,968	177–27,968
	N	178–15,651	177–15,604		177–15,596	178–16,026		
	E	179–15,718	178–15,966		177–15,590	179–27,968		
N50	M	1,316	1,288	1,288–1,338	1,317	1,303	1,300-1,326	1,288–1,326
	N	1,318	1,338		1,318	1,326		
	E	1,317	1,319		1,300	1,304		

*M, morning; N, noon; and E, evening. *Data presented as sum of three biological replicates.*

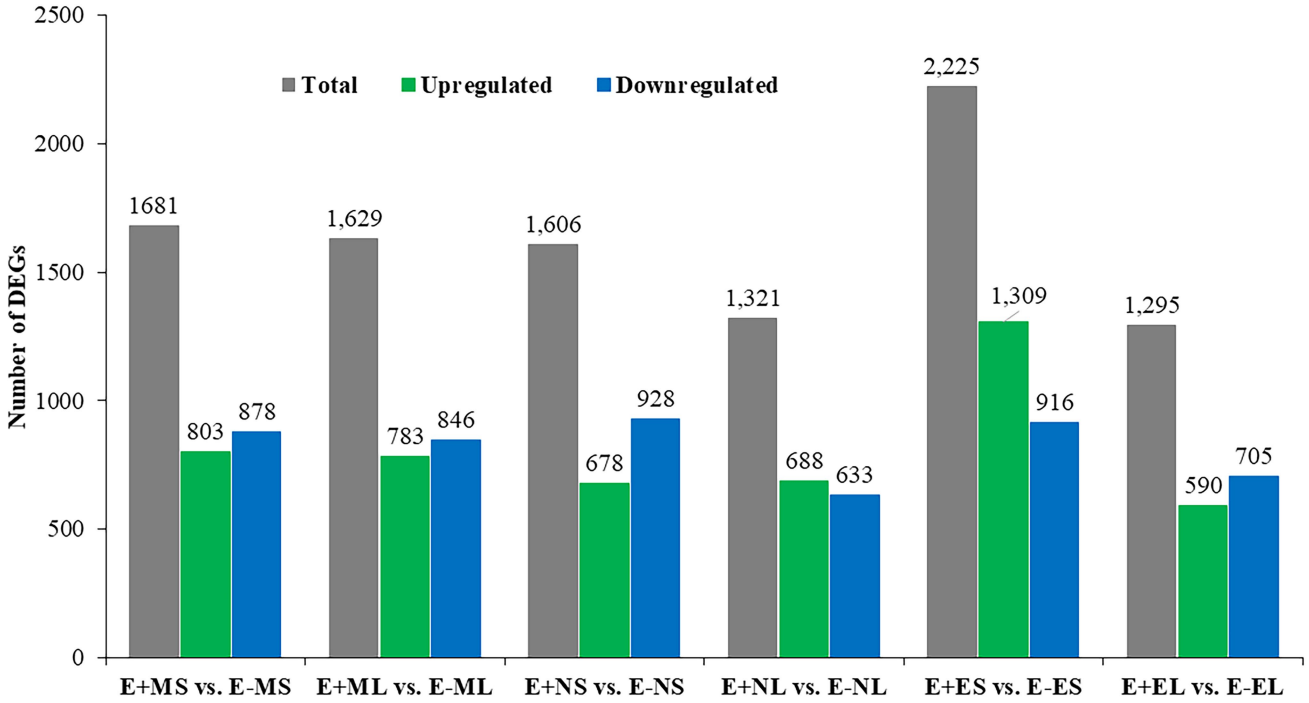


FIGURE 1 | Identification of differentially expressed genes between E+ and E– tissues under cold stress. Samples were: E+MS (endophyte positive, morning, pseudostem), E+ML (endophyte positive, morning, leaf blade), E+NS (endophyte positive, afternoon, pseudostem), E+NL (endophyte positive, afternoon, leaf blade), E+ES (endophyte positive, evening, pseudostem), E+EL (endophyte positive, evening, leaf blade), E–MS (endophyte-free, morning, pseudostem), E–ML (endophyte-free, morning, leaf blade), E–NS (endophyte-free, afternoon, pseudostem), E–NL (endophyte-free, afternoon, leaf blade), E–ES (endophyte-free, evening, pseudostem), and E–EL (endophyte-free, evening, leaf blade).

number of transcripts were expressed in S than L tissues at each time point (**Figure 1**). The upregulated transcripts were detected fewer than the downregulated ones in both tissues under the freezing (−3–0.5°C) temperature (S: 803 up and 878 down; L: 783 up and 846 down) in the morning. At afternoon (11–12°C), fewer upregulated (678) than downregulated (928) transcripts were identified in S tissues, but it was differing in L tissue (688 up- vs. 633 downregulated). In contrast to the afternoon, in the evening (12–10°C), more upregulated (1,309) than downregulated (916) transcripts were identified in S tissues, and fewer upregulated (590) than downregulated (705) transcripts were detected in L tissues (**Figure 1**).

The specific and overlapping DEGs among the comparisons were visualized in DiVenn (**Figure 2**). The result showed 463 DEGs were specific to E+NS vs. E−NS, 321 to E+NL vs. E−NL, 961 to E+ES vs. E−ES, and 470 were specific to E+EL vs. E−EL under normal cold condition in the afternoon and evening time. Among the morning time expressed transcripts, 97 DEGs were common between S (E+MS vs. E−MS) and L (E+ML vs. E−ML), of which 42 were upregulated in one but downregulated in other comparison. In addition, there were 556 DEGs were specific to E+MS vs. E−MS and 529 were specific to E+ML vs. E−ML, totaling of 1,085 were significantly up- and downregulated in the morning freezing conditions, and were not expressed in the normal cold condition in the afternoon and evening time (**Figure 2**).

The DEGs were used for linkage hierarchical clustering analysis (**Figure 3**). A distinct pattern of gene expression at transcriptional level under the three time points was observed. Cluster analysis showed that some genes upregulated in the morning were downregulated in the afternoon and evening time or vice versa. Heat map also showed that the expression profiles of the majority genes were different between the S and L tissues in all the time points (**Figure 3**). This result indicates that tall fescue responded to the stress conditions in time- and tissue-specific manners.

Gene Ontology Analysis of DEGs

Out of 5,757 significant DEGs, 1,099 got hit to 732 rice genes in the six comparisons, E+MS vs. E−MS, E+ML vs. E−ML, E+NS vs. E−NS, E+NL vs. E−NL, E+ES vs. E−ES and E+EL vs. E−EL

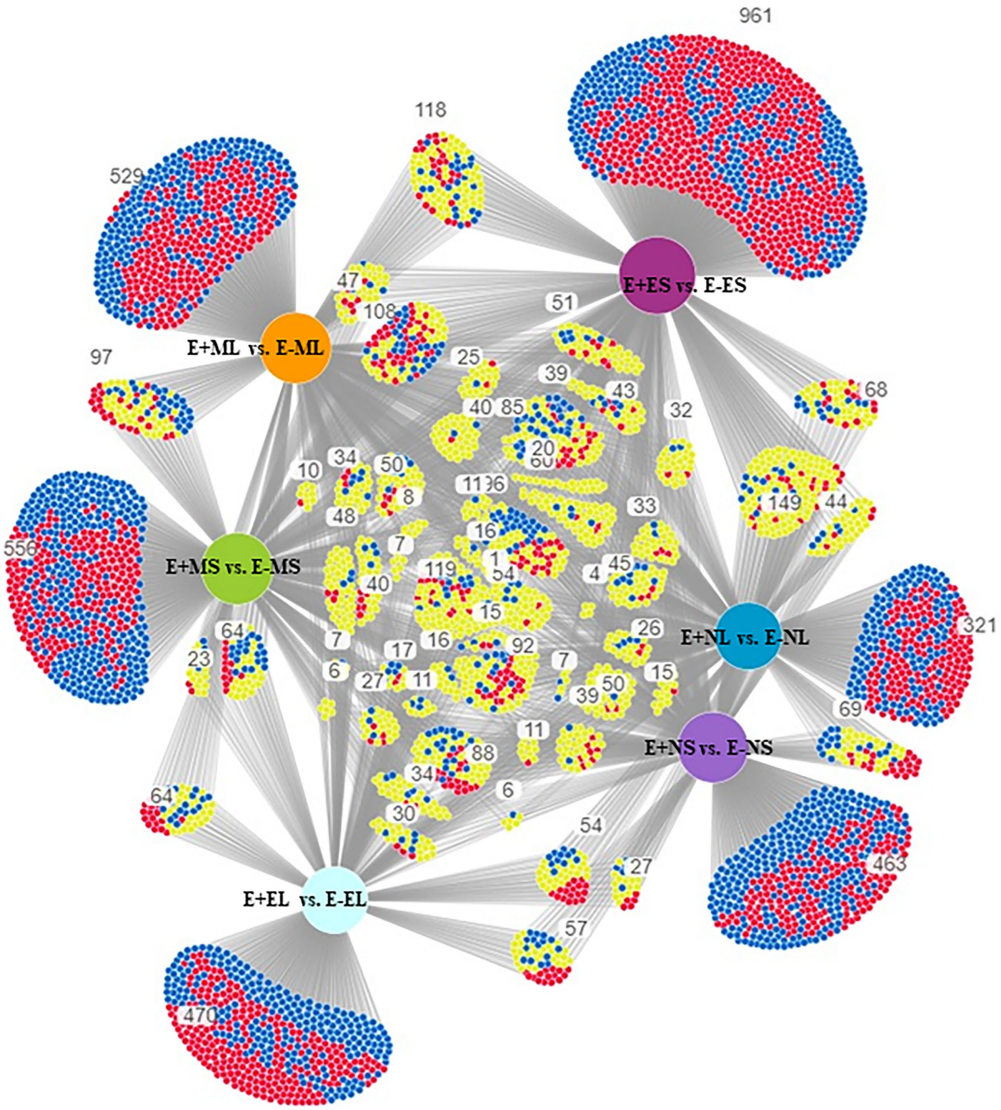


FIGURE 2 | Visualization of differential gene expression pattern among E+MS vs. E−MS, E+ML vs. E−ML, E+NS vs. E−NS, E+NL vs. E−NL, E+ES vs. E−ES, and E+EL vs. E−EL using DiVenn program. The red and blue nodes represent up- and downregulated genes, respectively. The yellow node represents upregulated in one dataset but downregulated in the other dataset. Abbreviations of the samples are the same as in **Figure 1**.

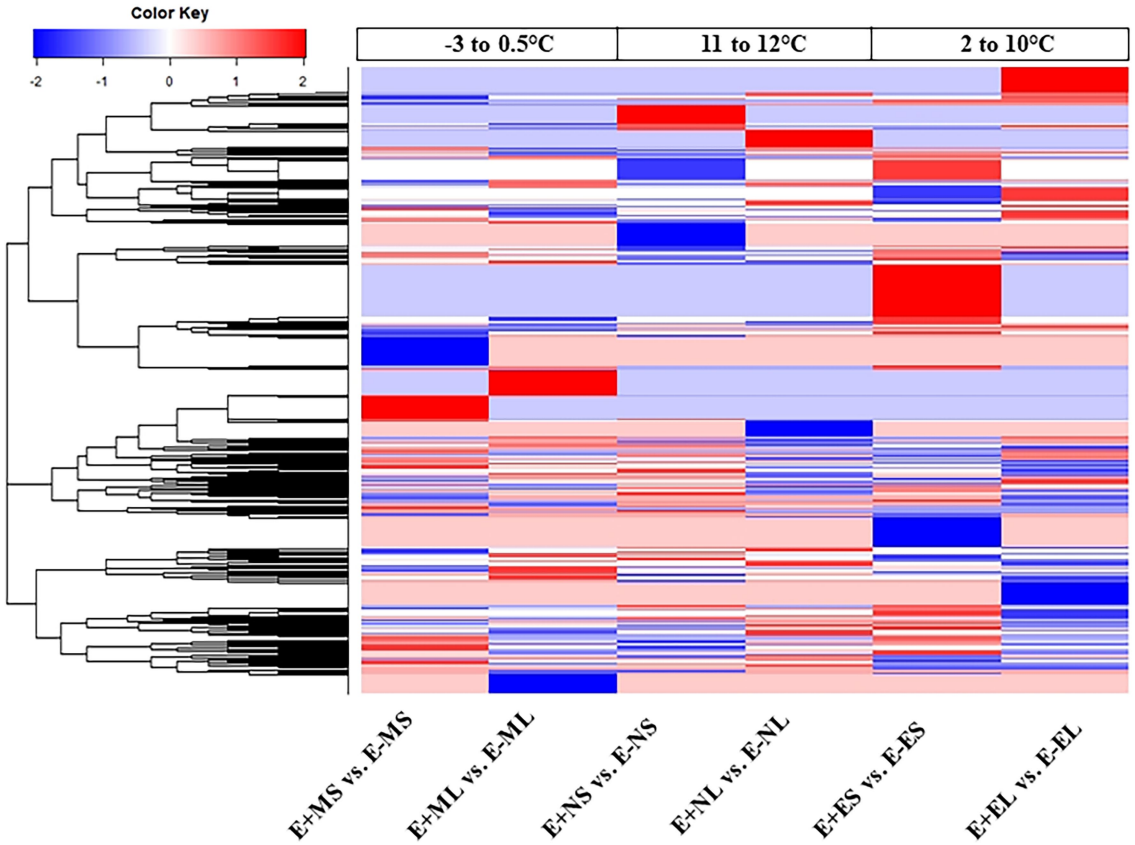


FIGURE 3 | Hierarchical clustering analysis of all differentially expressed genes from six comparisons. Each column represents a different sample with subject to endophyte status, cold stress, and tissue type. Red represents upregulated; blue, downregulated; and white, no change. Abbreviations of the samples are the same as in **Figure 1**.

and were used for GO analysis (**Supplementary Table 3**). We obtained 98 significant GO terms of three major categories, such as BP, MF, and CC (**Supplementary Tables 4**, **5**), and only the top 10 significant GO terms of each of the six comparisons were presented in **Figure 4**. Some of the identified GO terms were often contained common genes. Though our key objective was to identify DEGs under different time points, we tried to understand the function of genes those were expressed under the freezing temperatures (−3–0.5°C) in the morning *via* GO analysis. In E+MS vs. E−MS samples, 38 GO terms under BP, 31 under MF, and 11 under CC categories were identified. The most significant GO terms involved are "protein complex biogenesis" and "protein complex assembly" under BP; "nucleotide binding" and "purine nucleotide binding" under MF; and "membrane coat" and "coated membrane" under CC (**Figure 4**; **Supplementary Figure 4**; **Supplementary Table 5**). There were 86 genes under the "nucleotide binding" category (**Figure 4**), of which four genes: *LOC_Os04g58410.1* (tall fescue gene ID: Fa.36660.1), *LOC_Os03g51030.1* (Fa.8356.1), *LOC_Os03g51030.2* (Fa.8921.1 and Fa.10647.1), *LOC_Os03g54084.1* (Fa.6944.1), are involved in GO:0006974 (cellular response to DNA damage stimulus), GO:0033554 (cellular response to stress), GO:0071214 (cellular response to abiotic stimulus), GO:0071478 (cellular response to radiation), and GO:0071482 (cellular response to light stimulus).

Similarly, in E+ML vs. E−ML, GO analysis detected 21 GO terms under BP, 30 under MF, and nine under CC categories. The most significant GO terms involved are "cellular amino acid metabolic process," "cellular amine metabolic process," and "cellular nitrogen compound metabolic process" under BP; "pyrophosphatase activity," "nucleoside-triphosphatase activity," "hydrolase activity, acting on acid anhydrides, in phosphorus-containing anhydrides," "hydrolase activity, acting on acid anhydrides," and "nucleotide binding" under MF; and "membrane part" under CC category (**Figure 4**; **Supplementary Figure 5**; **Supplementary Table 5**). There were 32 genes under each of the GO terms "pyrophosphatase activity," "nucleoside-triphosphatase activity," "hydrolase activity," and "membrane part," and the other 78 genes under the "nucleotide binding" category (**Figure 4**), of which three genes, *LOC_Os04g32560.2* (Fa.47164.1), *LOC_Os12g12850.2* (Fa.23585.2) and *LOC_Os02g57530.2* (Fa.32995.1) are involved in GO:0033554 (cellular response to stress), GO:0006974 (response to DNA damage stimulus), GO:0071369 (cellular response to ethylene stimulus) and GO:0071495 (cellular response to endogenous stimulus).

Among the top 10 significant GO terms in each of the six comparisons, there was one significant GO term "amino acid activation" under BP and two GO terms "membrane coat" and "coated membrane" under CC, which were enriched only

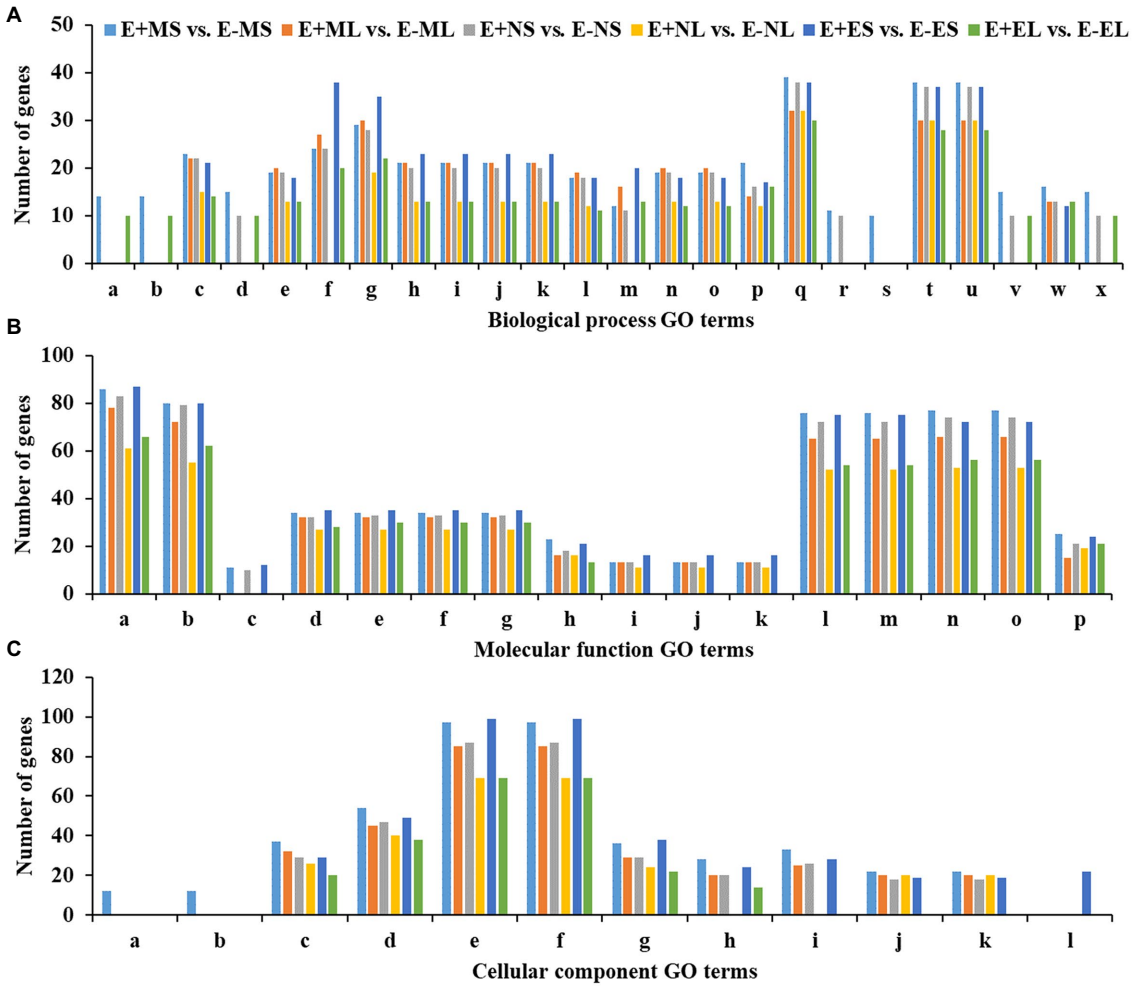


FIGURE 4 | Top 10 significant GO terms in each of the six comparisons obtained using differentially expressed genes of rice orthologues. Significantly ($p \leq 0.01$) enriched GO terms of **(A)**—biological process (BP), **(B)**—molecular function (MF) and **(C)**—cellular component (CC) are shown in *x*-axis and the number of genes of each GO term are displayed in *y*-axis. The GO terms for the biological process are (a) protein complex biogenesis, (b) protein complex assembly, (c) cellular nitrogen compound metabolic process, (d) macromolecular complex assembly, (e) amine metabolic process, (f) carbohydrate metabolic process, (g) small molecule metabolic process, (h) oxoacid metabolic process, (i) organic acid metabolic process, (j) carboxylic acid metabolic process, (k) cellular ketone metabolic process, (l) cellular amino acid metabolic process, (m) cellular carbohydrate metabolic process, (n) cellular amine metabolic process, (o) cellular amino acid and derivative metabolic process, (p) macromolecule localization, (q) localization, (r) tRNA metabolic process, (s) amino acid activation, (t) transport, (u) establishment of localization, (v) cellular component assembly, (w) protein localization, and (x) macromolecular complex subunit organization. The GO terms for the molecular function are: (a) nucleotide binding, (b) purine nucleotide binding, (c) protein tyrosine kinase activity, (d) nucleoside-triphosphatase activity, (e) pyrophosphatase activity, (f) hydrolase activity, acting on acid anhydrides, in phosphorus-containing anhydrides, (g) hydrolase activity, acting on acid anhydrides, (h) ATPase activity, (i) ATPase activity, coupled to transmembrane movement of substances, (j) ATPase activity, coupled to movement of substances, (k) hydrolase activity, acting on acid anhydrides, catalyzing transmembrane movement of substances, (l) adenyl nucleotide binding, (m) purine nucleoside binding, (n) purine ribonucleotide binding, (o) ribonucleotide binding, and (p) substrate-specific transporter activity. The GO terms for the cellular component are: (a) membrane coat, (b) coated membrane, (c) membrane part, (d) membrane, (e) cell, (f) cell part, (g) cytoplasm, (h) protein complex, (i) macromolecular complex, (j) integral to membrane, (k) intrinsic to membrane, and (l) cytoplasm part. Abbreviations of the samples are the same as in **Figure 1**.

E+MS vs. E−MS under the freezing stress (−3–0.5°C) in the morning (**Figure 4**). Most of the genes associated with the three GO categories mentioned above were related to kinase activity, binding, signaling, and transporter activity. Interestingly, there was only one molecular function GO term "lyase activity," which was not among the top 10 GO terms significantly enriched in E+ML vs. E−ML during the freezing stress in the morning than other cold stresses (**Supplementary Table 5**). There were 12 genes under "lyase activity" category, of which one gene, *LOC_Os04g37920.1* (Fa.63716.1) was identified in GO:0033554 (cellular response to stress) and GO:0006974 (response to DNA damage stimulus) categories.

KEGG Pathway Analysis of DEGs

We identified 85 and 90 KEGG pathways at E+MS vs. E−MS and E+ML vs. E−ML, respectively under freezing stress in the morning. Similarly, 81 and 71 pathways were identified at E+NS vs. E−NS and E+NL vs. E−NL, respectively under afternoon

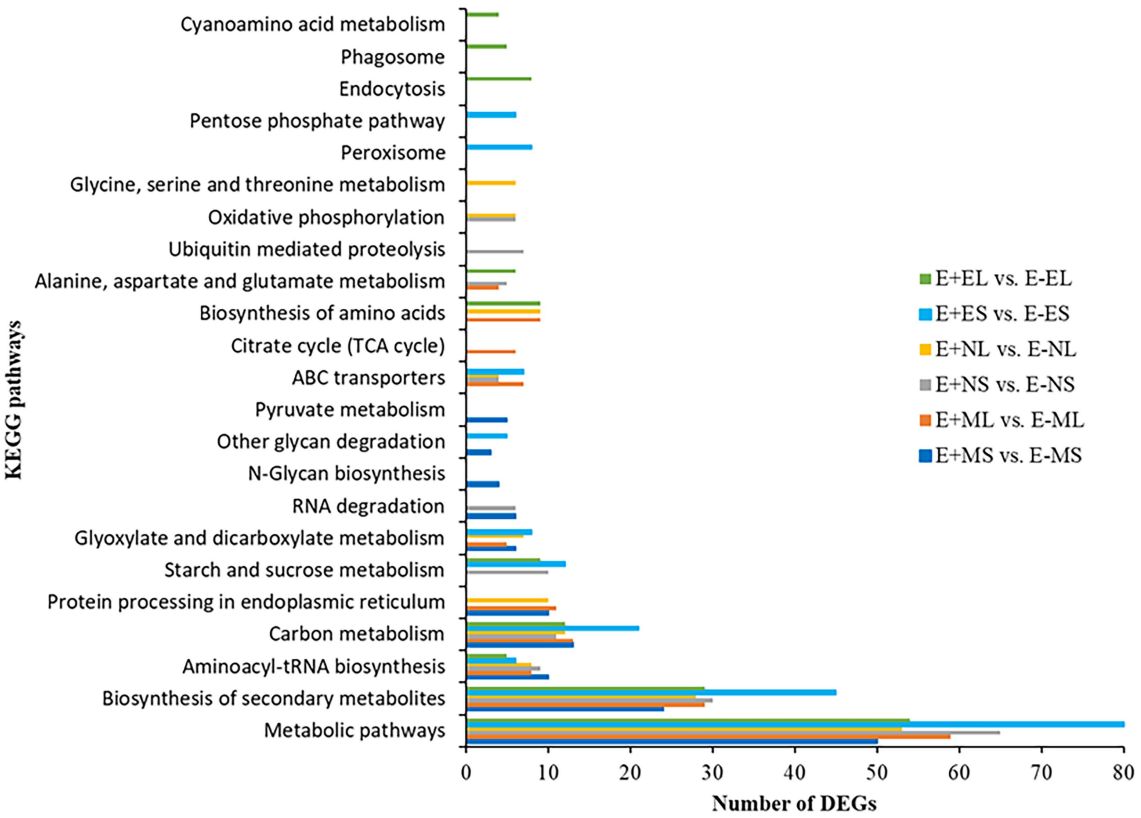


FIGURE 5 | The distribution of differentially expressed genes in KEGG pathways. The top 10 enriched KEGG pathways based on FDR-corrected value of p ($p < 0.05$) in each six comparisons, E+MS vs. E−MS, E+ML vs. E−ML, E+NS vs. E−NS, E+NL vs. E−NL, E+ES vs. E−ES, and E+EL vs. E−EL were displayed in y-axis and the number of DEGs under each KEGG pathways were displayed in x-axis. Abbreviations of the samples are the same as in **Figure 1**.

cold condition, and 86 pathways were identified in both E+ES vs. E−ES and E+EL vs. E−EL under evening cold period (**Supplementary Table 6**). Top 10 KEGG pathways of each six comparisons were shown in **Figure 5**. Among the top 10 enriched pathways, "metabolic pathways (dosa01100)," "biosynthesis of secondary metabolites (dosa01110)," and "carbon metabolism (dosa01200)" were contained maximum number of DEGs. Under different time points, metabolic pathway contained lowest number (50) of DEGs at morning freezing stress (E+MS vs. E−MS) and maximum number (80) at evening cold condition (E+ES vs. E−ES). Similarly, biosynthesis of secondary metabolites contained lowest number (24) of DEGs at morning freezing stress (E+MS vs. E−MS) and maximum number (45) at evening cold condition (E+ES vs. E−ES; **Figure 5**).

Prediction of Ortholog Genes of Rice as a Potential Candidate Responsible for Cold Tolerance in Tall Fescue

In this study, by annotating the gene function through GO analyses, we identified eight candidate genes specifically expressed in the morning, but neither in the afternoon nor in the evening. Four genes, *LOC_Os03g51030.1* (Fa.8356.1), *LOC_Os03g51030.2* (Fa.8921.1 and Fa.10647.1), *LOC_Os03g54084.1* (Fa.6944.1), and *LOC_Os04g58410.1* (Fa.36660.1) were upregulated (+6 to +22 FC) in S tissues in the morning; and the remaining four genes: *LOC_Os04g37920.1* (Fa.63716.1) and *LOC_Os02g57530.2* (Fa.32995.1) were downregulated (−8 to −27 FC), and *LOC_Os12g12850.2* (Fa.23585.2) and *LOC_Os04g32560.2* (Fa.47164.1) were upregulated (+21 to +38 FC) in L tissues in the morning (**Table 2**).

Validation of DEGs Using qRT-PCR

To verify the FC values of DEGs identified by pairwise comparison between E+ and E- tissues, eight DEGs, including seven (Fa6944.1, Fa8356.1, Fa10647.1, Fa36660.1, Fa79541.1, Fa120720.1, and Fa139004.1) upregulated and one (Fa5596.1) downregulated in E+MS vs. E−MS were evaluated using qRT-PCR method. The expression levels of the selected DEGs detected with qRT-PCR showed similar pattern with the expression levels calculated from RNA-Seq data (**Figure 6**), indicating the RNA-seq data were reliable in this study.

DISCUSSION

The objectives of this study were to characterize the tall fescue transcriptome, and to identify DEGs with and without presence of a fungal symbiont in tall fescue under the cold condition

TABLE 2 | List of eight candidate genes in tall fescue expressed under freezing temperatures in the morning.

Tall fescue gene ID	log2 FC*						Important GO terms[†]	Rice orthologs[¥]
	TE + MS vs. TE – MS	TE + ML vs. TE – ML	TE + NS vs. TE – NS	TE + NL vs. TE – NL	TE + ES vs. TE – ES	TE + EL vs. TE – EL		
Fa.63716.1		–26.87					Lyase activity GO:0033554 cellular response to stress GO:0006974 response to DNA damage stimulus	*LOC_Os04g37920.1*
Fa.8356.1	+22.11						Nucleotide binding GO:0071478 cellular response to radiation GO:0071482 cellular response to light stimulus GO:0071214 cellular response to abiotic stimulus	*LOC_Os03g51030.1*
Fa.8921.1 Fa.10647.1	+21.68 +22.42						GO:0071478 cellular response to radiation GO:0071482 cellular response to light stimulus GO:0071214 cellular response to abiotic stimulus	*LOC_Os03g51030.2*
Fa.6944.1	+7.46						GO:0071478 cellular response to radiation GO:0071482 cellular response to light stimulus GO:0071214 cellular response to abiotic stimulus	*LOC_Os03g54084.1*
Fa.36660.1	+6.56						GO:0006974 response to DNA damage stimulus GO:0033554 cellular response to stress	*LOC_Os04g58410.1*
Fa.47164.1		+37.57					Pyrophosphatase activity, nucleoside-triphosphatase activity, hydrolase activity, and nucleotide binding GO:0006974 response to DNA damage stimulus GO:0033554 cellular response to stress	*LOC_Os04g32560.2*
Fa.23585.2		+20.74		–22.57			GO:0006974 response to DNA damage stimulus GO:0033554 cellular response to stress	*LOC_Os12g12850.2*
Fa.32995.1		–7.69					Nucleotide binding and membrane part GO:0071369 cellular response to ethylene stimulus GO:0071495 cellular response to endogenous stimulus	*LOC_Os02g57530.2*

The expression values (log2 fold change) in each comparison, important GO terms and rice orthologs were provided.+, Upregulated; –, Downregulated; FC, Fold change.*

[¥]Rice gene ID was taken from the Rice Genome Annotation Project (rice.uga.edu).

[†]Hydrolase activity, acting on acid anhydrides, in phosphorus-containing anhydrides; and hydrolase-activity, acting on acid anhydrides referred as hydrolase activity.

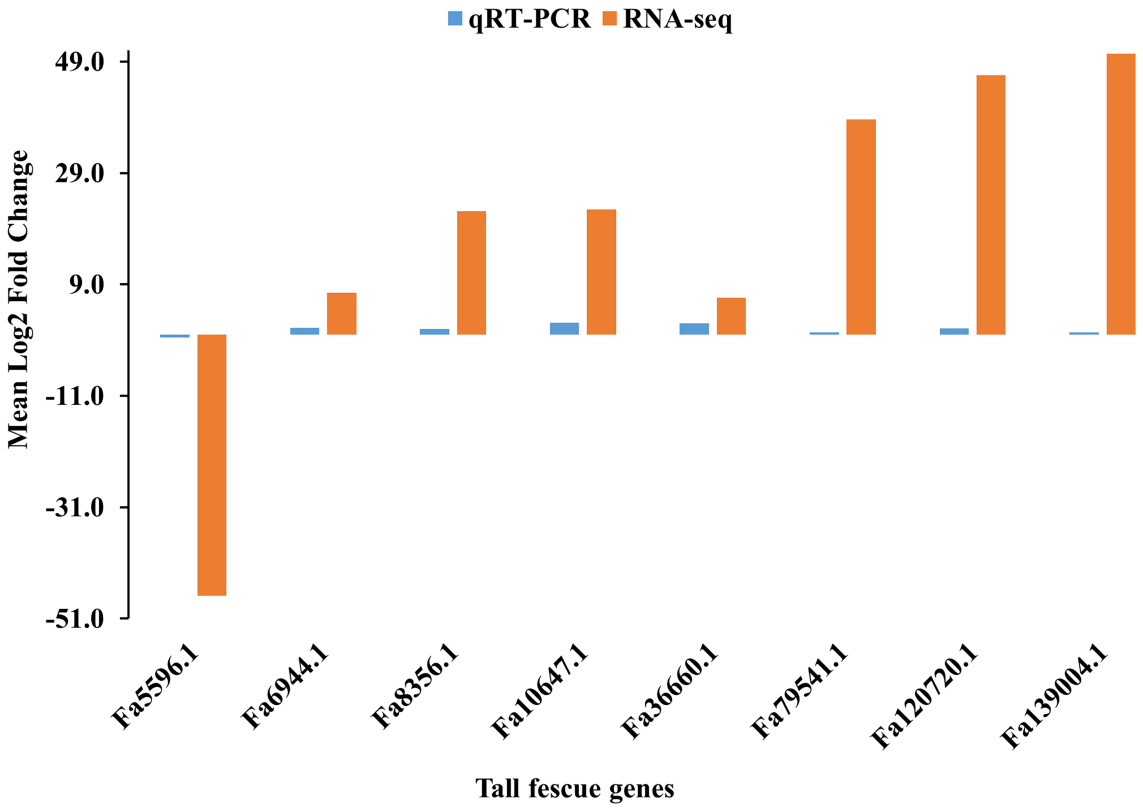


FIGURE 6 | Validation of RNA-seq data obtained from pseudostem of E+ and E– tall fescue collected at morning freezing temperatures through qRT-PCR analysis. The gene expression pattern was shown in *x*-axis and the log2 fold change value was shown in *y*-axis.

in the natural field environment. Tall fescue is an obligate outcrossing plants such that the established cultivar, Texoma E+ and E- in our experiment is a breeding population with a heterogeneous genetic makeup. Thus, a large sampling size is desirable to accurately capture change in gene expression under the experimental condition. Although the ideal sample size for a large-scale genomic experiment may be debatable, 25–30 individuals per population have been estimated to be adequate for population genetic studies using microsatellite markers (Hale et al., 2012). For each biological replicate in our experiment, we collected tissue samples from 10 individual tall fescue and pooled prior to RNA extraction. Thus, each sample group included 30 plants collectively, which would be sufficient for gene expression profiling. Our study is in agreement with (Schunter et al., 2014), where they suggested that seven individuals can be used for transcriptome sequencing in a non-model fish.

Due to lack of a well-annotated tall fescue reference genome, we generated 36 *de novo* assemblies, consist of S and L tissues of E+ and E- Texoma tall fescue genotypes at three time points and three replications (**Table 1**). Individual assembly were performed to keep right track of the DEGs in the E+ and E- Texoma genotypes under different cold conditions in the field as well as to know the transcript abundance in the individual samples (Li et al., 2018). Our transcriptome assemblies result in 234,883 transcripts, which may constitute important transcriptomic resources for understanding cold tolerance mechanism of this allohexaploid forage species. In a previous transcriptome study, *de novo* assembly obtained 199,399 contigs from novel endophyte (AR584) infected two tall fescue genotypes under water stress condition in a greenhouse study using the Illumina Genome Analyzer IIx system (Talukder et al., 2015). Recently, Dinkins et al. (2019) generated transcriptome resources from two tall fescue genotypes infected with common toxic endophyte, one with non-toxic strain (NTE19) and the other with hybrid endophyte species (FaTG-4) under water deficit condition in the greenhouse, and assembled against a tall fescue TF153K transcriptome assembly (Dinkins et al., 2017). Both studies are performed in controlled condition in the greenhouse, but we performed this transcriptomic study of AR584 infected Texoma tall fescue at natural field environment. The field condition is always more variable than that of the controlled growth chamber, due to the direct effect of sunlight, day length, soil microbial community, and genotype–environment interaction but can provide more naturalistic outcome. Among the complex interaction in the field, we considered cold stress is the main external environmental factor that contribute to alter gene expression profile in plant when the plants are in stressed.

By comparing the transcript abundance within E+ and E– Texoma tall fescue, we observed that the number of unique

transcripts were higher in S than that of L tissue in all three different temperatures (**Figure 1**). This may be due to the reason that the endophyte colonizes only in the pseudostem but not in leaf blade of continental tall fescue (Takach et al., 2012). More importantly, our results showed that novel endophyte had positive influence on gene expression over E- tall fescue under freezing temperatures (−3.0–0.5°C) in the morning (E+MS: 200,380 > E−MS: 190,514; E+ML: 190,568 > E−ML: 188,760) when the plants were in stress (**Table 1**). The number of unique transcripts were almost similar in S tissues, but slightly different in L tissues between E+ and E− tall fescue in the afternoon and evening temperature (10–12°C). The higher number of transcripts obtained in E+ over E− samples in all tissues examined under three temperatures conditions in this study (**Table 1**) might be due to endophyte's response. Thus, we considered that the plant does not need support from endophyte under normal cold condition for this cool-season grass species, but does need support to perform better under freezing stresses by altering their gene expression. However, previous studies reported that presence of endophyte does not improve the freezing tolerance of perennial ryegrass (Heineck et al., 2018) and the absence of endophyte does not reduce freezing tolerance of tall fescue (Casler and van Santen, 2008) without transcriptome analyses. Similarly, Dinkins et al. (2017) reported that the presence and/or absence of endophyte do not change global expression. Our transcriptome analyses suggest that there is a positive association between freezing tolerance and endophyte infection. This association would provide further insight into the relationship between endophyte infection and its host fitness to abiotic stresses.

DiVenn showed that 1,085 DEGs were specifically expressed only in the morning freezing conditions (**Figure 2**). Under morning freezing temperatures, plants triggered genes in response to extreme cold stress that was evidenced in GO analyses where three out of top 10 significant GO terms in each of the six comparisons (one under BP and two under CC category) were only found in the morning time (**Figure 4**; **Supplementary Table 5**). However, we did not able to analyze all the DEGs due to lack of available information in tall fescue genome and orthologous genes in related species.

The identification of 1,085 DEGs in Texoma tall fescue under morning freezing condition has led us to identify candidate genes responsible for plant survival under cold stress. As tall fescue is a non-model plant species (Talukder et al., 2021), we performed functional annotation of DEGs those got hit to rice ortholog genes through GO analysis and identified eight orthologous genes of rice as candidate genes for cold tolerance in tall fescue. These eight orthologous candidate genes were expressed in the morning freezing temperature only, and are involved in cellular response to stress, cellular response to abiotic, ethylene, and light stimulus and could be a possible route for extreme cold tolerance in tall fescue (**Table 2**).

Among them, the candidate gene, phytochrome A (*LOC_Os03g51030.1*/Fa.8356.1), phytochrome A (*LOC_Os03g51030.2*/Fa.8921.1 and Fa.10647.1), and phytochrome C (*LOC_Os03g54084.1*/Fa.6944.1) were only upregulated in S under morning freezing temperatures (**Table 2**; **Supplementary Table 3**). Light plays an important role in cold acclimation by accelerating the expression of *CORs* in different species and phytochromes have been identified as important factors in the transcriptional control of *CORs* (Lindlöf, 2010). Three types of phytochromes, such as phytochrome A, phytochrome B, and phytochrome C, were identified in the flowering plants (Agrama and Moussa, 1996) that respond to radiation in the environment. In tomato, phytochrome A perceived far-red light to positively and phytochrome B perceived red light to negatively regulate cold tolerance through abscisic acid-dependent jasmonate signaling (Wang et al., 2016). In rice, phytochrome B negatively controls cold tolerance by regulating *OsDREB1* gene expression through phytochrome interacting factor-like protein OsPIL16 (He et al., 2016). The phytochrome C as an essential light receptor for photoperiodic flowering in a temperate grass model *Brachypodium distachyon* (Woods et al., 2014).

The C-terminal small MutS-related (SMR) domain containing protein (*LOC_Os04g58410.1*/Fa.36660.1) identified from the MF GO term "nucleotide binding," involves in repair of damaged DNA and response to stress, was expressed in S tissue in the morning freezing temperatures (**Table 2**). The pentatricopeptide repeat (PPR) protein family contain a SMR domain in plants (Liu et al., 2013), and that MutS proteins repair the damaged DNA bases produced during DNA replication processes (Sachadyn, 2010). Recently, Zhang and Lu (2019) reported that the PPR-SMR proteins played an important roles in chloroplast biogenesis and gene expression and retrograde signaling between chloroplast to nucleus. It has been believed that retrograde signaling is necessary when the chloroplast is under stressed (Crawford et al., 2017).

Two other orthologous genes, such as *LOC_Os04g32560.2*/Fa.47164.1 and *LOC_Os12g12850.2*/Fa.23585.2, referred as the ATP-dependent Clp protease ATP-binding subunit clpA homolog CD4B putatively expressed in the chloroplast were upregulated in L tissues under morning freezing temperature. Okuzaki et al. (2021) reported that the acidic domain of the chloroplast ribonucleoprotein 31A (CP31A) is essential for cold tolerance in *Arabidopsis.* Under extreme low temperatures, plant growth and development were ceased and thus plant adjusted to different physiological and biochemical processes in response to cold stress. Exogenous ethylene level altered under cold stress. The putatively expressed ethylene receptor (*LOC_Os02g57530.2*/Fa.32995.1) was downregulated in L tissues under morning freezing temperature. Ethylene levels are negatively correlated with cold tolerance in *Medicago truncatula* (Zhao et al., 2014), but positively affect cold tolerance of tomato (*Lycopersicon esculentum;* Ciardi et al., 1997).

We identified an orthologous gene of rice, FAD binding domain of DNA photolyase domain containing protein (*LOC_Os04g37920.1*/Fa.63716.1), involves in cellular response to stress and repair DNA damage, is similar to cryptochrome (CRY) 1b (Li et al., 2021). CRY receives blue light and causes photomorphogenic responses, for example, cotyledon expansion, prevent hypocotyl elongation, and increased anthocyanin accumulation in plants (Ahmad and Cashmore, 1993). The

overexpression of *OsCRY1b* results in short leaf sheath and leaf blade phenotype in rice (Zhang et al., 2006). But, the *CRY1b* (*LOC_Os04g37920.1*/Fa.63716.1) was under expressed (−27 FC) in L tissues under morning temperature in tall fescue. This study would be very useful to develop hypothesis that can bring further understanding of underlying genetics of cold tolerance in tall fescue. By validating the eight DEGs through qRT-PCR analysis, we confirmed that we profiled the expression of DEGs of tall fescue accurately under three different time points in cold condition (**Figure 6**).

CONCLUSION

This study represents the first transcriptome analysis of E+ and E− Texoma tall fescue under freezing and chilling temperatures in the natural field environment. We generated 234,883 unique transcripts from 36 *de novo* assemblies. A total of 5,757 DEGs were identified between E+ and E− samples under three diurnal temperature conditions, of which 1,085 were only up- or downregulated under freezing temperatures in the morning. We were not able to analyze all the genes expressed differentially in two tissues under three temperature conditions, due to lack of available information in related species. Using GO analysis, eight candidate genes were identified from E+ vs. E− samples collected during morning freezing temperature that might help to understand the genetic basis of freezing tolerance in tall fescue. Moreover, the transcriptomic resources generated in this study would serve as valuable resources for grass breeders and to the research community for further structural annotation of tall fescue genome.

DATA AVAILABILITY STATEMENT

The original research presented in this study is publicly available. The raw RNA-seq datasets can be found online in the National Center for Biotechnology Information (NCBI)'s Sequence Read Archive (SRA) database [https://www.ncbi.nlm.nih.gov/sra] under the accession number: [PRJNA734807].

AUTHOR CONTRIBUTIONS

MSI and MS conceived the study. MSI conducted the experiment and wrote the manuscript. TK prepared the library for RNA-seq. MSI and NK analyzed RNA-seq data. MSI and GL prepared RNA for sequencing and performed qRT-PCR. MS critically revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research work was funded by the Noble Research Institute headquarter in Ardmore, Oklahoma, United States.

ACKNOWLEDGMENTS

We extended our gratitude to Josh Barbour, Melissa McMahon, and Konstantin Chekhovskiy for their help in plant tissue collection and processing. We also thank the Perennial Grass Breeding Group for giving access to collect plant materials from their field experiment and Forage Analysis Core of the Noble Research Institute for confirming endophyte status of the genotypes.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.803400/full#supplementary-material

Supplementary Figure 1 | Field conditions during sample collection in the morning, afternoon, and evening time. Air temperature data during sampling time was collected from local Mesonet.

Supplementary Figure 2 | Experiment design of the transcriptomic study. At each collection time, tissues were collected from three replicates.

Supplementary Figure 3 | BLASTx hit of tall fescue transcripts to reference genomes of other plant species. X-axis indicates plant species to identify orthologous genes, and Y-axis indicates number of transcripts/orthologous genes of the respective plant species.

Supplementary Figure 4 | Number of significant GO terms under morning freezing stress in E+ versus E− pseudostem tissue. **A**- biological process, **B**- molecular function, and **C**- cellular component GO terms.

Supplementary Figure 5 | Number of significant GO terms under morning freezing stress in E+ versus E− leaf blade tissue. **A**- biological process, **B**- molecular function, and **C**- cellular component GO terms.

Supplementary Table 1 | List of the primers used in the qRT-PCR experiment.

Supplementary Table 2 | List of differentially expressed genes identified in six comparisons, E+MS vs. E-MS, E+ML vs. E-ML, E+NS vs. E-NS, E+NL vs. E-NL, E+ES vs. E-ES and E+EL vs. E-EL. Differentially expressed genes are represented as log2 transformed values. Genes were considered significantly differentially expressed between E+ and E- samples when it passed the significance threshold requiring the adjusted p-value <0.05 and log2 fold change difference ≥ 5. Rice gene ID was taken from the Rice Genome Annotation Project (rice.uga.edu).

Supplementary Table 3 | List of rice ortholog genes corresponds to differentially expressed genes 732 identified in six comparisons, E+MS vs. E-MS, E+ML vs. E-ML, E+NS vs. E-NS, E+NL vs. E-NL, E+ES vs. E-ES and E+EL vs. E-EL. Gene name, log2 fold change, rice gene symbol and their description were provided here. Rice gene ID was taken from the Rice Genome Annotation Project (rice.uga.edu).

Supplementary Table 4 | Number and identification of significant GO terms (FDR ≤ 0.05) identified in six comparisons using rice orthologues.

Supplementary Table 5 | List of significant GO terms (FDR ≤ 0.05) identified in six comparisons using rice orthologs. GO term name, ontology type, description of the GO term, gene number in input list, gene number in background/reference list, p-value, and FDR value of each GO term were provided.

Supplementary Table 6 | List of the significant KEGG pathways (FDR corrected p-value <0.05).

REFERENCES

Agrama, H. A. S., and Moussa, M. E. (1996). Mapping QTLs in breeding for drought tolerance in maize (*Zea mays* L.). *Euphytica* 91, 89–97. doi: 10.1007/bf00035278

Ahmad, M., and Cashmore, A. R. (1993). HY4 gene of *A. thaliana* encodes a protein with characteristics of a blue-light photoreceptor. *Nature* 366, 162–166. doi: 10.1038/366162a0

Bacon, C. W., and Siegel, M. R. (1988). Endophyte parasitism of tall fescue. *J. Prod. Agric.* 1, 45–55. doi: 10.2134/jpa1988.0045

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *JR. Stat. Soc.* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Buckner, R. C., Powell, J. B., and Frakes, R. V. (1979). "Historical development," in *Tall Fescue*. eds. C. B. Robert and P. B. Lowell (United States: American science agronomy), 1–8.

Casler, M. D., and van Santen, E. (2008). Fungal endophyte removal does not reduce cold tolerance of tall fescue. *Crop Sci.* 48, 2033–2039. doi: 10.2135/cropsci2007.11.0615

Charlton, N. D., Craven, K. D., Afkhami, M. E., Hall, B. A., Ghimire, S. R., and Young, C. A. (2014). Interspecific hybridization and bioactive alkaloid variation increases diversity in endophytic *Epichloë* species of *Bromus laevipes*. *FEMS Microbiol. Ecol.* 90, 276–289. doi: 10.1111/1574-6941.12393

Chen, S., Huang, X., Yan, X., Liang, Y., Wang, Y., Li, X., et al. (2013). Transcriptome analysis in sheepgrass (*Leymus chinensis*): a dominant perennial grass of the Eurasian steppe. *PLoS One* 8:e67974. doi: 10.1371/journal.pone.0067974

Ciardi, J. A., Deikman, J., and Orzolek, M. D. (1997). Increased ethylene synthesis enhances chilling tolerance in tomato. *Physiol. Plant.* 101, 333–340. doi: 10.1111/j.1399-3054.1997.tb01005.x

Crawford, T., Lehotai, N., and Strand, Å. (2017). The role of retrograde signals during plant stress responses. *J. Exp. Bot.* 69, 2783–2795. doi: 10.1093/jxb/erx481

Dinkins, R. D., Nagabhyru, P., Graham, M. A., Boykin, D., and Schardl, C. L. (2017). Transcriptome response of *Lolium arundinaceum* to its fungal endophyte *Epichloë coenophiala*. *New Phytol.* 213, 324–337. doi: 10.1111/nph.14103

Dinkins, R. D., Nagabhyru, P., Young, C. A., West, C. P., and Schardl, C. L. (2019). Transcriptome analysis and differential expression in tall fescue harboring different endophyte strains in response to water deficit. *Plant Genome* 12:71. doi: 10.3835/plantgenome2018.09.0071

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883

Hale, M. L., Burg, T. M., and Steeves, T. E. (2012). Sampling for microsatellite-based population genetic studies: 25 to 30 individuals per population is enough to accurately estimate allele frequencies. *PLoS One* 7:e45170. doi: 10.1371/journal.pone.0045170

He, Y., Li, Y., Cui, L., Xie, L., Zheng, C., Zhou, G., et al. (2016). Phytochrome B negatively affects cold tolerance by regulating *OsDREB1* gene expression through phytochrome interacting factor-like protein OsPIL16 in rice. *Front. Plant Sci.* 7:1963. doi: 10.3389/fpls.2016.01963

Heineck, G. C., Watkins, E., and Ehlke, N. J. (2018). The fungal endophyte *Epichloë festucae* var. *lolii* does not improve the freezing tolerance of perennial ryegrass. *Crop Sci.* 58, 1788–1800. doi: 10.2135/cropsci2017.12.0731

Interrante, S. M., Stein, J. D., Trammell, M. A., Webb, S. L., and Butler, T. J. (2020). Tall fescue germplasm response to freezing temperatures. *Agri. Environ. Lett.* 5:e20017. doi: 10.1002/ael2.20017

Kim, D., Paggi, J. M., Park, C., Bennett, C., and Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* 37, 907–915. doi: 10.1038/s41587-019-0201-4

Li, Q., He, Y., Tu, M., Yan, J., Yu, L., Qi, W., et al. (2018). Transcriptome sequencing of two Kentucky bluegrass (*Poa pratensis* L.) genotypes in response to heat stress. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca* 47, 328–338. doi: 10.15835/nbha47111365

Li, C., Wang, X., Zhang, L., Zhang, C., Yu, C., Zhao, T., et al. (2021). OsBIC1 directly interacts with OsCRYs to regulate leaf sheath length through mediating GA-responsive pathway. *Int. J. Mol. Sci.* 23:287. doi: 10.3390/ijms23010287

Lindlöf, A. (2010). Interplay between low-temperature pathways and light reduction. *Plant Signal. Behav.* 5, 820–825. doi: 10.4161/psb.5.7.11701

Liu, S., Melonek, J., Boykin, L. M., Small, I., and Howell, K. A. (2013). PPR-SMRs: ancient proteins with enigmatic functions. *RNA Biol.* 10, 1501–1510. doi: 10.4161/rna.26172

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2-\Delta\Delta CT$ method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8

Marco-Puche, G., Lois, S., Benítez, J., and Trivino, J. C. (2019). RNA-Seq perspectives to improve clinical diagnosis. *Front. Genet.* 10:1152. doi: 10.3389/fgene.2019.01152

Okuzaki, A., Rühle, T., Leister, D., and Schmitz-Linneweber, C. (2021). The acidic domain of the chloroplast RNA binding protein CP31A is supporting cold-tolerance of *Arabidopsis thaliana*. *J. Exp. Bot.* 72, 4904–4914. doi: 10.1093/jxb/erab165

Peleg, Z., and Blumwald, E. (2011). Hormone balance and abiotic stress tolerance in crop plants. *Curr. Opin. Plant Biol.* 14, 290–295. doi: 10.1016/j.pbi.2011.02.001

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295. doi: 10.1038/nbt.3122

Rogers, J. K., and Locke, J. M. (2013). *Tall Fescue: History, Application, Establishment and Management*. (Ardmore, USA: Noble Research Institute).

Sachadyn, P. (2010). Conservation and diversity of MutS proteins. *Mutat. Res.* 694, 20–30. doi: 10.1016/j.mrfmmm.2010.08.009

Schardl, C. L., Young, C. A., Hesse, U., Amyotte, S. G., Andreeva, K., Calie, P. J., et al. (2013). Plant-symbiotic fungi as chemical engineers: multi-genome analysis of the Clavicipitaceae reveals dynamics of alkaloid loci. *PLoS Genet.* 9:e1003323. doi: 10.1371/journal.pgen.1003323

Schunter, C., Garza, J. C., Macpherson, E., and Pascual, M. (2014). SNP development from RNA-seq data in a nonmodel fish: how many individuals are needed for accurate allele frequency prediction? *Mol. Ecol. Resour.* 14, 157–165. doi: 10.1111/1755-0998.12155

Shi, Y., Ding, Y., and Yang, S. (2015). Cold signal transduction and its interplay with phytohormones during cold acclimation. *Plant Cell Physiol.* 56, 7–15. doi: 10.1093/pcp/pcu115

Sun, L., Dong, S., Ge, Y., Fonseca, J. P., Robinson, Z. T., Mysore, K. S., et al. (2019). DiVenn: An interactive and integrated web-based visualization tool for comparing gene lists. *Front. Genet.* 10:421. doi: 10.3389/fgene.2019.00421

Takach, J. E., Mittal, S., Swoboda, G. A., Bright, S. K., Trammell, M. A., Hopkins, A. A., et al. (2012). Genotypic and chemotypic diversity of *Neotyphodium* endophytes in tall fescue from Greece. *Appl. Environ. Microbiol.* 78, 5501–5510. doi: 10.1128/aem.01084-12

Talukder, S. K., Azhaguvel, P., Mukherjee, S., Young, C. A., Tang, Y., Krom, N., et al. (2015). *De novo* assembly and characterization of tall fescue transcriptome under water stress. *The plant. Genome* 8:50. doi: 10.3835/plantgenome2014.09.0050

Talukder, S. K., Islam, M. S., Krom, N., Chang, J., and Saha, M. C. (2021). Drought responsive putative marker-trait association in tall fescue as influenced by the presence of a novel endophyte. *Front. Plant Sci.* 12:797. doi: 10.3389/fpls.2021.729797

Tian, T., Liu, Y., Yan, H., You, Q., Yi, X., Du, Z., et al. (2017). agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res.* 45, W122–W129. doi: 10.1093/nar/gkx382

Waller, J. C. (2009). "Endophyte effects on cattle," in *Tall Fescue for the Twenty-first Century*. eds. D. B. Hannaway, C. B. West and H. Fribourg (United States: ASA), 289–310.

Wang, F., Guo, Z., Li, H., Wang, M., Onac, E., Zhou, J., et al. (2016). Phytochrome A and B function antagonistically to regulate cold tolerance via abscisic acid-dependent jasmonate signaling. *Plant Physiol.* 170, 459–471. doi: 10.1104/pp.15.01171

Warnes, G.R., Bolker, B., Bonebakker, L., Gentleman, R., Liaw, W., Lumley, T., et al. (2015). Gplots: various R programming tools for plotting data. *The comprehensive R archive Network*

Woods, D. P., Ream, T. S., Minevich, G., Hobert, O., and Amasino, R. M. (2014). Phytochrome C is an essential light receptor for photoperiodic flowering in the temperate grass, *Brachypodium distachyon*. *Genetics* 198, 397–408. doi: 10.1534/genetics.114.166785

Xie, C., Mao, X., Huang, J., Ding, Y., Wu, J., Dong, S., et al. (2011). KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res.* 39(Suppl. 2), W316–W322. doi: 10.1093/nar/gkr483

Zhang, C. (2009). *Identification and Characterization of Cold-Responsive Genes in Perennial Ryegrass*. Doctor of Philosophy Graduate Theses and Dissertations Iowa State University, Ames, Iowa.

Zhang, Y.-C., Gong, S.-F., Li, Q.-H., Sang, Y., and Yang, H.-Q. (2006). Functional and signaling mechanism analysis of rice cryptochrome 1. *Plant J.* 46, 971–983. doi: 10.1111/j.1365-313X.2006.02753.x

Zhang, Y., and Lu, C. (2019). The enigmatic roles of PPR-SMR proteins in plants. *Adv. Sc.* 6:1900361. doi: 10.1002/advs.201900361

Zhang, L., Ren, J., Tong, L., Aide, W., and Tan, D. (2016). *De novo* transcriptome sequencing of cold-treated Kentucky bluegrass (*Poa pratensis*) and analysis of the genes involved in cold tolerance. *J. Hort.* 3:182. doi: 10.4172/2376-0354.1000182

Zhao, M., Liu, W., Xia, X., Wang, T., and Zhang, W.-H. (2014). Cold acclimation-induced freezing tolerance of *Medicago truncatula* seedlings is negatively regulated by ethylene. *Physiol. Plant.* 152, 115–129. doi: 10.1111/ppl.12161

Conflict of Interest: MSI, MS, NK, TK, and GL were employed by Noble Research Institute LLC.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 06 July 2022
doi: 10.3389/fpls.2022.804630

Contrasting Water Withholding Responses of Young Maize Plants Reveal Link Between Lipid Peroxidation and Osmotic Regulation Corroborated by Genetic Analysis

Vlatko Galić[1†], Selma Mlinarić[2†], Matea Marelja[2], Zvonimir Zdunić[1,3], Andrija Brkić[1], Maja Mazur[1], Lidija Begović[2] and Domagoj Šimić[1,3]*

[1] Department of Maize Breeding and Genetics, Agricultural Institute Osijek, Osijek, Croatia, [2] Department of Biology, Josip Juraj Strossmayer University of Osijek, Osijek, Croatia, [3] Centre of Excellence for Biodiversity and Molecular Plant Breeding (CroP-BioDiv), Zagreb, Croatia

Linking biochemistry and genetics of tolerance to osmotic stress is of interest for understanding plant adaptations to unfavorable conditions. The aims of this study were to investigate the variability in responses of panel of elite maize inbred lines to water withholding for stress-related traits through association study and to identify pathways linked to detected associations for better understanding of maize stress responses. Densely genotyped public and expired Plant Variety Protection Certificate (ex-PVP) inbred lines were planted in controlled conditions (16-h/8-h day/night, 25°C, 50% RH) in control (CO) and exposed to 10-day water withholding (WW). Traits analyzed were guaiacol peroxidase activity (GPOD), total protein content (PROT), lipid peroxidation (TBARS), hydrogen peroxide accumulation (H_2O_2), proline accumulation (proline), and current water content (CWC). Proline accumulation was found to be influenced by H_2O_2 and TBARS signaling pathways acting as an accumulation-switching mechanism. Most of the associations detected were for proline (29.4%) and TBARS (44.1%). Gene ontology (GO) enrichment analysis showed significant enrichment in regulation of integral membrane parts and peroxisomes along with regulation of transcription and polysaccharide catabolism. Dynamic studies involving inbreds with extreme phenotypes are needed to elucidate the role of this signaling mechanism in regulation of response to water deficit.

Keywords: GWAS, ontology, lipid peroxidation, proline, signaling, abiotic stress

OPEN ACCESS

Edited by:
Prasanta Kumar Subudhi, Louisiana State University, United States

Reviewed by:
Suravoot Yooyongwech, Mahidol University, Thailand
Ratna Karan, University of Florida, United States

***Correspondence:**
Lidija Begović
lbegovic@biologija.unios.hr

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

Received: *29 October 2021*
Accepted: *30 May 2022*
Published: *06 July 2022*

Citation:
Galić V, Mlinarić S, Marelja M, Zdunić Z, Brkić A, Mazur M, Begović L and Šimić D (2022) Contrasting Water Withholding Responses of Young Maize Plants Reveal Link Between Lipid Peroxidation and Osmotic Regulation Corroborated by Genetic Analysis. Front. Plant Sci. 13:804630. doi: 10.3389/fpls.2022.804630

INTRODUCTION

One of the key targets of maize (*Zea mays* L.) breeding is tolerance to abiotic stress conditions. Phenotyping for stress responses represents the key for success in breeding while the underlying trait physiology mostly remains unclear (Masuka et al., 2012). Plants adapt to sub-optimal conditions by morpho-physiological adjustments, with vast number of mechanisms on different organizational levels at their disposal (Pareek et al., 2010). However, genotypic variability for these adjustments exists and some genotypes are expected to cope with abiotic stress conditions better than others (Slafer and Araus, 2007; Tardieu, 2012). Water deficit represents one of the main abiotic stresses in the field conditions in rain-fed areas and can affect the plant

growth and development at any time from emergence to yield formation causing an outburst of physiological responses (Wang et al., 2019). For example, water deficit affecting the plant during reproductive stages can cause the formation of smaller number of kernels or kernel abortion. In grain filling, it results in smaller grains and premature senescence, whereas the effects of drought in early plant development received relatively less attention despite the fact that water deficit at this stage can cause within-field variability in plant size, deteriorate the stands, and make the crop more susceptible to diseases (Farooq et al., 2012; Aslam et al., 2015). Climate change causes a significant alteration of spatiotemporal patterns of drought occurrence (Stagge et al., 2017; Grillakis, 2019), with more recent droughts lasting longer and occurring at less predictable times. Water deficit results in a number of adverse effects such as reduction of plant turgidity, reactive oxygen species (ROS) build-up, decrease in photosynthetic efficiency, and ultimately, the plant death. However, the plant mechanisms to cope with these adverse effects include the osmotic adjustments by synthetizing the osmotically active compounds and increase in enzymatic activity to detoxify the effects of ROS (Anjum et al., 2017).

During the oxidative stress periods, plant cell suffers damage at many levels, i.e., outer cell layers, cytoplasmic components, nucleus, and nucleic contents, inevitably leading to selective cell death. ROS exert highly oxidative cell surroundings, interacting with lipids (primarily polyunsaturated fats, PUFAs), proteins, and nucleic acids often resulting in limitations to biological yield (Czarnocka and Karpiński, 2018). One family of products of this highly detrimental interaction is the products of lipid biomolecule peroxidation called malondialdehydes (MDAs). MDAs are highly reactive in oxidative surroundings expected in oxidative stress due to their α-β-unsaturated carbonyl group and are thus known as reactive carbonyl species (RCS). The well-known instability and reactivity of these species makes it unfeasible to measure them directly, so the products of their secondary activity, reacting with thiobarbituric acid (i.e., *thiobarbituric acid reactive substances*, TBARS), are measured instead. RCS are in parallel produced both enzymatically (lipoxygenase activity) and non-enzymatically (ROS-mediated), and both processes also occur in healthy organisms (Farmer and Mueller, 2013). However, temporary increase in MDA levels in stress conditions represents acclimation process that activates regulatory gene networks involved in plant defense and development such as dehydration/heat shock-related genes and genes involved in antioxidant machinery (Morales and Munné-Bosch, 2019). Moreover, MDA can cause the transcriptional reprogram of a cell, activating transcription of abiotic-stress related genes, making them effective signal molecules (Weber et al., 2004). Another versatile plant signal molecule is hydrogen peroxide (H_2O_2). H_2O_2 is a ROS byproduct of metabolism, mostly built-up during stress-induced respiratory burst of plant plasma-membrane NADPH oxidases by superoxide-dismutase from more toxic oxygen species, or from the process of β-oxidation of lipid molecules in membrane bound microbodies peroxisomes (Corpas et al., 2020). In other cell compartments, lower doses of H_2O_2 show a limited cellular toxicity, as it is easily accumulated in plant cells by downregulation of its peroxisome-localized degradation enzymes such as catalase, ascorbate peroxidase, glutathione peroxidase, and so on. Which makes for a robust signaling molecule (Hossain et al., 2015). Generally, the H_2O_2 degradation process in maize is carried out by two classes of peroxidases; the ones utilizing its substrate in lignification and organogenesis, such as guaiacol peroxidases, and the others scavenging the peroxide molecules utilizing pyridine nucleotides, GSH, cytochrome *c* and ascorbate as electron donors (Prasad et al., 1995). The former group is involved in the young plant development, whereas both groups are involved in stress responses (Gechev et al., 2006) and signaling (Kidwai et al., 2020). One of the main tasks of H_2O_2 in stress-signaling appears to be the regulation of osmolyte synthesis, specifically proline, through transcriptional upregulation of proline-biosynthesis genes (Yang et al., 2009), and downregulation of its degradation pathways. Since the downregulation of peroxisomal H_2O_2-scavenging enzymes appears to be the main source of signaling H_2O_2 in cells (Su et al., 2019), and the peroxisomes also serve as the alternative cell energy supply by lipid catabolism, it is possible that by alterations of the peroxisomal regulation, some other signaling cascades become dominant or more pronounced.

Maize breeding relies on several germplasm resources (Lee and Tracy, 2009), key of which is the elite germplasm available after the expiration of plant variety protection (PVP) certificate, the so-called ex-PVP germplasm (Mikel and Dudley, 2006). In modern maize breeding, with maize holding the majority of world seed market (FAO/IHS Markit Agribusiness Consulting, 2019), the ex-PVP germplasm still prevails the new inbred registrations (Mikel, 2011; White et al., 2020). This extremely valuable germplasm resource consists of thousands of genotypic accessions (Romay et al., 2013) with traceable pedigrees and available comprehensive genotypic and phenotypic data (Canaran et al., 2008). Many studies were conducted based on this resource; however, studies combining the physiological assessment with dense genotypic data in elite germplasm are scarce. Moreover, the studies reporting results of association analysis for traits assessing oxidative status, lipid peroxidation, and proline accumulation in maize are scarce. Breeding for tolerance to early osmotic stress might be a meaningful endeavor due to the adverse effects of stress at this stage on stands and crop health status (Farooq et al., 2012; Aslam et al., 2015), corroborated by high variability of available germplasm resources in terms of ancestry (Lee and Tracy, 2009), genotype (White et al., 2020), and consequently the phenotype (Galić et al., 2020). It was hypothesized that integration of data assessing plant osmotic status (CWC, proline, PROT), plant oxidative status (TBARS, H_2O_2), and dense genotypic data might enhance detection of new important loci or pathways for adaptation to osmotic stress, as well as to highlight the gene ontology (GO) enrichment, thus facilitating the discovery of new regulatory networks involved in plant response to water withholding. An enrichment analysis of GO is an efficient methodology for the assessment of functions linked to large gene lists increasing the likelihood of interpretation of the detected biological processes and regulatory networks (Tian et al., 2017). In the advent of high-throughput molecular techniques identifying more and more genes and generating the big data, combining GO with

association mapping helps to bridge the gap in translation of genomes to phenomes (Pan et al., 2019).

Based on that, the aims of this study were to investigate the contrasts in responses of panel of diverse, elite ex-PVP and public maize inbred lines in proline accumulation lipid peroxidation and oxidative status to water withholding in early stages of growth and combine these results with dense genotypic data to increase precision and true detection rate in association analysis. We further aimed to establish the relationships between the assessed phenotypic traits and use the results of genetic analysis to determine the enriched biological processes involved in maize osmotic regulation. Our study represents a novel effort in maize to corroborate the findings in stress biochemistry with enrichment analysis of underlying genetic associations in a panel of elite inbred lines with global significance.

MATERIALS AND METHODS

Plant Material and Experimental Design

The experimental design was previously described in Galić et al. (2020). Briefly, seeds of maize inbred lines were freely collected according to the US Plant Variety Protection Act upon the expiration of their respective certificates from US National Plant Germplasm System (NPGS) according to their Distribution Policy (https://npgsweb.ars-grin.gov/gringlobal/distribution, accessed May 5, 2022). Seeds were transferred with their enclosed passport documents and USDA-APHIS Plant Export and Phytosanitary Certificates in fall of 2017. The maize inbred lines used to carry out this research were not subjected to any form of modification. Seeds were planted in field in growing season 2018 and selfed to obtain seeds in sufficient quantity for experiments. Selfing was successful for 109 inbreds. Experiments were conducted with awareness of the requirements the IUCN Policy Statement on Research Involving Species at Risk of Extinction and the Convention on the Trade in Endangered Species of Wild Fauna and Flora with intention to comply with all relevant institutional, national, and international guidelines and legislation. Experiments were set in controlled conditions (25°C, RH = 50%, 16-h/8-h day/night, 200 μmol/m^2/s) in trays (510 mm × 350 mm × 200 mm) divided to 15 rows with 7 boxes (50 mm × 35 mm). Each tray was filled with 8.67 kg (20 L) of air-dry soil [pH ($CaCl_2$) = 5.7, N ($NH_4^+ + NO_3^-$) = 70 mg/L, P (P_2O_5) = 50 mg/L, K (K_2O) = 90 mg/L, EC = 40 mS/m] and the planting was performed to 2 cm depth. The experiment was set with single water withholding treatment (WW) and control (CO). A total of three biological replicates (trays) of each genotype were planted, with one tray containing a single biological replicate of 15 genotypes (rows), with single row (7 planting boxes) each to enable shuffling. Watering regime for WW was optimized in preliminary trials to obtain the 50% reduction in fresh weight per plant in WW treatment compared to CO. Plants in CO were watered with spray bottle in planting and every 2 days thereafter with 8 ml of tap water per plant. Amount of water added to CO was determined by weighting the soil after 3-day drainage following complete saturation of soil with water (field water capacity). Plants in WW were watered in planting with full dose (8 ml) and two times thereafter (last watering on 4th day) with half a dose of water added per plant in C (4 ml). After that, water was withheld up to the 14th day since planting (10 days of water withholding) when the aboveground parts of three equally developed plants per genotype in each replicate were harvested. A total of eight trays were left without plants in the same conditions for soil weighting. Trays were treated in the same way as the experimental ones, with four trays representing Co and the other four representing WW. Trays were weighted every day. The 1 g samples were taken for analysis of current water content (CWC), whereas the rest was frozen on liquid nitrogen and left in −80°C freezer for further analyses. Tubes with 1 g samples were dried for 24 h in digital laboratory oven on 80°C. Current water content (CWC) was calculated:

$$CWC = \frac{Fresh\ weight - Dry\ weight}{Fresh\ weight} * 100 \quad (1)$$

Biochemical Analyses

All biochemical analyses were carried out in three biological replicates, each further measured in three technical (lab) replicates. Proline content (proline) was determined according to Carillo and Gibon (2011). About 20 mg of fresh seedling tissue was extracted in 400 μl (ethanol: water, 40:60 v/v) overnight at 4°C. For measurements, 50 μl of extract was used. Measurements were taken on microplate reader (Tecan, Spark) at 520 nm. Proline content was calculated from the standard curve using proline as standard and expressed as nmol/mg of fresh weight (FW).

Analysis of TBARS (thiobarbituric acid reactive substances) was performed according to the method described by Jambunathan (2010). After tissue homogenization in liquid nitrogen, about 0.2 g of plant tissue was extracted by the addition of 1 ml of 0.1% trichloroacetic acid (TCA). The samples were centrifuged for 5 min at 6,000 g at 4°C. After centrifugation, 0.5 ml of the supernatant was separated into a screw cap tube and 1 ml of TBA in TCA (0.5% thiobarbituric acid solution in 20% trichloroacetic acid solution) was added. The blank contained 1.5 mL of TBA in TCA. The reaction mixture was vortexed and incubated in a water bath for 30 min at 95°C followed by centrifugation for 15 min at 18,000 rpm at 4°C. The absorbance was measured at 532 and 600 nm. Obtained results were expressed as nmol/g of fresh weight (FW).

Concentration of H_2O_2 was determined by the method according to Mukherjee and Choudhuri (1983). After tissue homogenization in liquid nitrogen, 0.1 g of powder was extracted with 1 ml of cold acetone. The reaction mixture was vortexed and centrifuged for 5 min at 6,000 g and 4°C. The supernatant was separated, and 400 μl of titanium sulfate solution and 500 μl of concentrated ammonium hydroxide (NH_4OH) were added. The reaction mixture was centrifuged at 15,000 rpm for 10 min at 4°C. The supernatant was decanted and the resulting precipitate was dissolved by the addition of 1 ml of 2M H_2SO_4 solution. The absorbance was measured at 415 nm, and the H_2O_2 concentration was expressed as nmol/g of sample fresh weight (FW).

The method described by Siegel and Galston (1967) for determination of guaiacol peroxidase (GPOD) was adapted for

microplate reader. Briefly, approximately 0.2 g of previously powdered seedling tissue was extracted with 1 ml of 0.1 M phosphate buffer (pH 7.0). Samples were centrifuged at 18,000 rpm at 4°C. After centrifugation, samples were re-extracted and supernatants were pooled. Reaction mixture was prepared using 8 mM H_2O_2 and 90 mM guaiacol in 1:1 ratio (v/v). All measurements were taken in triplicates in 96-well plates using microplate reader (Tecan, Spark) by adding 150 μl of phosphate buffer, 40 μl guaiacol/H_2O_2 mix, and 10 μl of extract. Absorbance was read at 436 nm using extinction coefficient of 25.5 $mM^{-1}cm^{-1}$, and results were expressed as units per gram of sample fresh weight (U/FW).

Total protein content (PROT) in the samples was determined by the Bradford method (Bradford, 1976) adapted for microplate reader. Proteins were measured in same extract used for peroxidase activity assay. Briefly, 5 μl of sample and 250 μl of Bradford reagent were mixed and absorbance was read at 595 nm. The preparation of standard curve dilutions of bovine serum albumin (BSA) was used (0.125–1.4 mg/ml). The protein concentration was expressed as mg/g fresh weight.

Analyses of Phenotypic Traits

Pearson's product-moment correlations were calculated between genotypic means of all traits to establish trait connections within as well as between control and WW treatment. Genotypic means represented a mean value of three technical replicates over three biological replicates, totally 9 data points per genotype in each treatment. For each trait, fold-change values were calculated to examine specific patterns of reactions as $Tr_{rel} = \frac{Tr_{ww} - Tr_{co}}{Tr_{co}}$, where Tr_{rel} represents trait fold-change in water withholding (Tr_{ww}) relative to trait value in control conditions (Tr_{co}). Correlations were also calculated between trait fold-changes to establish connections between different physiological indicators. Correlation coefficients were displayed using the *R/gplots* package (Warnes et al., 2013) function *heatmap.2*. Relationships between traits were imposed using a correlation distance-based clustering.

To inspect the distinct patterns of reactions between groups of genotypes, unsupervised K-means clustering analysis was carried out in R/*factoextra* library (Kassambara and Mundt, 2020) with Tr_{rel} values as input. To determine the optimal number of clusters, a Silhouette statistic was computed (Charrad et al., 2014), and the optimal number of clusters was 2 (**Supplementary Figure S1**).

Variance components of all traits were calculated in R/*sommer* library (Covarrubias-Pazaran, 2016). Models were specified with unstructured error variances as follows:

$$y_{ijk} = G_i + (T_j) + GT_{ij} + \varepsilon_{ijk} \quad (2)$$

where y_{ijk} was value of *i*-th genotype in *j*-th treatment in *k*-th replicate, G_i represents random effects of genotype, (T_j) represents fixed treatment effects, GT_{ij} is the random genotype-treatment interaction term, and ε_{ijk} is the overall model error term. Estimates of errors of variance components and the trait repeatabilities were calculated in package's *pin* calculator. Trait repeatabilities were calculated as: $H^2 = \frac{\sigma_G^2}{\sigma_G^2 + \frac{\sigma_{GxT}^2}{n_t} + \frac{\sigma_\varepsilon^2}{n_t * n_r}}$, where σ_G^2, σ_{GxT}^2, and σ_ε^2 represent variance components of genotype, genotype by treatment interaction and error, respectively, whereas n_t is the number of treatments and n_r is the number of replicates. Satterthwaite's estimates of *p*-values for differences between treatment effects were calculated in R/*lmerTest* library (Kuznetsova et al., 2017).

Genotypic Data Manipulation and Analysis

Genotyping was performed at Cornell University with genotyping by sequencing approach with protocol from Elshire et al. (2011). Genotypic data were the part of the genotyping efforts of the US National Plant Gemplasm System Gene Bank consisting of tens of thousands of maize accessions (Romay et al., 2013). The genotypic data were retrieved from Panzea (panzea.org) repository (Canaran et al., 2008) as partially imputed calls with ∼955,000 SNP positions on AGPv4 B73 reference alignment. Dataset 1 was constructed by filtering the original partially imputed GBS data to maximum 10% missing data, minor allele frequency of 0.02, and no heterozygotes. Filtering resulted in 107,527 positions. Positions were imputed by using LinkImpute methodology (Money et al., 2015) with 30 sites in high LD and 10 nearest neighbors. For Dataset 2, LD pruning of positions was performed in Plink 1.9 (Purcell et al., 2007) with –indep 50 5 0.95 flag to prevent false positives in association analysis. Pruning resulted in 70,130 variants. Dataset 2 was used for PCoA, kinship, and association analysis. Association analysis was performed in mixed-effects linear model framework (MLM+Q+K) with Q matrix calculated in principal coordinate analysis with 7 assumed axes and identity-by-state kinship matrix (K) as covariate in Tassel software (Bradbury et al., 2007) version 5.2.67. To further control false detection rate (FDR), cluster affiliations of different genotypes from K-means analysis were used as covariate. The Tr_{rel} values were used as phenotypes in association analysis. Arbitrary –log(p) threshold of 4 was used to declare significant associations according to the results of Bian and Holland (Bian and Holland, 2017) that showed the stable predictive abilities of the loci detected in the range of –log(P) thresholds from value of 4 to Bonferroni-corrected value in oligogenic and polygenic traits. Bonferroni threshold was also determined following the simpleM procedure described by Gao et al. (2008, 2010). Briefly, of the 70,130 filtered and imputed markers, the effective number of markers (M_{eff}) was determined to be 13,966 and the significance threshold of $\alpha = 0.05$ was divided with the M_{eff} which resulted in –log(p) value of 5.446.

The scan for genes associated with detected positions was carried out using a Ensembl Plants service (Howe et al., 2020), BioMart tool (Kinsella et al., 2011). The scan was limited to protein-coding genes 120 kbp from the detected associations in both directions, according to the results of linkage disequilibrium in this association panel (Galić et al., 2020). Only the protein-coding genes were analyzed. All 182 detected genes were subjected to AgriGo version 2.0 analysis for enrichment of biological processes, pathways, and cellular components (Tian et al., 2017), but only 162 genes with known pathways were

TABLE 1 | Mean values ± standard deviations and *p*-values of differences in means in control (CO) and 10-day water withholding treatment (WW) for GPOD (AU/g FW) PROT (mg/g FW), TBARS (ng/g FW), proline (nmol/mg FW), H_2O_2 (nmol/g FW), and CWC (%) followed by variance components and repeatabilities ± standard errors of repeatabilities and the results of K-means clustering analysis with cluster means of trait fold-change and the accompanying *p*-values of pairwise *t*-tests between cluster means.

	GPOD	PROT	TBARS	Proline	H_2O_2	CWC
Treatment effects						
CO	0.94 ± 0.23	0.72 ± 0.21	4.14 ± 1.18	1.85 ± 0.66	1.66 ± 0.55	93.52 ± 0.85
WW	1.00 ± 0.29	0.67 ± 0.25	6.15 ± 2.00	5.33 ± 3.55	2.05 ± 0.71	90.92 ± 1.67
p-value	0.00808	0.362	<0.001	<0.001	<0.001	<0.001
Variance components and repeatabilities						
σ^2_G	0.036	1.404	1.181	1.068	0.269	0.862
σ^2_{GxT}	0.028	0.998	1.461	5.360	0.117	0.627
σ^2_e	0.014	0.467	0.193	0.263	0.049	0.781
R^2	0.68 ± 0.06	0.71 ± 0.06	0.61 ± 0.08	0.28 ± 0.14	0.80 ± 0.04	0.66 ± 0.07
K-means cluster analysis						
Cluster 1 ($n = 65$)	0.010 ± 0.22	−0.002 ± 0.421	0.321 ± 0.265	0.982 ± 0.691	0.188 ± 0.299	−0.312 ± 0.152
Cluster 2 ($n = 44$)	0.195 ± 0.37	−0.107 ± 0.324	0.863 ± 0.517	3.400 ± 2.475	0.407 ± 0.353	−0.545 ± 0.243
p	0.003964	0.1453	<0.001	<0.001	0.001127	<0.001

analyzed. AgriGo uses the Fisher's test and the *z-scores* to retrieve the enriched terms, taking into the account total number of one organism's genes annotated with GO or of the user-provided background, the number of genes mapped to the background in the query list, the total number of genes in one GO term and the counts of overlapped genes as well as the means, and standard deviations of sample counts. The software's acyclic drawer is based on semantic similarity measurement (SSM) described in Wang et al. (2007).

RESULTS

Treatment Effects and Variance Components

Water withholding treatment induced visible drying of soil in terms of field water capacity (FWC). On the day 0 (last watering in WW), the mean soil water content in WW was 75.4% compared to 90.1% in CO. The difference was induced by adding only a half dose of water to the soil in WW compared to CO at day 0. On the day of sampling (day 10), the mean soil water content dropped to 22.8% FWC (**Supplementary Figure S2**). In CO, the soil water content was maintained between 72.5 and 93.2% FWC. Water withholding treatment induced significant changes in all measured biochemical traits (**Table 1**) except PROT ($p = 0.362$). TBARS was in average significantly increased in WW treatment compared to control (48.5% increase), as well as proline (288.1%), H_2O_2 (23.5%), and GPOD (6.38%). Significant decrease was observed only in CWC (2.61%). Non-zero estimates of variance components were observed for GPOD, TBARS, proline, H_2O_2, and CWC (**Table 1**). Genotype by treatment interaction was larger than genotypic variance only in proline and TBARS. All repeatability estimates were larger than zero, spanning from 0.28 (proline) to 0.80 (H_2O_2).

Correlation Analysis

Correlation analysis of trait values between CO and WW treatment showed moderate to strong, significant correlations between all traits (**Figure 1A**). Lowest correlation between CO and WW was observed for proline (0.455, not shown), whereas the highest was observed for H_2O_2 (0.690). Very weak to weak positive and negative correlations were detected between all traits in CO, whereas in WW, most of the correlation strengths increased. Moderate positive correlation was observed in WW between GPOD and TBARS (0.555) and between TBARS and proline (0.595). Contrarily, the strong negative correlation was observed in WW between CWC and proline (−0.637), whereas the correlation between TBARS and CWC was moderate negative (−0.404).

Trait fold-change was calculated and used for further analyses because the change in the assessed parameters is of true interest for understanding plant adaptation to osmotic stress conditions and differentiate sensitive genotypes from tolerant ones. When the correlations were calculated between trait fold-changes, clustering of traits by correlation patterns was observed (**Figure 1B**). A number of two separate clusters were formed: one with TBARS, proline, H_2O_2, and GPOD and another with PROT and CWC. Strongest positive correlation was observed between fold-change values of TBARS and proline (0.583). Significant weak positive correlations were also detected between H_2O_2 and TBARS (0.368) and between H_2O_2 and proline (0.220). Strongest negative correlation was observed between fold-change values of proline and CWC (−0.578). Other two significant negative correlations were detected between fold-change values of CWC and TBARS (−0.477) and between CWC and H_2O_2 (0.300).

Crossover Reactions Indicate Involvement of TBARS in Osmotic Stress Signaling

To further analyze the relationship patterns of reactions in lipid peroxidation, hydrogen peroxide build-up, and proline

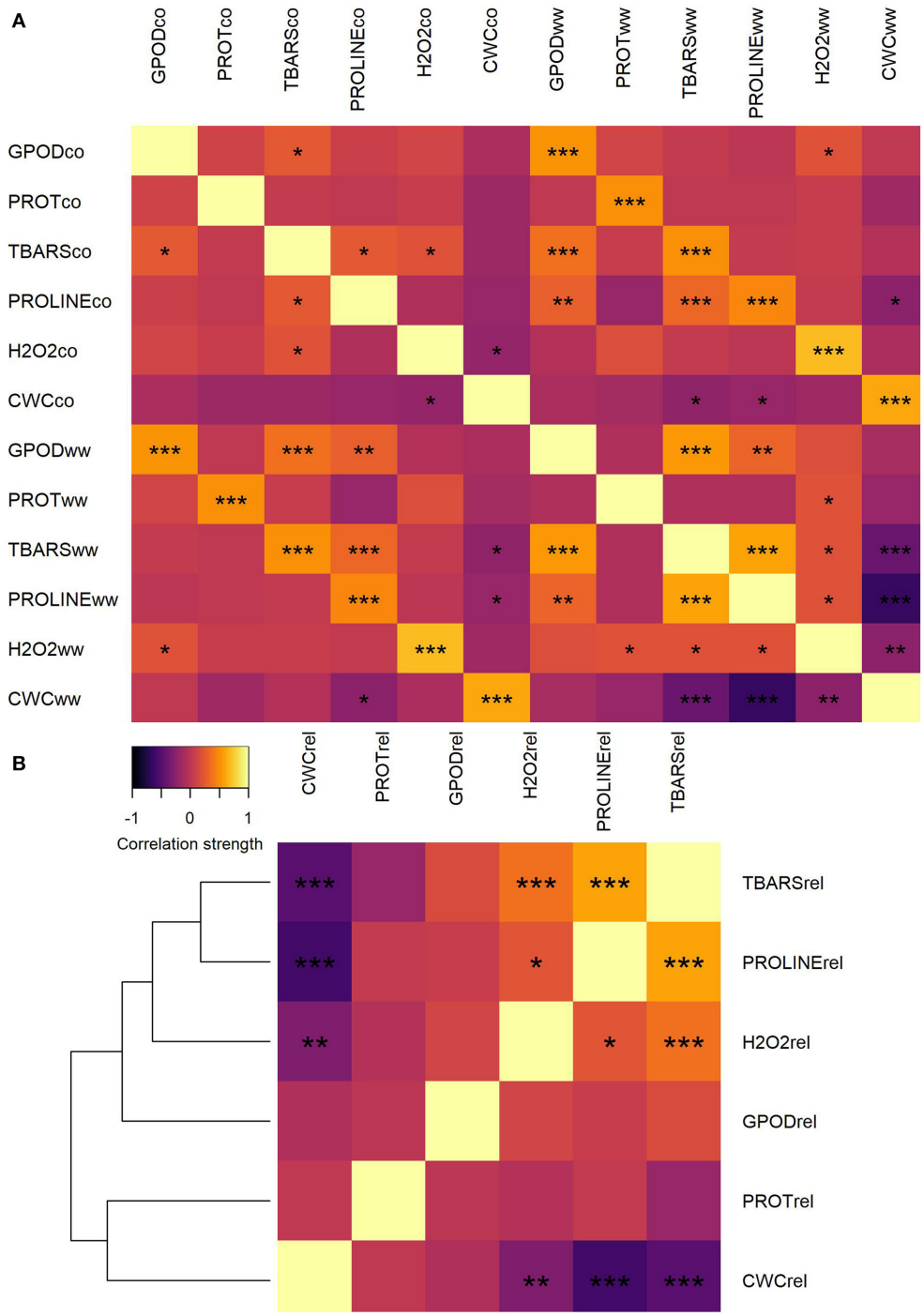


FIGURE 1 | Heatmap of correlations between trait mean values in CO and WW **(A)** and heatmap and correlation-distance cladogram of correlations between trait fold-change values in WW compared to CO **(B)**. *, **, and *** represent significance of correlations at $\alpha = 0.05$, $\alpha = 0.01$, and $\alpha = 0.001$ levels, respectively.

accumulation, trait fold-changes were plotted on a 2-y-axis plot (**Figure 2A**). Highest values of proline accumulation were observed in accessions where TBARS fold-change crossed the value of fold-change in H_2O_2. Contrarily, in genotypes in which fold-change in H_2O_2 was larger than the fold-change in TBARS, proline accumulation appeared to be lower. Genotypes were divided in two groups following this pattern, the one in which fold-change of TBARS was larger than fold-change in H_2O_2 (TBARS > H_2O_2), and another in which fold-change in H_2O_2 was larger compared to fold-change in TBARS (H_2O_2 >

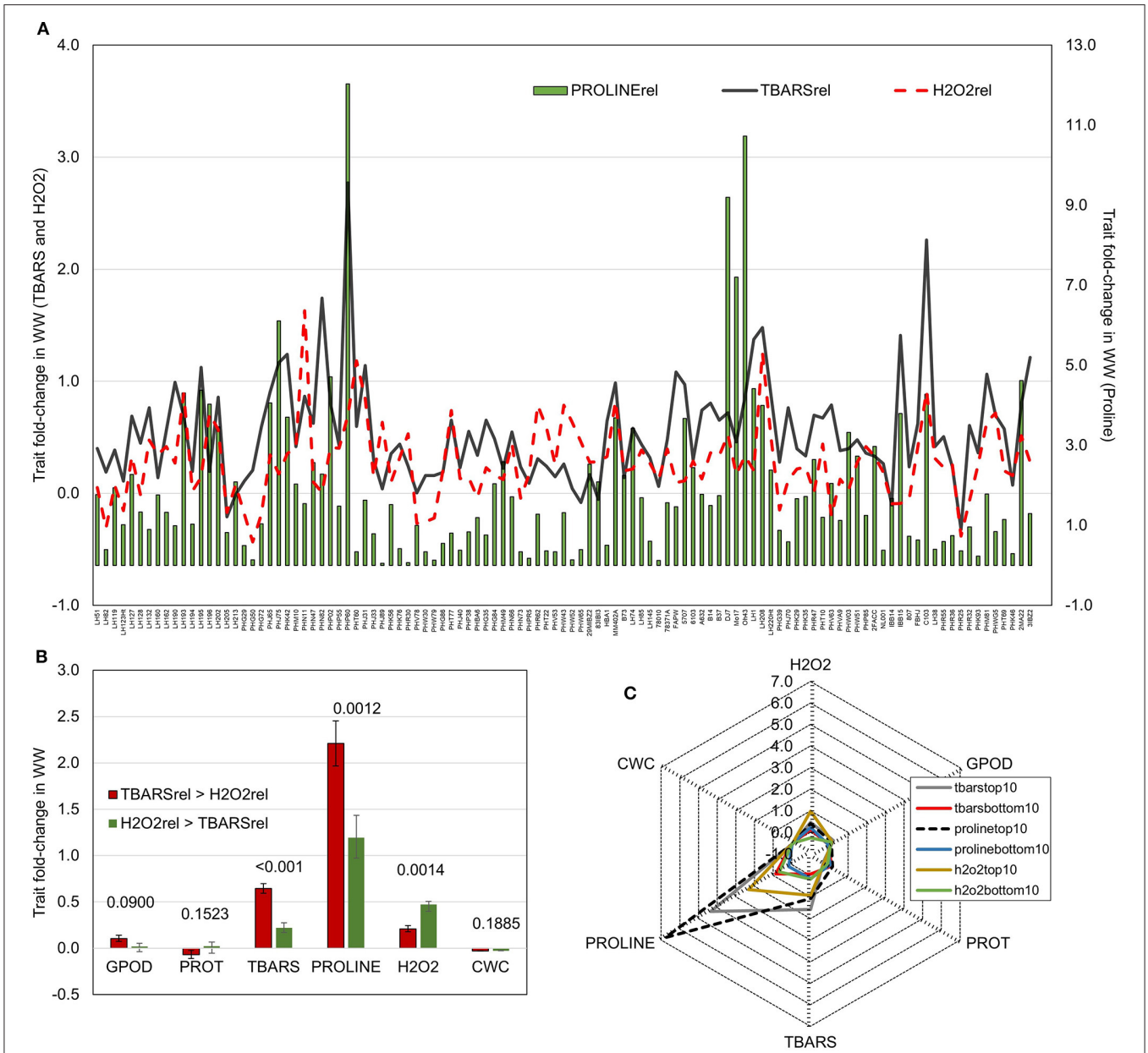


FIGURE 2 | (A) Fold-change in WW compared to CO of TBARS and H_2O_2 accumulation (left y-axis) and proline (right y-axis); **(B)** Fold changes in WW compared to CO of GPOD, PROT, TBARS, proline, H_2O_2, and CWC in groups with larger fold change of TBARS than H_2O_2 (TBARS > H_2O_2) and with larger fold change of H_2O_2 than TBARS (H_2O_2 > TBARS). *p*-values of pairwise *t*-tests between groups are given above bars; **(C)** radar-plots of fold changes WW compared to CO of GPOD, PROT, TBARS, proline, H_2O_2, and CWC in top and bottom 10 performers in fold change of TBARS, proline, and H_2O_2.

TBARS). The analysis showed significantly larger fold-change in TBARS (0.645 vs. 0.221) and proline (2.21 vs. 1.19) in TBARS > H_2O_2 group, and significantly lower fold-change in H_2O_2, whereas other analyzed traits were not significantly affected (**Figure 2B**). Moreover, analysis of the top 10 scorers for proline, TBARS, and H_2O_2 fold-change revealed that highest mean fold-change in TBARS was accompanied by second-highest fold-change in proline accumulation of 4.35 (after top 10 scorers for proline with mean of 6.77), followed by top 10 scorers in H_2O_2 with mean fold-change in proline of 2.31 (**Figure 2C**).

Most interestingly, correlation analysis of trait fold-changes between two groups (TBARS > H_2O_2 and H_2O_2 > TBARS) showed strong positive significant correlation between H_2O_2 and TBARS in H_2O_2 > TBARS group (**Table 2**), compared to moderate significant correlation between these traits in TBARS > H_2O_2 group, indicating possible activation of the alternative pathway of MDA

TABLE 2 | TBARS (ng/g FW), proline (nmol/mg FW) and H_2O_2 (nmol/g FW) means ± standard errors of mean in WW compared to CO, and the correlations (bold) between the traits in two contrasting groups (**Figure 5A**), TBARS > H_2O_2, and H_2O_2 > TBARS.

	Group	TBARS	PROLINE	H_2O_2
CO	TBARS > H2O2	3.958 ± 0.125	1.828 ± 0.074	1.7 ± 0.064
	H2O2 > TBARS	4.698 ± 0.23	1.917 ± 0.124	1.535 ± 0.084
WW	TBARS > H2O2	6.324 ± 0.23	5.684 ± 0.403	1.994 ± 0.074
	H2O2 > TBARS	5.622 ± 0.317	4.244 ± 0.583	2.216 ± 0.153
PROLINE	**TBARS > H2O2**	***0.590***	**–**	**–**
	H2O2 > TBARS	**0.181**	**–**	**–**
H_2O_2	**TBARS > H2O2**	***0.564***	***0.391***	**–**
	H2O2 > TBARS	***0.812***	**0.118**	**–**

Correlations significant at α = 0.05 level are shown in italic.

build-up. The increase in correlation strength from non-significant weak in H_2O_2 > TBARS to moderate to strong significant positive in TBARS > H_2O_2 group indicated possible involvement of this increase in lipid peroxidation in proline accumulation signaling.

K-means clustering with trait fold-change showed two distinct clusters of reactions (**Figure 3**). Clustering explained 53.3% of total variation in first two dimensions. In first cluster, 65 genotypes with moderate reactions to WW treatment were grouped, whereas in the other, the reactions of 44 genotypes to WW were more pronounced. Cluster designations of inbreds are available in **Supplementary Table S1**, available online. Differences in changes in reactions between clusters were significant for all traits except PROT (**Table 1**, p = 0.1453). The largest, more than three-fold difference between clusters was observed for proline (346%). Following the groups identified in **Figure 2**, groups H_2O_2 > TBARS and TBARS > H_2O_2 were given different symbols to further analyze their arrangement within clusters showing different responses to WW. Interestingly, only seven genotypes from H_2O_2 > TBARS were located in cluster 2, harboring genotypes with significantly higher proline accumulation (**Table 1**), whereas the remaining 20 genotypes were located in cluster 1, including the genotype with most extreme total phenotypic response in the two analyzed dimensions (PHW65).

The selection of each cluster representatives for further analysis was carried with following heuristics. First, the three nearest-to-centroid points were selected, with centroid representing an imaginary center of cluster (average response in reduced 2d hyperplane), as cluster representatives. Second, the farthest genotype in both dimensions (PHW65 in cluster 1 and PHP60 in cluster 2) was selected. The analysis of candidate's responses (**Figure 4**) showed apparent differences in proline accumulation, accompanied by subtler changes in other biochemical parameters. To accompany the analysis of relationship between accumulation of H_2O_2 and TBARS in context of proline accumulation (**Figure 2**), the relative response of H_2O_2 was subtracted from relative response in TBARS and the simple linear regression showed highly significant relationship (R^2= 0.876) between this difference and the proline accumulation (**Figure 4**).

Association Analysis and Candidate Genes

Allelic effects in association analysis in all analyzed traits followed normal distribution and no considerable deviations of effects were detected up to the value of –log(p) of 4 (**Figure 5**). Inflations of effects on the right tails of distributions for proline and TBARS (**Figures 5B,F**) indicated the presence of loci crossing the calculated Bonferroni threshold. Totally, 34 associations were declared significant (**Figure 5**; **Table 3**), most of which were detected for proline (29.4%) and TBARS (44.1%). A total of three of the 34 associations crossed the calculated Bonferroni threshold of 5.446 (**Table 3**), two of which were detected for proline (**Figure 5B**), one on chromosome three (PROLINE2@3), and another on chromosome eight (PROLINE6@8), along with a single association for TBARS, TBARS14@9 (**Figure 5F**). PROLINE2@3 was located on physical position 189,739,999 bp with peak –log(p) value of 5.970 and four genotypes carrying the minor variant. PROLINE6@8 was located on physical position 21,838,456 bp with peak –log(p) 5.897 and four genotypes carrying the minor allele. Single association that crossed Bonferroni threshold detected for TBARS (TBARS14@9) was located on chromosome 9, physical position 139,353,870 bp with –log(p) value of 5.616 and 12 genotypes with minor allele. Many of the variants detected as different associated loci represent same associations, however, with different numbers of genotypes carrying the minor allele. For example, associations PROT1@2 and PROT2@2 are 18.2 kbp apart with 13 and 11 minor-allele carriers. GPOD2@8 and GPOD3@8 probably also represent the same association as the distance between these loci is only 76 bp. TBARS6@6 and TBARS7@6 are 5.3 kbp apart, whereas TBARS13@9, TBARS14@9, and TBARS15@9 are 49.5 kbp and 7 bp apart, respectively. However, GPOD3@8 and GPOD4@8 might not represent the same association, as the 143.5 kbp distance exceeds the 120 kbp linkage disequilibrium block size. Within regions carrying the significant associations, 120 kbp in both directions from the peak physical locations 182 candidate genes from various metabolic pathways were found in the BioMart analysis (**Supplementary Table S3**) with 869 putative transcripts.

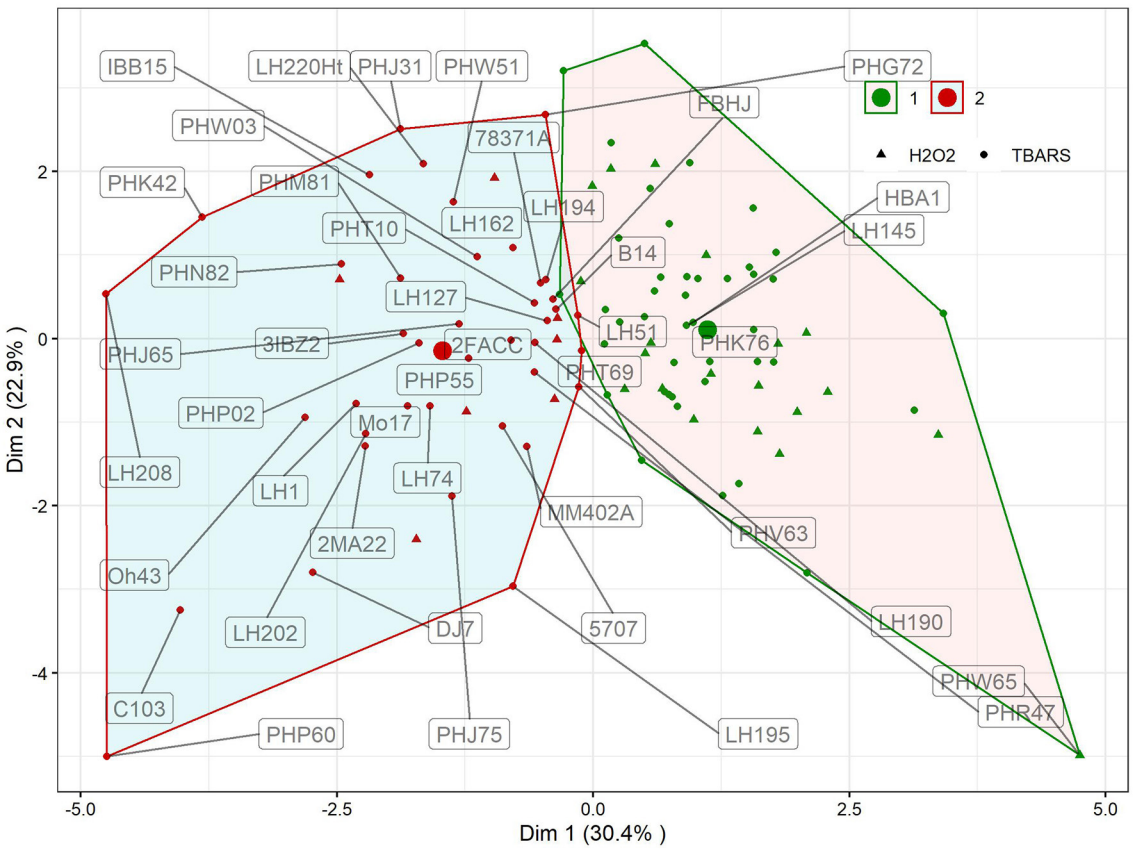


FIGURE 3 | Results of K-means clustering of genotypic means using trait fold-change values of GPOD, PROT, TBARS, proline, H_2O_2, and CWC in WW compared to CO. Inbred designations are given for all inbreds belonging to the cluster 2 and TBARS > H_2O_2 group, along with inbreds surrounding the centroid in cluster 1 and the most extrema phenotype in cluster 1 (PHW65). Large dots show centroid of each cluster.

Gene Ontology Enrichment Analysis

The GO analysis with 162 candidate genes from BioMart analysis passing the quality check with AgriGO 2.0 online mining tool showed significant enrichment of cellular components (**Figure 6**) and biological processes (**Supplementary Figure S3**). Highly significant negative regulation of the integral membrane parts was detected in cellular component analysis (**Figure 6**, $p < 10^{-9}$), along with significant ($p < 0.05$) negative regulation of microbodies peroxisomes. Biological processes analysis (**Supplementary Figure S3**) showed large number of positively regulated processes ($p < 0.05$), linked to regulation of DNA-dependent transcription and polysaccharide catabolic processes.

DISCUSSION

Variation of Stress-Related Traits in Water Withholding Treatment

The studies reporting quantitative genetic analysis of biochemical parameters involved in stress response are scarce, although these parameters harbor information about the well-known biological processes, such as detoxification of ROS, lipid peroxidation, or proline accumulation and might thus represent well-worth traits in breeding for osmotic stress tolerance and consequently higher yields (Tardieu, 2012). For example, in the study on sunflower hybrids by Khalil et al. (2016), higher broad-sense heritabilities were reported compared to repeatabilities detected in our study, which could be explained by the different crops, different stages and using hybrids compared to inbred lines. Several studies reported loci associated with proline accumulation in barley (Fan et al., 2015; Jang et al., 2020) and rice (Sayed et al., 2012), hydrogen peroxide build-up (Gill et al., 2019; Kumar and Nadarajah, 2020), and TBARS in rice (Jiang et al., 2009), wheat (Ma et al., 2015), and cotton (Yasir et al., 2019); however, there are no available results for these important physiological processes in maize up to this date. Non-zero genetic variances in variance component analysis (**Table 1**) imply feasibility of breeding directly for these traits, although the fold change compared to control might be more useful in screening of maize accessions due to the functional diversity of analyzed traits even in non-stressful conditions. Designed experiments conducted in controlled conditions lack in diverse conditions and stressors that plant must cope with in field (Farooq et al., 2012). However, any drought-related trait can confer drought tolerance if addressed to a proper climatological scenario (Tardieu, 2012), and analysis

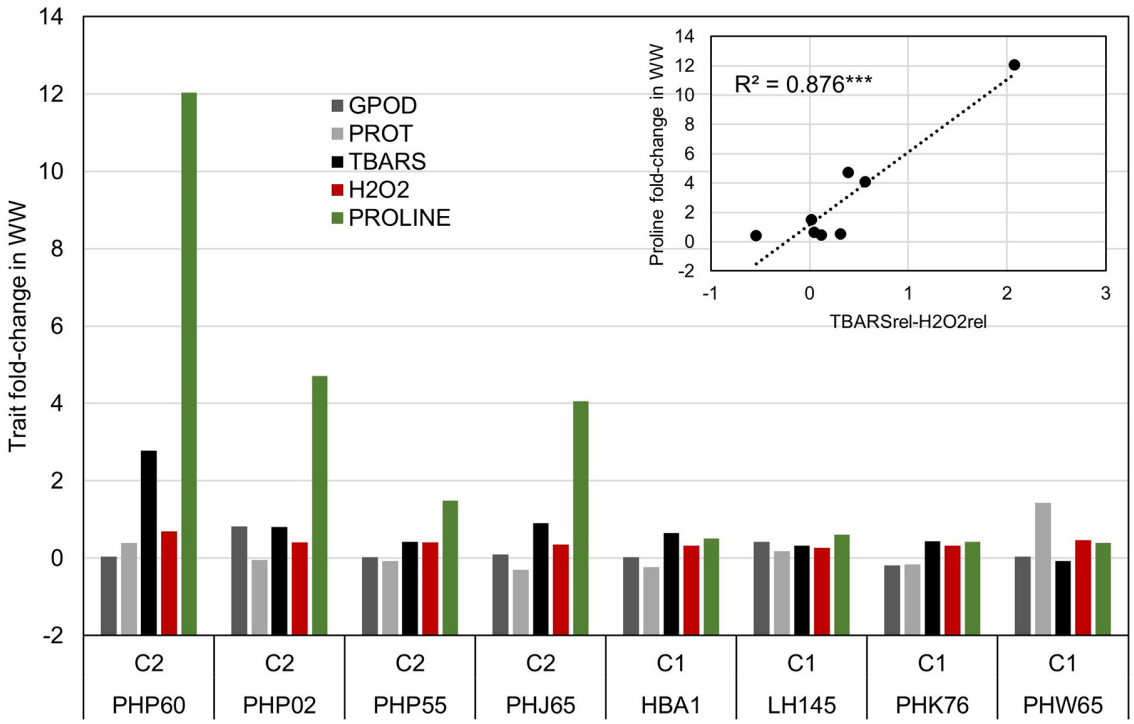


FIGURE 4 | Analysis of trait fold-change values in inbred representatives surrounding the centroids of clusters in **Figure 3** and the most extreme phenotypes in cluster one (PHW65) and two (PHP60). Linear model shows the relationship between the difference in trait fold-change values of TBARS and H_2O_2 and Proline accumulation in WW treatment in eight selected inbreds.

of responses to water withholding represents low-cost means for high throughput mass-screening of potentially favorable genotypes. Furthermore, it was found that the responses of maize to certain osmotic pressure were strongly correlated between controlled and field conditions (Chapuis et al., 2012).

The lowest repeatability of proline accumulation caused by the highest relative genotype by treatment interaction component is in accordance with proline manifold physiological functions and multiple pathways of biosynthesis (Verslues et al., 2014), along with differences in capacities of different genotypes to accumulate proline with increase of osmotic pressure (Khalil et al., 2016) causing the crossovers of genotype reactions. Increase in correlation strengths between different traits in WW treatment (**Figure 2A**) implies adaptive activation of plant physiological mechanisms to alleviate the stress effects (Bustos-Korts et al., 2018). Expectedly, reduction of plant CWC was negatively associated with TBARS, proline, and H_2O_2 (**Figure 2B**). The decrease in CWC value is expected in water deficit conditions (Avramova et al., 2016) and was shown to be negatively associated with photosynthetic efficiency and plant development in maize hybrids and antioxidant enzyme activity (Holá et al., 2017). The CWC reduction implies loss of leaf water, but without assessment of saturated leaf mass, thus providing only a loose estimate of leaf relative water content. However, its reduction is a good estimate of the effects of the drop in available soil water content especially in monitoring of drought development (Zhou et al., 2021).

Crossovers of Genotype Reactions Indicate TBARS Signaling Function

Water withholding stress induces stomatal closure leading to impaired CO_2 fixation and consequently excessive production of ROS such as hydrogen peroxide (H_2O_2) among others (Gill and Tuteja, 2010). To cope with water scarcity, plants developed different enzymatic and non-enzymatic mechanisms of ROS scavenging. Water withholding treatment caused generation of ROS in treated maize genotypes leading to increased levels of TBARS, H_2O_2, proline, and the activity of GPOD.

Previous studies demonstrated that H_2O_2 has dual role in plants by acting as a signaling molecule at low concentrations, thus triggering adaptation to stressful conditions (Gupta et al., 2016) while at higher concentrations can trigger programmed cell death (Gill and Tuteja, 2010). Its homeostasis is maintained at different cell parts and organelles such as peroxisomes by enzyme catalase (Hossain et al., 2015), in cytosol and chloroplasts by the ascorbate peroxidase (Guo et al., 2020b) and various peroxidases in mitochondria, such as guaiacol peroxidase (Tognolli et al., 2002). In this study, H_2O_2 levels showed induction of oxidative stress and the ROS scavenging capacity in maize genotypes exposed to water withholding stress. H_2O_2 is involved in numerous physiological processes in plants such as development, senescence, cell cycle, photosynthesis, and stomatal movement (Huang et al., 2019).

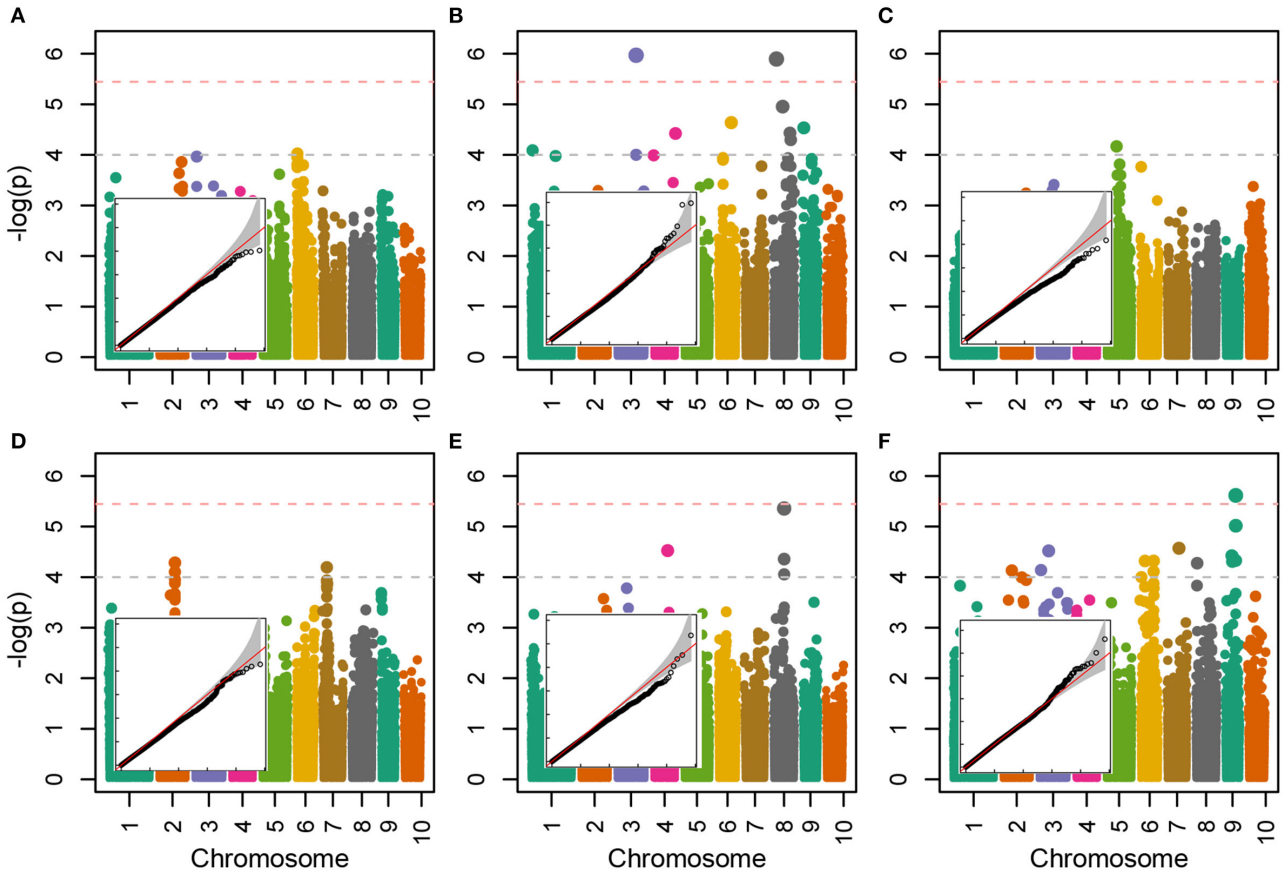


FIGURE 5 | Manhattan and quantile-quantile plots of the –log(p) values of associations of the trait fold-change values between CO and WW for CWC **(A)**, proline **(B)**, H_2O_2 **(C)**, PROT **(D)**, GPOD **(E)** and TBARS **(F)**. Horizontal dashed lines represent arbitrary threshold of 4 (gray) and Bonferroni threshold of 5.45 (light red).

Peroxidases are the enzymes responsible for scavenging hydrogen peroxide and reactive intermediary forms of O_2 under stress conditions. Guaiacol peroxidase (GPOD) is involved in oxidative stress response by catalyzing reduction of H_2O_2, thus decreasing its negative effects (Gill and Tuteja, 2010) using phenols (guaiacol) as a substrate. Another interesting feature of GPOD is the fact that it is involved in the process of lignification during the young plant development (Kidwai et al., 2020). It has been reported that peroxidases can play important role in ROS scavenging in maize under stressful conditions (Rohman et al., 2016), but GPOD is also a part of the antioxidant system involved in stress-acclimation resulting in transcriptional cell modifications (Gechev et al., 2006). Increase of peroxidase activity reduces ROS accumulation and also has the ability to consequently regulate the level of lipid peroxidation (Huang et al., 2019) to some extent.

Lipid peroxidation in parallel occurs in healthy organisms and is carried out both enzymatically (lipoxygenase activity) and non-enzymatically (ROS-mediated) (Farmer and Mueller, 2013), and thus, the change in TBARS accumulation in sub-optimal conditions is of true interest for understanding of trait implications. The activation of lipid peroxidation mechanism is characterized by three stages that include initiation, progression, and termination when ROS levels reach threshold on the cellular level, thus initiating production of lipid-derived radicals. Oxidation of lipids in the cell membranes leads to its instability by decrease of membrane fluidity and increase of membrane leakage which consequently inactivates different transport mechanisms such as receptors, ion channels, and enzymes (Gill and Tuteja, 2010). Rohman et al. (2016) reported higher MDA levels in maize seedlings more susceptible to drought stress while Zhang et al. (2014) showed that after exposure of maize seedlings to sudden drought stress, MDA levels were more increased in comparison with gradual drought stress. Although the increase in MDA values is usually interpreted in terms of stress damage, it was found that increase in lipid peroxidation can also shift the transcriptional profile of a cell (Weber et al., 2004). Furthermore, recently, it was speculated that the MDA might serve as a stress-signaling molecules in plants activating dehydration/heat shock-related genes and genes involved in antioxidant machinery (Morales and Munné-Bosch, 2019).

Accumulation of osmolytes, such as proline, under drought stress was reported in various plant species. Proline is a member of the glutamate family and plays versatile role in maintaining

TABLE 3 | Summary of detected associations for Tr_{rel} for GPOD (AU/g FW) PROT (mg/g FW), TBARS (ng/g FW), proline (nmol/mg FW), H_2O_2 (nmol/g FW) and CWC (%) crossing the –log(p) significance threshold of 4.

Trait	Name	Chr.	Pos.	–log(p)	Candidate gene name
CWC	CWC1@6	6	21457505	4.028	RNA-metabolizing metallo-beta-lactamase family protein
PRO	PROLINE1@1	1	9009032	4.09	beta amylase2
					Vegetative storage protein
	PROLINE2@3	**3**	**189739999**	**5.97**	Osmotin-like protein OSM34
	PROLINE3@3	3	189789986	4.003	NHL25
	PROLINE4@4	4	243872645	4.423	Peroxisomal membrane protein PEX16
					Probable phosphoinositide phosphatase SAC9
	PROLINE5@6	6	152847056	4.639	Coronatine-insensitive protein 1
	PROLINE6@8	**8**	**21838456**	**5.897**	Gigantea1
	PROLINE7@8	8	120343076	4.955	Protein NETWORKED 2D
	PROLINE8@8	8	162493033	4.433	NADPH:quinone oxidoreductase
					Protein SHI RELATED SEQUENCE 6
	PROLINE9@8	8	165274233	4.298	Delta3,5-delta2,4-dienoyl-CoA isomerase
	PROLINE10@9	9	12076699	4.535	Serine carboxypeptidase-like 29
H2O2	H2O21@5	5	74828595	4.169	Polyketide cyclase/dehydrase and lipid transport superfamily protein
PROT	PROT1@2	2	170426412	4.284	Aldo-keto reductase/ oxidoreductase
	PROT2@2	2	170444658	4.102	
	PROT3@7	7	33140990	4.194	UPF0235 protein
GPOD	GPOD1@4	4	181787137	4.524	Cyclin-dependent kinase inhibitor 1
	GPOD2@8	8	133675700	4.357	peroxidase64
	GPOD3@8	8	133675776	5.357	
	GPOD4@8	8	133819272	4.055	peroxidase2
TBARS	TBARS1@2	2	39693383	4.131	3-ketoacyl-CoA thiolase 2 peroxisomal
					Bowman-Birk type bran trypsin inhibitor
					Probable 2-oxoglutarate-dependent dioxygenase
	TBARS2@3	3	9767531	4.139	
	TBARS3@3	3	125562863	4.518	Lipid phosphate phosphatase delta
	TBARS4@6	6	17174793	4	Calreticulin-2
	TBARS5@6	6	90470962	4.319	GATA transcription factor 24
	TBARS6@6	6	154170989	4.32	Tubby-like F-box protein 10
	TBARS7@6	6	154176318	4.111	Glutathione S-transferase family protein
	TBARS8@7	7	138229143	4.571	Metallothionein-like protein 2A
	TBARS9@8	8	14999168	4.275	HSP20-like chaperones superfamily protein
					PASTICCINO 2
	TBARS10@9	9	104341296	4.424	Calcium-dependent lipid-binding (CaLB domain) plant phosphoribosyltransferase family protein
	TBARS11@9	9	110991638	4.415	Zinc finger protein CONSTANS-LIKE 5
	TBARS12@9	9	111784658	4.307	cryptochrome 3
	TBARS13@9	9	139304310	4.328	Hexosyltransferase
	TBARS14@9	**9**	**139353870**	**5.616**	
	TBARS15@9	9	139353877	5.016	

In bold are the associations crossing the calculated Bonferroni threshold of 5.45. Candidate gene IDs and gene names were selected from BioMart analysis. Details of the detected associations with candidate gene IDs and their functions are available in ***Supplementary Table S3****.*

water status, membrane stability, inhibiting protein oxidation and ROS scavenging under osmotic stress, thus contributing to drought-stress tolerance (Hayat et al., 2012). Genotypes exhibiting higher proline accumulation in water deficit are considered to be more drought tolerant (Gill and Tuteja, 2010; Ozturk et al., 2021). Yang et al. (2009) demonstrated that in

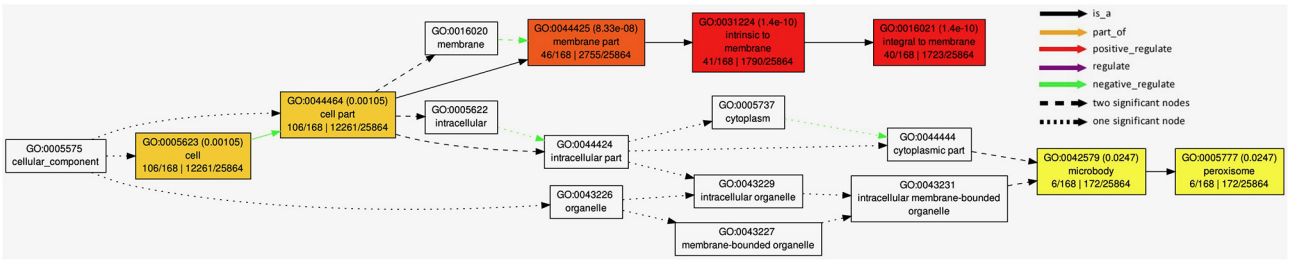


FIGURE 6 | Results of AgriGo gene ontology (GO) enrichment analysis with genes detected by BioMart tool (**Supplementary Table S3**) for cellular components.

maize seedlings, H_2O_2 could be involved in mechanisms of signal transduction that can lead to proline accumulation as well as regulation of its biosynthesis and degradation. Namely, proline accumulation induced by H_2O_2 is formed *via* combination of consecutive activation of the two biosynthetic pathways and inhibition of the proline degradation by H_2O_2. This was also confirmed in our study by the moderate positive significant correlation between proline and H_2O_2 in WW (**Figure 2A**) as well as their respective changes in WW compared to CO (**Figure 2B**).Our analysis of fold-changes between CO and WW indeed confirmed the relationship between the H_2O_2 build-up and proline accumulation as the top 10 scorers for H_2O_2 change in WW showed more than 2-fold increase in proline accumulation (**Figure 2C**). However, the top 10 scorers in TBARS accumulation showed even larger, more than 4-fold increase in proline accumulation. The division of the inbreds to two groups with fold change of H_2O_2 > TBARS (n = 27), and TBARS > H_2O_2 (n = 82) further confirmed the observed pattern. This finding is in accordance with the assumed signaling functions of MDA in the activation of heat shock and dehydration-related genes (Morales and Munné-Bosch, 2019), especially since there is a known relationship between heat shock proteins and proline accumulation in osmotic stress conditions (Augustine, 2016).

K-Means Clustering, Association Analysis, and Gene Ontology Enrichment Corroborate the Interplay Between H_2O_2 and TBARS in Osmotic Regulation

K-means clustering (**Figure 3**) showed that the assessed genotypes can be readily divided into two clusters of reactions to WW: one with larger change in trait values in WW compared to CO and another with lower change (**Table 2**) indicating the feasibility of division of cultivars to drought sensitive and drought tolerant (Gill and Tuteja, 2010) at this growth stage. Due to the lack of link between cluster designations and available pedigree and admixture data, K-means cluster affiliations were used as covariate to control FDR, thus compensating for relatively modest sample size (Aschard et al., 2017). Interestingly, cluster 2 was mostly populated by inbreds from TBARS > H_2O_2 group, bearing only seven accessions of the H_2O_2 > TBARS group, scattered near zero at both dimensions. This corroborated our speculation of the involvement of TBARS in proline accumulation, as inbreds belonging to cluster 2 showed 3.4 times increase in proline content in WW treatment, and 54.2% larger increase in TBARS compared to inbreds in cluster 1. It is also worth mentioning that the accession (PHW65) with the most extreme phenotype considering all analyzed traits in cluster 1 belonged to H_2O_2 > TBARS group, whereas in cluster 2, inbred PHP60, showing the most extreme phenotype, belonged to the group TBARS > H_2O_2. Noteworthy, the three nearest accessions to the centroid of the cluster 2, as well as the most extreme phenotype of cluster 2, belong to the Iodent breeding pool, namely, PHP60, PHP55, PHP02, and PHJ65, while closest to the centroid of the cluster 1, the inbreds HBA1 and LH145, belonging to Lancaster and B14 breeding pools, are found. Iodent group represents a relatively novel germplasm pool, which was not commercially utilized prior to 1980s (Mikel and Dudley, 2006) and might thus be associated with the genetic variability underlying the detected signals, especially as the more recent breeding efforts were inclined toward drought tolerance (Troyer, 2009). The Iodent breeding pool might thus indeed carry the favorable alleles for osmotic stress tolerance. This was confirmed in a recent study on 209 diverse inbred lines at seedling stage, where Iodent pool founder line, PH207, was selected among the four most drought tolerant accessions (Guo et al., 2020a). It is worth to emphasize that all except one inbred (PHW65) in the **Figure 4** belong to TBARS > H_2O_2 group, so the involvement of lipid peroxidation—hydrogen peroxide crosstalk cannot be unambiguously pinpointed as a sole contributor to the proline accumulation. However, inspection of the regression results in this group of inbreds, with very high coefficient of determination between difference between H_2O_2 and TBARS fold change in water withholding and proline accumulation, along with other aforementioned patterns, indicates possible important role of this pathway in osmoprotection of young plants.

Details on the associations detected in association analysis (**Figure 5**; **Table 3**) are available in **Supplementary Table S2** available online, along with most probable gene candidates identified from full gene list from BioMart (Kinsella et al., 2011) scan (**Supplementary Table S3**). The genes identified in BioMart were used in AgriGo online mining tool (Tian et al., 2017) for the gene ontology enrichment analysis. The highly significant negative regulation of intrinsic membrane parts revealed from genes detected in association analysis in GO enrichment analysis (**Figure 6**) was probably caused by the damage to

the cell membrane causing membrane lipid-bilayer distortion. The distortion of the lipid-bilayer affects the functions of the surrounded proteins, causing the alteration of their functionality (Lee, 2004). The detected enrichment was supported by the involvement of 40 detected genes in gene network regulating these components. Other negatively regulated cell components were peroxisomes. Interestingly, H_2O_2 is a well-known stress signaling molecule, with peroxisomes being their main source (Hossain et al., 2015; Su et al., 2019). It is possible that the negative regulation of the resource of this important signaling molecule represents the part of the signaling cascade causing the differences in proline accumulation between groups observed in **Figures 2B,C**. This is further corroborated by the rates of accumulation of H_2O_2 in organelles. For example, the peroxide accumulation in mitochondria does not show large variation throughout the day, whereas the rates of its formation can be 30 to 100 times higher in chloroplasts and peroxisomes (Hossain et al., 2015). Furthermore, H_2O_2 synthesis in peroxisomes is associated with photorespiration, or, more specifically, oxidation of glycolate during the photosynthetic carbon oxidation cycle (Niu and Liao, 2016), impacting the gene expression and metabolic enzyme activity. The crosstalk in proline synthesis between lipid peroxidation and H_2O_2 is further corroborated by the genes linked to associations TBARS10@9 and TBARS13@9-TBARS15@9, coding for calcium-dependent lipid-binding (CaLB domain) plant phosphoribosyltransferase family protein and hexosyltransferase, both acting as glycosyl transferases in our study (**Supplementary Table S2**). On the other hand, the study by Ben Rejeb et al. (2015) showed that accumulation of the H_2O_2 generated by the NADPH oxidases in *Arabidospis thaliana* plants increases the proline biosynthesis by positive regulation of proline biosynthesis genes. Interestingly, the association detected in our study (PROLINE8@8) was located in the region harboring the NADPH quinone oxidoreductase linked to NADPH oxidation, also involved in the detoxication process of lipid peroxides (Mano et al., 2002), possibly representing another link between the TBARS and proline accumulation. Also, the dehydration response genes were detected, such as OSM34 coding for osmotin, related to association PROLINE2@3. Osmotines are plant sentinel proteins expressed during the osmotic stress, providing plant cell the means to retain osmolarity by metabolic changes and solutes compartmentalization (Ozturk et al., 2021), such as proline. There is abundant evidence that the expression of osmotin genes triggers proline accumulation in osmotic stress in many species (for review, see Anil Kumar et al., 2015). Other genes that possibly corroborate the relationship assumed by Morales and Munné-Bosch (2019) that the MDA might act as signaling molecules in stress are heat shock-related genes such as HSP20 linked to association TBARS9@8 (**Figure 5**; **Table 3**). It was found that the heat shock proteins, acting as chaperones, are readily expressed in osmotic stress conditions, helping in binding, folding, displacing, and degrading other proteins (Ozturk et al., 2021). Furthermore, there is a known crosstalk between heat shock proteins and proline in heat stress, mediated by nitric oxide (Alamri et al., 2019), possibly active in osmotic stress as well.

Most interestingly, the correlation analysis (**Table 2**) showed the lower correlation between TBARS and H_2O_2 in TBARS > H_2O_2 group compared to correlation observed in H_2O_2 > TBARS, possibly implying the enzymatic rather than ROS-mediated origin of TBARS (Farmer and Mueller, 2013). This was accompanied by the increase in correlation strength between TBARS and proline in TBARS > H_2O_2 group implying the activation of additional proline synthesis mechanisms. The H_2O_2 and lipid peroxidation homeostasis might thus play a critical role in proline synthesis. In a study by Terzi et al. (2014), it was shown that H_2O_2 treatment in maize affects levels of proline and MDA in leaves in osmotic stress conditions. The process might be mediated through activation of polysaccharide catabolic processes, as proline synthesis is an energy-consuming task. The differentially regulated polysaccharide catabolism (**Supplementary Figure S3**) covered the gene related to detected association PROLINE1@1 (**Table 3**) coding for beta-amylase2, and the beta amylases are known for their role in providing energy for proline synthesis in drought stress (Zanella et al., 2016). Finally, it was found that 24 of 168 genes analyzed in AgriGo tool were involved in significant enrichment of regulation of DNA-dependent transcription, causing the transcriptional reprogram of cells.

In conclusion, the reactions of inbred lines assessed in this study allowed detection of a potentially significant regulatory signaling mechanism for response to water withholding at young plant growth stage. Namely, the results of this mass-screening indicate that in response to water withholding, the group of inbreds in which accumulation of products of lipid peroxidation (TBARS) surpassed the accumulation of H_2O_2, a well-known signaling molecule, in average increased their proline content by nearly a double. Furthermore, the analysis of inbred responses to water withholding per se corroborated by the K-means cluster analysis showed lower variability in all assessed biochemical traits in H_2O_2 > TBARS group accompanied by nearly no change in proline content. For example, the representatives of the two K-means clusters showed radical differences in proline accumulation, where inbred PHP60 (cluster 2, TBARS > H_2O_2) was able to increase proline content nearly 12-folds in WW, whereas inbred PHW65 (cluster 1, H_2O_2 > TBARS) produced barely detectable response to osmotic changes detectable by the used methods. The association mapping combined with gene ontology enrichment analysis showed significant phenotypic effects of the linkage regions harboring genes involved in osmotic-stress signaling and osmolyte accumulation, as well as negative regulation of peroxisomes, corroborating the phenotypic analysis. However, the contrasting responses to water withholding of the two groups of inbreds in this study do not necessarily reflect to the final (agronomic) performance of the genotypes, so three important aspects have to be elucidated in further research to establish the implications of these findings: (i) the dynamic analysis in genotypes with contrasting responses (e.g., top 10 and bottom 10 TBARS scorers) needs to capture the switching of this mechanism, and the difference in dynamics of proline accumulation between the groups; (ii) the transcriptome of known regulatory genes, and the implications of this cascade through the developmental cycle haves to be studied, and

(iii) the field trials in relevant environments need to provide the connection between the contrasting performance at early growth and agronomic performance in terms of yield quantity and quality.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**. Genomic data for the analyzed accessions is available as supplementary data at: https://doi.org/10.3390/plants9020275.

AUTHOR CONTRIBUTIONS

VG, SM, LB, and DŠ conceived the study and prepared the first manuscript draft. VG, ZZ, AB, and MMaz conducted the experiments. SM, MMar, and LB performed laboratory analyses. VG analyzed results. ZZ and DŠ acquired funding. All authors contributed to the article and approved the submitted version.

FUNDING

This research was funded by the EU project Biodiversity and Molecular Plant Breeding, Grant Number KK.01.1.1.01.0005, of the Centre of Excellence for Biodiversity and Molecular Plant Breeding (CroP-BioDiv), Zagreb, Croatia.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.804630/full#supplementary-material

REFERENCES

Alamri, S. A., Siddiqui, M. H., Al-Khaishany, M. Y., Khan, M. N., Ali, H. M., and Alakeel, K. A. (2019). Nitric oxide-mediated cross-talk of proline and heat shock proteins induce thermotolerance in *Vicia faba* L. *Environ. Exp. Bot.* 161, 290–302. doi: 10.1016/j.envexpbot.2018.06.012

Anil Kumar, S., Hima Kumari, P., Shravan Kumar, G., Mohanalatha, C., and Kavi Kishor, P. B. (2015). Osmotin: a plant sentinel and a possible agonist of mammalian adiponectin. *Front. Plant Sci.* 6, 163. doi: 10.3389/fpls.2015.00163

Anjum, S. A., Ashraf, U., Tanveer, M., Khan, I., Hussain, S., Shahzad, B., et al. (2017). Drought induced changes in growth, osmolyte accumulation and antioxidant metabolism of three maize hybrids. *Front. Plant Sci.* 8, 69. doi: 10.3389/fpls.2017.00069

Aschard, H., Guillemot, V., Vilhjalmsson, B., Patel, C. J., Skurnik, D., Ye, C. J., et al. (2017). Covariate selection for association screening in multiphenotype genetic studies. *Nat. Genet.* 49, 1789–1795. doi: 10.1038/ng.3975

Aslam, M., Maqbool, M. A., and Cengiz, R. (2015). *Drought Stress in Maize (Zea mays L.)* Cham, Germany: Springer.

Augustine, S. (2016). "Function of heat-shock proteins in drought tolerance regulation of plants," in Drought Stress Tolerance in Plants, Vol 1: Physiology and Biochemistry, eds M. A. Hossain, S. H. Wani, S. Bhattacharjee, D. J. Burritl, and L. S. P. Tran (Cham, Germany: Springer), 163–185.

Avramova, V., Nagel, K. A., Abdelgawad, H., Bustos, D., Duplessis, M., Fiorani, F., et al. (2016). Screening for drought tolerance of maize hybrids by multi-scale analysis of root and shoot traits at the seedling stage. *J. Exp. Bot.* 67, 2453–2466. doi: 10.1093/jxb/erw055

Ben Rejeb, K., Lefebvre-De Vos, D., Le Disquet, I., Leprince, A. S., Bordenave, M., Maldiney, R., et al. (2015). Hydrogen peroxide produced by NADPH oxidases increases proline accumulation during salt or mannitol stress in Arabidopsis thaliana. *New Phytol.* 208, 1138–1148. doi: 10.1111/nph.13550

Bian, Y., and Holland, J. B. (2017). Enhancing genomic prediction with genome-wide association studies in multiparental maize populations. *Heredity* 118, 585–593. doi: 10.1038/hdy.2017.4

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248–254. doi: 10.1016/0003-2697(76)90527-3

Bustos-Korts, D., Romagosa, I., Borràs-Gelonch, G., Casas, A. M., Slafer, G. A., and Van Eeuwijk, F. A. (2018). "Genotype by environment interaction and adaptation," in *Encyclopedia of Sustainability Science and Technology*, eds R. Savin and G. A. Slafer (Berlin: Springer Science+Business Media, LLC), 29–71.

Canaran, P., Buckler, E. S., Glaubitz, J. C., Stein, L., Sun, Q., Zhao, W., et al. (2008). Panzea: an update on new content and features. *Nucleic Acids Res.* 36, 2007–2009. doi: 10.1093/nar/gkm1022

Carillo, P., and Gibon, Y. (2011). *PROTOCOL: Extraction and Determination of Proline*. PrometheusWiki. Available online at: https://prometheusprotocols.net/function/tissue-chemistry/primary-metabolites/extraction-and-determination-of-proline/ (accessed July 10, 2022).

Chapuis, R., Delluc, C., Debeuf, R., Tardieu, F., and Welcker, C. (2012). Resiliences to water deficit in a phenotyping platform and in the field: how related are they in maize? *Eur. J. Agron.* 42, 59–67. doi: 10.1016/j.eja.2011.12.006

Charrad, M., Ghazzali, N., Boiteau, V., and Niknafs, A. (2014). Nbclust: an R package for determining the relevant number of clusters in a data set. *J. Stat. Softw.* 61, 1–36. doi: 10.18637/jss.v061.i06

Corpas, F. J., González-Gordo, S., and Palma, J. M. (2020). Plant peroxisomes: a factory of reactive species. *Front. Plant Sci.* 11, 853. doi: 10.3389/fpls.2020.00853

Covarrubias-Pazaran, G. (2016). Genome-assisted prediction of quantitative traits using the r package sommer. *PLoS ONE* 11, e0156744. doi: 10.1371/journal.pone.0156744

Czarnocka, W., and Karpiński, S. (2018). Friend or foe? Reactive oxygen species production, scavenging and signaling in plant response to environmental stresses. *Free Radic. Biol. Med.* 122, 4–20. doi: 10.1016/j.freeradbiomed.2018.01.011

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., et al. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* 6, e19379. doi: 10.1371/journal.pone.0019379

Fan, Y., Shabala, S., Ma, Y., Xu, R., and Zhou, M. (2015). Using QTL mapping to investigate the relationships between abiotic stress tolerance (drought and salinity) and agronomic and physiological traits. *BMC Genomics* 16, 1–11. doi: 10.1186/s12864-015-1243-8

FAO/IHS Markit Agribusiness Consulting (2019). *Analysis on Sales and Profitability Within the Seed Sector*. London, UK.

Farmer, E. E., and Mueller, M. J. (2013). ROS-mediated lipid peroxidation and RES-activated signaling. *Annu. Rev. Plant Biol.* 64, 429–450. doi: 10.1146/annurev-arplant-050312-120132

Farooq, M., Hussain, M., Wahid, A., and Siddique, K. H. M. (2012). "Drought stress in plants: an overview," in *Plant Responses to Drought Stress*, ed R. Aroca (Berlin: Springer Berlin Heidelberg), 1–33.

Galić, V., Mazur, M., Brkić, A., Brkić, J., Jambrović, A., Zdunić, Z., et al. (2020). Seed weight as a covariate in association and prediction studies for biomass traits in maize seedlings. *Plants* 9, 1–18. doi: 10.3390/plants9020275

Gao, X., Becker, L. C., Becker, D. M., Starmer, J. D., and Province, M. A. (2010). Avoiding the high bonferroni penalty in genome-wide association studies. *Genet. Epidemiol.* 34, 100–105. doi: 10.1002/gepi.20430

Gao, X., Starmer, J., and Martin, E. R. (2008). A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genet. Epidemiol.* 32, 361–369. doi: 10.1002/gepi.20310

Gechev, T. S., Van Breusegem, F., Stone, J. M., Denev, I., and Laloi, C. (2006). Reactive oxygen species as signals that modulate plant stress responses and programmed cell death. *BioEssays* 28, 1091–1101. doi: 10.1002/bies.20493

Gill, M. B., Zeng, F., Shabala, L., Zhang, G., Yu, M., Demidchik, V., et al. (2019). Identification of QTL related to ROS formation under hypoxia and their association with waterlogging and salt tolerance in Barley. *Int. J. Mol. Sci.* 20, 699. doi: 10.3390/ijms20030699

Gill, S. S., and Tuteja, N. (2010). Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. *Plant Physiol. Biochem.* 48, 909–930. doi: 10.1016/j.plaphy.2010.08.016

Grillakis, M. G. (2019). Increase in severe and extreme soil moisture droughts for Europe under climate change. *Sci. Total Environ.* 660, 1245–1255. doi: 10.1016/j.scitotenv.2019.01.001

Guo, J., Li, C., Zhang, X., Li, Y., Zhang, D., Shi, Y., et al. (2020a). Transcriptome and GWAS analyses reveal candidate gene for seminal root length of maize seedlings under drought stress. *Plant Sci.* 292, 110380. doi: 10.1016/j.plantsci.2019.110380

Guo, K., Li, Z., Tian, H., Du, X., Liu, Z., Huang, H., et al. (2020b). Cytosolic ascorbate peroxidases plays a critical role in photosynthesis by modulating reactive oxygen species level in stomatal guard cell. *Front. Plant Sci.* 11, 446. doi: 10.3389/fpls.2020.00446

Gupta, K., Sengupta, A., Chakraborty, M., and Gupta, B. (2016). Hydrogen peroxide and polyamines act as double edged swords in plant abiotic stress responses. *Front. Plant Sci.* 7, 1343. doi: 10.3389/fpls.2016.01343

Hayat, S., Hayat, Q., Alyemeni, M. N., Wani, A. S., Pichtel, J., and Ahmad, A. (2012). Role of proline under changing environments: a review. *Plant Signal. Behav.* 7, 1456–1466. doi: 10.4161/psb.21949

Holá, D., Benešová, M., Fischer, L., Haisel, D., Hnilicka, F., Hnilicková, H., et al. (2017). The disadvantages of being a hybrid during drought: a combined analysis of plant morphology, physiology and leaf proteome in maize. *PLoS ONE* 12, e0176121. doi: 10.1371/journal.pone.0176121

Hossain, M. A., Bhattacharjee, S., Armin, S. M., Qian, P., Xin, W., Li, H. Y., et al. (2015). Hydrogen peroxide priming modulates abiotic oxidative stress tolerance: insights from ROS detoxification and scavenging. *Front. Plant Sci.* 6, 420. doi: 10.3389/fpls.2015.00420

Howe, K. L., Contreras-Moreira, B., De Silva, N., Maslen, G., Akanni, W., Allen, J., et al. (2020). Ensembl genomes 2020-enabling non-vertebrate genomic research. *Nucleic Acids Res.* 48, D689–D695. doi: 10.1093/nar/gkz890

Huang, H., Ullah, F., Zhou, D. X., Yi, M., and Zhao, Y. (2019). Mechanisms of ROS regulation of plant development and stress responses. *Front. Plant Sci.* 10, 800. doi: 10.3389/fpls.2019.00800

Jambunathan, N. (2010). "Plant stress tolerance," in *NIH Public Access*, ed R. Sunkar (Berlin/Heidelberg, Germany: Springer Science+Business Media, LLC) 1–14.

Jang, S., Han, J. H., Lee, Y. K., Shin, N. H., Kang, Y. J., Kim, C. K., et al. (2020). Mapping and validation of QTLs for the amino acid and total protein content in brown rice. *Front. Genet.* 11, 240. doi: 10.3389/fgene.2020.00240

Jiang, J., Zhuang, J. Y., Fan, Y. Y., and Bo, S. H. (2009). Mapping of QTLs for leaf malondialdehyde content associated with stress tolerance in rice. *Rice Sci.* 16, 72–74. doi: 10.1016/S1672-6308(08)60059-1

Kassambara, A., and Mundt, F. (2020). *factoextra: Extract and Visualize the Results of Multivariate Data Analyses*. Available online at: https://cran.r-project.org/package=factoextra (accessed June 10, 2022).

Khalil, F., Rauf, S., Monneveux, P., Anwar, S., and Iqbal, Z. (2016). Genetic analysis of proline concentration under osmotic stress in sunflower (*Helianthus annuus* L.). *Breed. Sci.* 66, 463–470. doi: 10.1270/jsbbs.15068

Kidwai, M., Ahmad, I. Z., and Chakrabarty, D. (2020). Class III peroxidase: an indispensable enzyme for biotic/abiotic stress tolerance and a potent candidate for crop improvement. *Plant Cell Rep.* 39, 1381–1393. doi: 10.1007/s00299-020-02588-y

Kinsella, R. J., Kähäri, A., Haider, S., Zamora, J., Proctor, G., Spudich, G., et al. (2011). Ensembl BioMarts: a hub for data retrieval across taxonomic space. *Database* 2011, 1–9. doi: 10.1093/database/bar030

Kumar, I. S., and Nadarajah, K. (2020). A meta-analysis of quantitative trait loci associated with multiple disease resistance in rice (*Oryza sativa* L.). *Plants* 9, 1–28. doi: 10.3390/plants9111491

Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* 82, 1–26. doi: 10.18637/jss.v082.i13

Lee, A. G. (2004). How lipids affect the activities of integral membrane proteins. *Biochim. Biophys. Acta Biomembr.* 1666, 62–87. doi: 10.1016/j.bbamem.2004.05.012

Lee, E. A., and Tracy, W. F. (2009). "Modern maize breeding," in *Handbook of Maize: Genetics and Genomics*, eds J. Bennetzen and S. Hake (Berlin/Heidelberg: Springer Science+Business Media, LLC), 151–160.

Ma, J., Du, G., Li, X., Zhang, C., and Guo, J. (2015). A major locus controlling malondialdehyde content under water stress is associated with Fusarium crown rot resistance in wheat. *Mol. Genet. Genomics* 290, 1955–1962. doi: 10.1007/s00438-015-1053-3

Mano, J., Torii, Y., Hayashi, S., ichiro, Takimoto, K., Matsui, K., Nakamura, K., et al. (2002). The NADPH: Quinone oxidoreductase P1-ζ-crystallin in Arabidopsis catalyzes the α,β-hydrogenation of 2-alkenals: detoxication of the lipid peroxide-derived reactive aldehydes. *Plant Cell Physiol.* 43, 1445–1455. doi: 10.1093/pcp/pcf187

Masuka, B., Araus, J. L., Das, B., Sonder, K., and Cairns, J. E. (2012). Phenotyping for abiotic stress tolerance in maize. *J. Integr. Plant Biol.* 54, 238–249. doi: 10.1111/j.1744-7909.2012.01118.x

Mikel, M. A. (2011). Genetic composition of contemporary U.S. commercial dent corn germplasm. *Crop Sci.* 51, 592–599. doi: 10.2135/cropsci2010.06.0332

Mikel, M. A., and Dudley, J. W. (2006). Evolution of North American dent corn from public to proprietary germplasm. *Crop Sci.* 46, 1193–1205. doi: 10.2135/cropsci2005.10-0371

Money, D., Gardner, K., Migicovsky, Z., Schwaninger, H., Zhong, G. Y., and Myles, S. (2015). LinkImpute: fast and accurate genotype imputation for nonmodel organisms. *G3 Genes Genomes Genet.* 5, 2383–2390. doi: 10.1534/g3.115.021667

Morales, M., and Munné-Bosch, S. (2019). Malondialdehyde: facts and artifacts. *Plant Physiol.* 180, 1246–1250. doi: 10.1104/pp.19.00405

Mukherjee, S. P., and Choudhuri, M. A. (1983). Implications of water stress-induced changes in the levels of endogenous ascorbic acid and hydrogen peroxide in Vigna seedlings. *Physiol. Plant.* 58, 166–170. doi: 10.1111/j.1399-3054.1983.tb04162.x

Niu, L., and Liao, W. (2016). Hydrogen peroxide signaling in plant development and abiotic responses: crosstalk with nitric oxide and calcium. *Front. Plant Sci.* 7, 230. doi: 10.3389/fpls.2016.00230

Ozturk, M., Turkyilmaz Unal, B., García-Caparrós, P., Khursheed, A., Gul, A., and Hasanuzzaman, M. (2021). Osmoregulation and its actions during the drought stress in plants. *Physiol. Plant.* 172, 1321–1335. doi: 10.1111/ppl.13297

Pan, Q., Wei, J., Guo, F., Huang, S., Gong, Y., Liu, H., et al. (2019). Trait ontology analysis based on association mapping studies bridges the gap between crop genomics and Phenomics. *BMC Genomics* 20, 1–13. doi: 10.1186/s12864-019-5812-0

Pareek, A., Sopory, K., and Bohnert, H. J. (2010). *Abiotic Stress Adaptation in Plants*. Dordrecht: Springer.

Prasad, T. K., Andersen, M. D., and Stewart, C. R. (1995). Localization and characterization of peroxidases in the mitochondria of chilling-acclimated maize seedlings. *Plant Physiol.* 108, 1597–1605. doi: 10.1104/pp.108.4.1597

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795

Rohman, M. M., Begum, S., Talukder, M. Z. A., Akhi, A. H., Amiruzzaman, M., Ahsan, A. F. M. S., et al. (2016). Drought sensitive maize inbred shows more oxidative damage and higher ROS scavenging enzymes, but not glyoxalases than a tolerant one at seedling stage. *Plant Omics* 9, 220–232. doi: 10.21475/poj.16.09.04.pne31

Romay, M. C., Millard, M. J., Glaubitz, J. C., Peiffer, J. A., Swarts, K. L., Casstevens, T. M., et al. (2013). *Comprehensive Genotyping of the USA National Maize Inbred Seed Bank*. Available online at: http://genomebiology.com/content/pdf/gb-2013-14-6-r55.pdf (accessed June 10, 2022).

Sayed, M. A., Schumann, H., Pillen, K., Naz, A. A., and Léon, J. (2012). AB-QTL analysis reveals new alleles associated to proline accumulation and leaf wilting under drought stress conditions in barley (*Hordeum vulgare* L.). *BMC Genet.* 13, 1–2. doi: 10.1186/1471-2156-13-61

Siegel, B. Z., and Galston, A. W. (1967). The isoperoxidases of *Pisum sativum*. *Plant Physiol.* 42, 221–226. doi: 10.1104/pp.42.2.221

Slafer, G. A., and Araus, J. L. (2007). "Physiological traits for improving wheat yield under a wide range of conditions," in *Scale and Complexity in Plant Systems Research*, eds J. H. J. Spiertz, P. C. Struik, and H. H. van Laar (Dordrecht: Springer Netherlands), 147–156.

Stagge, J. H., Kingston, D. G., Tallaksen, L. M., and Hannah, D. M. (2017). Observed drought indices show increasing divergence across Europe. *Sci. Rep.* 7, 1–10. doi: 10.1038/s41598-017-14283-2

Su, T., Li, W., Wang, P., and Ma, C. (2019). Dynamics of peroxisome homeostasis and its role in stress response and signaling in plants. *Front. Plant Sci.* 10, 705. doi: 10.3389/fpls.2019.00705

Tardieu, F. (2012). Any trait or trait-related allele can confer drought tolerance: just design the right drought scenario. *J. Exp. Bot.* 63, 25–31. doi: 10.1093/jxb/err269

Terzi, R., Kadioglu, A., Kalaycioglu, E., and Saglam, A. (2014). Hydrogen peroxide pretreatment induces osmotic stress tolerance by influencing osmolyte and abscisic acid levels in maize leaves. *J. Plant Interact.* 9, 559–565. doi: 10.1080/17429145.2013.871077

Tian, T., Liu, Y., Yan, H., You, Q., Yi, X., Du, Z., et al. (2017). AgriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res.* 45, W122–W129. doi: 10.1093/nar/gkx382

Tognolli, M., Penel, C., Greppin, H., and Simon, P. (2002). Analysis and expression of the class III peroxidase large gene family in Arabidopsis thaliana. *Gene* 288, 129–138. doi: 10.1016/S0378-1119(02)00465-1

Troyer, A. F. (2009). "Development of hybrid corn and the seed corn industry," in *Handbook of Maize: Genetics and Genomics*, eds J. L. Bennetzen, and S. Hake (New York, NY: Springer).

Verslues, P. E., Lasky, J. R., Juenger, T. E., Liu, T. W., and Nagaraj Kumar, M. (2014). Genome-wide association mapping combined with reverse genetics identifies new effectors of low water potential-induced proline accumulation in Arabidopsis. *Plant Physiol.* 164, 144–159. doi: 10.1104/pp.113.224014

Wang, B., Liu, C., Zhang, D., He, C., Zhang, J., and Li, Z. (2019). Effects of maize organ-specific drought stress response on yields from transcriptome analysis. *BMC Plant Biol.* 19, 335. doi: 10.1186/s12870-019-1941-5

Wang, J. Z., Du, Z., Payattakool, R., Yu, P. S., and Chen, C. F. (2007). A new method to measure the semantic similarity of GO terms. *Bioinformatics* 23, 1274–1281. doi: 10.1093/bioinformatics/btm087

Warnes, G. R., Bolker, B. M., Bonebakker, L., Gentleman, Robert, Liaw, A., Lumley, T., et al. (2013). *Package 'gplots.'* R package version 3.1.3, CRAN, Available online at: https://CRAN.R-project.org/package=gplots (accessed June 10, 2022).

Weber, H., Chételat, A., Reymond, P., and Farmer, E. E. (2004). Selective and powerful stress gene expression in Arabidopsis in response to malondialdehyde. *Plant J.* 37, 877–888. doi: 10.1111/j.1365-313X.2003.02013.x

White, M. R., Mikel, M. A., de Leon, N., and Kaeppler, S. M. (2020). Diversity and heterotic patterns in North American proprietary dent maize germplasm. *Crop Sci.* 60, 100–114. doi: 10.1002/csc2.20050

Yang, S. L., Lan, S. S., and Gong, M. (2009). Hydrogen peroxide-induced proline and metabolic pathway of its accumulation in maize seedlings. *J. Plant Physiol.* 166, 1694–1699. doi: 10.1016/j.jplph.2009.04.006

Yasir, M., He, S., Sun, G., Geng, X., Pan, Z., Gong, W., et al. (2019). A genome-wide association study revealed key SNPs/genes associated with salinity stress tolerance in upland cotton. *Genes* 10, 829. doi: 10.3390/genes10100829

Zanella, M., Borghi, G. L., Pirone, C., Thalmann, M., Pazmino, D., Costa, A., et al. (2016). β-amylase 1 (BAM1) degrades transitory starch to sustain proline biosynthesis during drought stress. *J. Exp. Bot.* 67, 1819–1826. doi: 10.1093/jxb/erv572

Zhang, L. X., Lai, J. H., Liang, Z. S., and Ashraf, M. (2014). Interactive effects of sudden and gradual drought stress and foliar-applied glycinebetaine on growth, water relations, osmolyte accumulation and antioxidant defence system in two maize cultivars differing in drought tolerance. *J. Agron. Crop Sci.* 200, 425–433. doi: 10.1111/jac.12081

Zhou, H., Zhou, G., He, Q., Zhou, L., Ji, Y., and Lv, X. (2021). Capability of leaf water content and its threshold values in reflection of soil–plant water status in maize during prolonged drought. *Ecol. Indic.* 124, 107395. doi: 10.1016/j.ecolind.2021.107395

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 18 July 2022
doi: 10.3389/fpls.2022.905842

Genome-Scale Profiling and High-Throughput Analyses Unravel the Genetic Basis of Arsenic Content Variation in Rice

Sang-Beom Lee[1], Gyeong-Jin Kim[2], Jung-Du Shin[3], Woojin Chung[4], Soo-Kwon Park[1], Geun-Hyoung Choi[2], Sang-Won Park[5] and Yong-Jin Park[6]**

[1] Crop Foundation Research Division, National Institute of Crop Science, Wanju, South Korea, [2] Residual Agrochemical Assessment Division, National Institute of Agriculture Science, Wanju, South Korea, [3] Bio-Technology of Multidisciplinary Sciences Co., Wanju, South Korea, [4] Department of Environmental Energy Engineering, Kyonggi University, Suwon, South Korea, [5] Reserch Policy Bureau, Rural Development Administration, Wanju, South Korea, [6] Department of Plant Resources, College of Industrial Sciences, Kongju National University, Yesan, South Korea

OPEN ACCESS

Edited by:
Mukesh Jain,
Jawaharlal Nehru University, India

Reviewed by:
Xin-Yuan Huang,
Nanjing Agricultural University, China
Xuehai Zhang,
Henan Agricultural University, China

***Correspondence:**
Yong-Jin Park
yjpark@kongju.ac.kr
Sang-Won Park
swpark@korea.kr

Specialty section:
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

Received: *28 March 2022*
Accepted: *15 June 2022*
Published: *18 July 2022*

Citation:
Lee S-B, Kim G-J, Shin J-D, Chung W, Park S-K, Choi G-H, Park S-W and Park Y-J (2022) Genome-Scale Profiling and High-Throughput Analyses Unravel the Genetic Basis of Arsenic Content Variation in Rice. Front. Plant Sci. 13:905842. doi: 10.3389/fpls.2022.905842

Ionomics, the study of the composition of mineral nutrients and trace elements in organisms that represent the inorganic component of cells and tissues, has been widely studied to explore to unravel the molecular mechanism regulating the elemental composition of plants. However, the genetic factors of rice subspecies in the interaction between arsenic and functional ions have not yet been explained. Here, the correlation between As and eight essential ions in a rice core collection was analyzed, taking into account growing condition and genetic factors. The results demonstrated that the correlation between As and essential ions was affected by genetic factors and growing condition, but it was confirmed that the genetic factor was slightly larger with the heritability for arsenic content at 53%. In particular, the cluster coefficient of *japonica* (0.428) was larger than that of *indica* (0.414) in the co-expression network analysis for 23 arsenic genes, and it was confirmed that the distance between genes involved in As induction and detoxification of *japonica* was far than that of *indica*. These findings provide evidence that *japonica* populations could accumulate more As than *indica* populations. In addition, the *cis*-eQTLs of AIR2 (arsenic-induced RING finger protein) were isolated through transcriptome-wide association studies, and it was confirmed that AIR2 expression levels of *indica* were lower than those of *japonica*. This was consistent with the functional haplotype results for the genome sequence of AIR2, and finally, eight rice varieties with low AIR2 expression and arsenic content were selected. In addition, As-related QTLs were identified on chromosomes 5 and 6 under flooded and intermittently flooded conditions through genome-scale profiling. Taken together, these results might assist in developing markers and breeding plans to reduce toxic element content and breeding high-quality rice varieties in future.

Keywords: arsenic, ionomics, eQTLs, co-expression network, transcriptome-wide association studies

INTRODUCTION

Ionomics is defined as the study of the composition of mineral nutrients and trace elements in organisms and represents the inorganic components in the cells and tissues of an organism (Salt et al., 2008). Early ionomics focused on the identification or characterization of *Arabidopsis* and yeast mutants (Lahner et al., 2003); however, its application has been extended to high-throughput element profiling to find genes involved in controlling the ionome of an organism (Danku et al., 2013).

Ionomics is a powerful approach to quickly analyzing several samples using the inductively coupled plasma mass spectrometry (ICP-MS). In addition, since transcript and ion profiling can be performed not only on cells regulating plant physiology but also on the whole organism simultaneously, many complex studies on ionomes combined with the genome, transcriptome, proteome, and metabolites have recently been published (Baxter, 2010).

Reportedly, As, a toxic non-essential element, accumulates in rice and other crops due to irrigation with contaminated groundwater, industrialization, mining activity, and use of arsenical pesticides (Duxbury et al., 2003; Liao and Ou, 2005; Williams et al., 2007; Chen et al., 2016). As contamination is not limited to water sources, soil contamination with As induces infertile ears and reduces plant growth, especially in agricultural areas (Kumarathilaka et al., 2018a). Arsenic is a major threat to human health because it is absorbed from the soil and accumulates in edible plant parts (Finnegan and Chen, 2012).

In the environment, As exists in organic and inorganic forms, and the contents of inorganic species, such as arsenate [As(V)] and arsenite [As(III)], are higher in soil than the organic arsenic species, such as monomethylarsonic acid (MMAA) and dimethylarsinic acid (DMAA). The ratio of As(III) and As(V) varies depending on the redox state and redox potential (Eh) of soil; if Eh is negative, As(V) is transformed to As(III), and the proportion of As(III) increases (Yamaguchi et al., 2014). Recent studies based on climate change prediction models have shown that arsenic concentrations in the soil are affected by Co_2 and temperature. As(III) concentration decreased with increasing CO_2 concentration, but the toxicity of As(III) to humans increased proportionally with an increase in the temperature compared with that of As(V) (Muehe et al., 2019).

Rice, one of the staple crops consumed by more than 50% of the world's population, is known to be roughly 10 times more likely to assimilate As than other crops (Williams et al., 2007), and therefore, it could be a major source of As toxicity in humans. The mechanism of absorption As(V) is closely related to the phosphate uptake system, whereas As(III) is absorbed by the aquaporins of roots (Ullrich-eberius et al., 1989). It is known that several phosphate transporter genes (e.g., *OsPht1, 1-OsPht1,* and *Pht*) are involved in As(V) absorption, and nodulin 26-like intrinsic proteins (NIPs) have been shown to absorb As(III) (Paszkowski et al., 2002; Bienert et al., 2008; Isayenkov and Matouis, 2008; Ma et al., 2008; Kirk et al., 2009).

The accumulation of ionomes, including arsenic, in plants is regulated by these genes involved in uptake, binding, transport, and sequestration (Baxter, 2009). Therefore, studies to verify the correlation between arsenic and other inorganic components in various populations and to elucidate genetic factors involved in this have been reported, so far (Zhang et al., 2008; Yang et al., 2018). However, the interactions between ionomes are affected not only by genotypes, but also by the environment or the interaction between the environment and the genotypes (Garcia-Oliveira et al., 2009; Hu et al., 2013; Huang and Salt, 2016).

In this study, the interactions between As and other ions were analyzed under non-stress and stress conditions for rice core collections, including *temperate japonica*, *tropical japonica*, *indica,* and *aus*, to elucidate the impact of environmental and genotypic differences. In addition, it was attempted to identify the genetic factors regulating As genes through a transcriptome-wide association study on 23 genes known to transport, detoxify, or stress-response (Yang et al., 2012; Nguyen et al., 2014; Most and Papenbrock, 2015; Yamaji et al., 2015; Hwang et al., 2016, 2017; Shi et al., 2016; Das et al., 2017, 2018a,b; Salt, 2017; Latowski et al., 2018; Sun et al., 2018; Wang et al., 2019; Tiwari et al., 2020; Singh et al., 2021).

MATERIALS AND METHODS

Plant Materials

An allele-mining set of 166 accessions was developed using a heuristic algorithm for 4,046 rice accessions collected from 60 countries held by RDA-Genebank (Kim et al., 2007; Zhao et al., 2010). After the rice core set was constructed, association mapping was conducted on eating quality and amylose content, and a rice core collection for the current 430 accessions was established (Lu and Park, 2012a,b; Zhao et al., 2013). A rice core collection was cultivated in non-contaminated and contaminated paddy soil in 2016 and 2017, and then, 273 accessions overlapped by year were selected. The following subspecies of selected 273 accessions were used: 192 temperate *japonica*, 19 tropical *japonica*, 49 *indica*, 8 *aus*, 3 *admixture*, and 2 *aromatic*.

Field Experiment and Inorganic Component Analysis

Field experiments were conducted on general paddy soil (latitude: 36.670, longitude: 126.855) at Kongju National University and contaminated paddy soil at Hakyeong Mine (latitude: 36.573, longitude: 126.819), Yesan-gun, Chungcheongnam-do, Republic of Korea. The 273 rice accessions were grown in contaminated soil over two years (2016 and 2017), and the ears of each individual were sampled at the yellow ripe stage to analyze accumulation patterns of ionomes in brown rice (**Supplementary File 1**). In particular, field experiments were conducted under flooded (2016) and intermittently flooded conditions (2017) due to the narrow arable land of the contaminated soil.

The chemical properties of soil were analyzed following the National Academy Aggregation Science (NAAS, 2010). First, soil samples were mixed with distilled water at a 1:5 ratio and kept for 1 h. The pH and EC were then measured (Orion 3 Star, Thermo, USA). To measure cation exchangeability for Ca, Mg, K, and Na in soil, soil samples were mixed with 1 mL of distilled water, 21 mL of HCl, and 7 mL of HNO_3. Subsequently, the mixture

was decomposed using Kjeldahl (C. Gerhardt GmbH & Co., Northants, UK). Afterward, 1 M NH_4OAc (pH 7.0) was added to the decomposed mixture, mixed by shaking, and filtered using a Whatman No. 42 filter paper (Kang et al., 2018).

After the 273 brown rice accessions were ground using a cyclone miller (PX-MFC 90 D, KINEMATICA, Switzerland), they were decomposed by adding 4 mL of HNO_3 and 1 mL of distilled water in a microwave oven (UltraWAVE, Milestone, USA). The volume of the decomposed samples was adjusted to 25 mL with distilled water. Micro-elements in plants (As, Zn, Fe, Cu, Mn, Na, and Se) were analyzed using ICP-MS (7700E, Agilent Technologies, USA), and the macro-elements in soil (Ca and Mg) were analyzed using inductively coupled plasma-optical emission spectrometry (Integra XL, GBC, AUS). The R^2 of the standard calibration curves for each element was 0.999 or more using multi-element standard (Agilent, USA), and it was confirmed that the recovery rates for inorganic components were 80–120% using IRMM-804 rice flour.

Meanwhile, broad-sense heritability for As content was calculated by QTLmax Global (2022) with the following formula: $H^2 = \frac{\sigma_g^2}{\sigma_g^2+\sigma_e^2}$, where "$\sigma_g^2$" is the genetic variance, "σ_e^2" is the genotype by the annual environmental effect, "σ^2" is the error variance, and "e" is the annual environmental effect (Le Sech and Christian, 1991).

DNA Extraction and Whole-Genome Sequencing

DNA and RNA were extracted from the 15-day-old seeds after heading to analyze the SNP and arsenic gene expression levels. Some amino acids or metabolites play important roles in plant growth and development, as well as plant resistance to various stresses (Li et al., 2010). Amino acid metabolism is affected by high night temperature at the early milky stage. As a result, the seed weight and grain quality of rice are determined by changes in gene expression patterns (Liao et al., 2015).

The 15-day-old young seeds were collected and powdered using mortar–pestle. Afterward, genomic DNA was extracted using DNeasy® Plant Mini Kit (QIAGEN, Germany). The concentration of the sample was adjusted to 30 ng μL-1. For next-generation sequencing library preparation, quality control (QC) analysis was performed to ensure that the fragment of the DNA was of the desired size. The extracted DNA samples were quantified using the Quant-iT™ dsDNA High-Sensitivity Assay Kit (Invitrogen, Carlsbad, CA, USA) on an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Optical density was measured using Tecan F200 (Tecan, Switzerland), and the quality of the extracted DNA was confirmed by electrophoresis on a 0.7% agarose gel. Short-read sequences were obtained using HiSeq 2500 (Illumina), and next-generation sequencing was performed for genome analysis (**Supplementary File 2**).

RNA Sequencing and Transcriptome Analyzes

Total RNA from the samples was extracted using the Total RNA Prep Kit for plant tissues (QIAGEN, Germany). The quality of the extracted RNA was confirmed by electrophoresis on 0.7% agarose gel, and its absorbance was measured using a UV spectrophotometer (UV-2600, SHIMADAZU). The purity and quantification were performed using NanoDrop ND-1000 (Dupont Agricultural Genomics Laboratory). The concentration of the sample was adjusted to 20 ng μL^{-1}. The short-read sequence obtained from RNA sequencing was aligned, and the Bowtie (version 1.1.2) and Tophat (version 2.1.0) were used to compare and map with International Rice Genome Sequencing Project 1.0 (Heo et al., 2017; **Supplementary File 3**).

Model Selection for GWAS

GWAS is the most powerful statistical tool to analyze the association between traits and SNP markers, and various genetic models are applied to identify quantitative traits. Whole-genome resequencing data were imputed using the Beagle (Browning and Browning, 2007), and then, 3,110,974 SNPs for the rice core collection (273 accessions) were obtained from 808,686 SNPs by adjusting the minor allele frequency (MAF) to be less than 5% and removing the proportion of missing SNPs by 80%. In addition, accurate genome-scale profiling for As was performed by applying the linear mixed model (LMM).

The formula of the LMM is as follows:

$$y = X\beta + Zu + \varepsilon, \quad (1)$$

where "y" is the observed phenotype, "β" indicates the marker information that is a fixed effect, "u" indicates the object information that is a random effect, "ε" is the random residual effect, and "X" and "Z" are the associated design metrics. The random effect assumes that u is proportional to the normal distribution for the mean and covariates [$u \sim N(0, G)$] and that ε is proportional to the normal distribution for the means and sum of the error squares [$\varepsilon \sim N(0, I\sigma_\varepsilon^2)$] (Piepho et al., 2008).

On the contrary, in the 2016 and 2017 GWAS results applying GLM, the lambda values were 1.234507 and 1.58868, respectively, which were inflated compared with the LMM (0.781438 and 0.950722) (**Supplementary File 4**).

Significant QTLs with $-\log_{10}(p) > 5$ from GWAS results applied with LMM were selected, and linkage disequilibrium was analyzed by calculating allele frequency as r^2 for each QTL. Candidate genes were identified at the range of ±50 kbp using genome browser of the Rice Annotation Project (https://rapdb.dna.affrc.go.jp/) (Kawahara et al., 2013; Sakai et al., 2013) for QTLs with a low recombination rate ($r^2 \leq 1$) (**Supplementary File 5**). In addition, functional haplotypes for candidate genes were investigated by using genomic sequence information, including the promoter and coding regions. The phenotypic variation for the haplotype group was statistically validated with a one-way ANOVA test.

Statistical and Network Analyses

Jamovi (version 1.6.12) was used for the correlation analysis between inorganic components. The one-way analysis of variance (ANOVA) was performed for genotypic differences (*japonica*, *indica*, and *aus*) of the rice core collection. Cytoscape (version 3.7.2) was used for the network analyses between inorganic

components based on the Pearson correlation coefficient. However, the network analysis excluded *Admixture* and *Aromatic* collections with small sample numbers ($n < 5$).

RESULTS

Variations in the Ionome Among the 273 Rice Accessions

The variations of As, Se, Na, Ca, Mn, Fe, Cu, Zn, and Mg contents accumulated in brown rice of 273 accessions were analyzed. Under the non-contaminated soil condition, As contents of rice varieties ranged from 0.0407 to 0.1775 mg kg^{-1}, with an average content of 0.0833 mg kg^{-1}. In 2016, As contents of varieties grown under contaminated soil conditions ranged from 0.1306 to 0.6923 mg kg^{-1}, with an average content of 0.2768 mg kg^{-1}. This average content increased by approximately 332% compared with the non-contaminated soil condition. The As content of the cultivars in 2017 was grown under the same polluted soil conditions as in 2016, As contents ranged from 0.0391 to 0.4184 mg kg^{-1}, with an average of 0.1471 mg kg^{-1}, a decrease of approximately 53% compared with the previous year. Functional inorganic components, such as Na, Ca, Mn, Fe, Cu, Zn, and Mg, showed statistically significant differences ($p < 0.05$) in all environmental conditions. It was confirmed that under the contaminated soil condition in 2017, other inorganic components except for Na, Cu, Fe, and Mg increased compared with the contaminated soil condition of the previous year (**Figure 1**). However, there were no significant differences in Se contents under the contaminated soil conditions in both years. In the 273 rice accessions grown in contaminated soil, the broad-sense heritability (H^2) of As content in 2016 and 2017 was about 54%, which was found to be affected by both environmental and genetic factors (**Figure 2**).

Interaction of Ionomes in Rice Core Collection

It was confirmed that the accumulation of ionomes in rice is affected by environmental conditions, and the contents of ionomes vary greatly depending on the rice varieties. This is predicted to be due to the genotypic differences due to subspecies, and the ionome content and their interactions were analyzed by classifying them into *temperate japonica*, *tropical japonica*, *indica*, and *aus*. Compared with other subspecies, *temperate japonica* exhibited more complex interactions between ionomes under both non-contaminated and contaminated soil conditions (**Figures 3A,B**). On the contrary, the correlations between ionomes in the rice grown under contaminated soil conditions were increased in all subspecies in the second year (2017) compared with that in the previous year (**Figures 3B,C**). The direction of correlation between As and other functional minerals varied depending on subspecies and environmental conditions.

Co-Expression Network Analysis

The subspecies of rice were largely classified into *japonica* and *indica* population; then, the relationship between expression levels and genotypes of 23 genes known to be involved in As transport, detoxification, or As-induced stress response was analyzed for each population (**Table 1**). The expression networks for 23 As genes were performed using RNA-seq data from the rice accessions grown in unpolluted paddy soils.

The network average clustering coefficient for 23 As genes expression levels was 0.428 in the *japonica* population, which was larger than that in the *indica* population (0.414). That is, it was confirmed that the expression network of the *indica* population was denser than that of *japonica*, so seven essential genes (STR5, STR6, STR8, MYB1, multidrug and toxic compound extrusion 2 [MATE2], NIP1.1, and NIP3.2) were identified in the *japonica* population, and these genes are related to sulfur transferase and As transporter. Seven essential genes (STR5, STR8, GRX4, MRP1, ACR2.1, RCS, and AIR2) involved in sulfur transferase, As translation, and As induction were identified in the *indica* population (**Figure 4**). As a result of comparing As contents accumulated in grains of rice core collection cultivated in contaminated soil based on network analyses for rice subspecies, the As content of *japonica* under the flooded condition was 0.283 mg kg^{-1}, which was significantly higher than that of *indica* (0.253 mg kg^{-1}). In addition, the arsenic content of *japonica* was 0.153 even under the intermittently flooded condition, which was significantly higher than that of *indica* (0.126 mg kg^{-1}).

Expression QTL Analyzes

Twelve genes were identified as essential genes in the expression network for As genes. To identify the genetic factors regulating the expression of these genes, the association between the whole genome and the gene expression level was analyzed for the 273 rice accessions. As a result, the associations between expression levels and genotypes for STR5, STR8, and AIR2 were significant as $-\log_{10}(p) > 5$ (**Figure 5**). The expression levels of STR5 were associated with *trans*-eQTLs ranging from 25.4 to 33.6 Mb on chromosome 1, and the $-\log_{10}(p)$ for *trans*-eQTLs ranged from 5.03 to 7.21 (**Figure 5A**). The 18,967 SNPs were identified at *trans*-eQTLs of STR5. These SNPs were clustered into three groups by variations, and 90% *japonica* varieties and 97% *indica* varieties were included in Group 1 and Group 3, respectively. STR5 expression levels of each group were 0.23, 0.35, and 0.30, but there was no statistical significance (**Table 2**). A total of 121 *trans*-eQTL genes were identified at the *trans*-eQTLs, including 5 at 25 Mb and 31 at 33 Mb. Os01g0635400 (protein-binding) and transcription-regulating Os01g0635550 (ZF-HD homeobox) were detected at 25 Mb, and genes involved in nodulin 20 (Os01g0786500), ABC transporter (Os01g0786000), peroxidase (Os01g0787000), and transcription factor (Os01g0788800) were detected at 33 Mb (**Supplementary File 6**).

The expression levels of STR8 and AIR2 showed high association at *cis*-eQTLs on chromosome 2 and chromosome 11, respectively. The *cis*-eQTLs for STR8 were identified at 3.0 to 3.1 Mb and 19.4 to 20.0 Mb on chromosome 2. The $-\log_{10}(p)$ for *cis*-eQTLs ranged from 5.1 to 8.9 (**Figure 5B**). The 226 SNPs at 3 Mb were identified and clustered into three groups by SNPs variations. All *japonica* and 63% *indica* varieties were included in Group 1, and 26% *indica* varieties were included in Group 3. The expression levels of STR8 in each group were 0.22, 0.33, and 0.21, but there

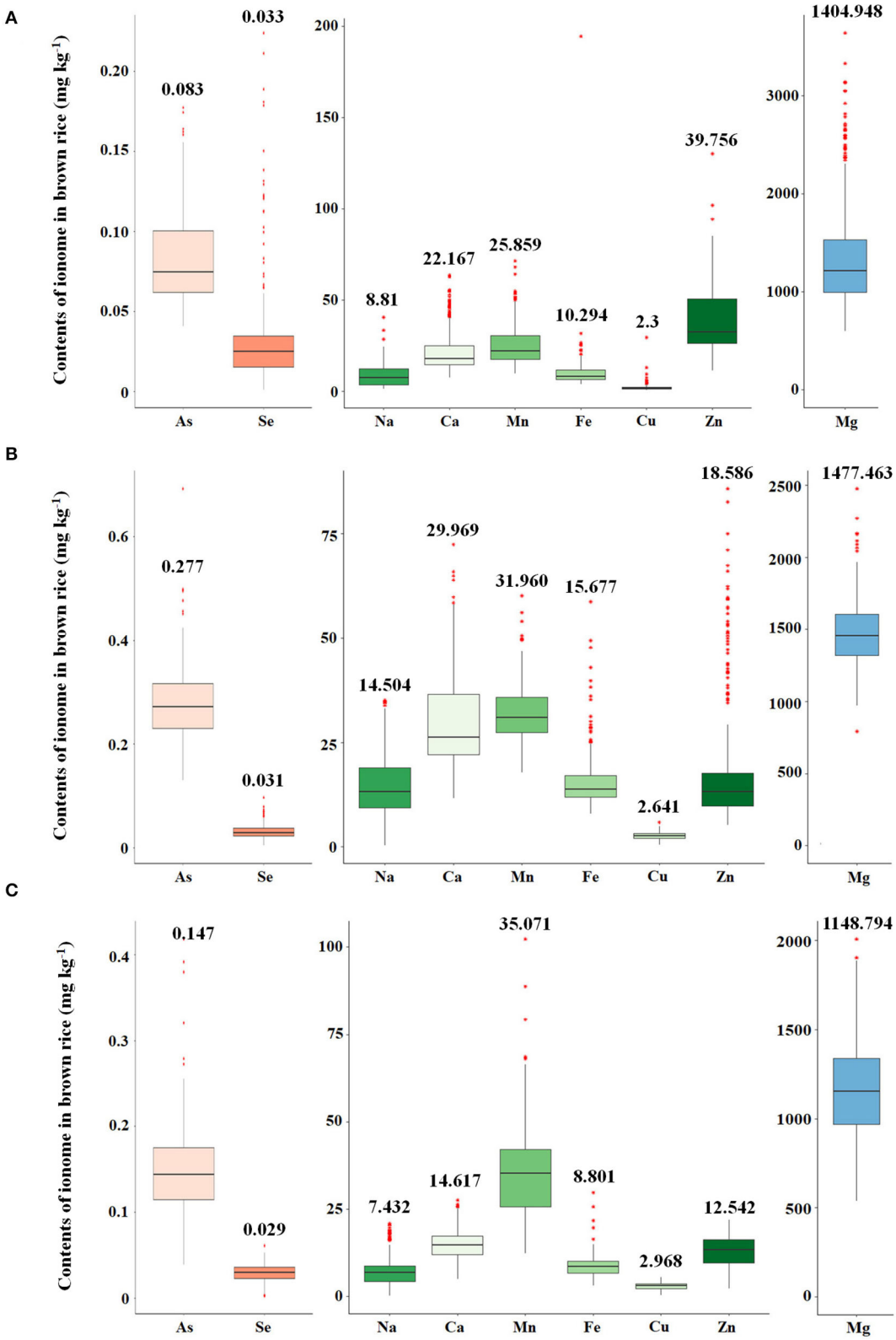


FIGURE 1 | Variation of arsenic and functional ion content for the 273 rice accessions. **(A)** Ion contents in 273 rice grains in non-contaminated soil. **(B)** Ion contents in 273 rice grains under the flooded condition in contaminated soil. **(C)** Ion contents of 273 rice grains under the intermittently flooded condition in contaminated soil.

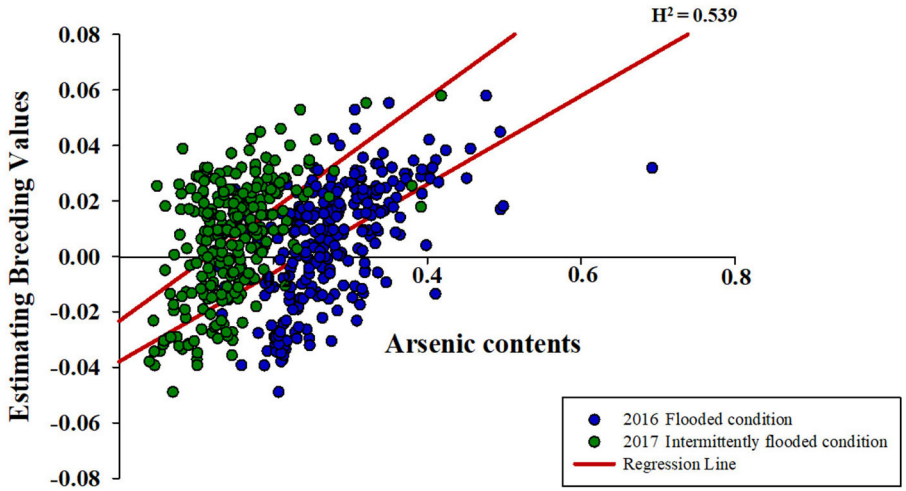


FIGURE 2 | Correlations between estimating breeding values (EBVs) and arsenic contents. Green circle indicates arsenic contents of the 273 rice accessions grown under the flooded condition in contaminated soil. Blue circle indicates arsenic contents of the 273 rice accessions grown under the intermittently flooded condition in contaminated soil.

was no significant difference between groups. The 376 SNPs were identified and clustered into two groups at *cis*-eQTLs within 19.4–20 Mb. All *japonica* cultivars and 72% *indica* varieties were included in Groups 1 and 2, respectively, but there was no association between STR8 expression and SNPs variations at *cis*-eQTLs (**Table 2**). The 20 *cis*-eQTL genes were detected at *cis*-eQTLs of STR8. The *cis*-eQTLs genes involved in leucine-rich repeat domain (e.g., Os02g0156400) and arsenate reductase (Os02g0157600) were identified at 3 Mb. In addition, five cis-eQTLs genes, including chloroplast development protein (Os02g0539600), were identified at 19.2–20 Mb (**Supplementary File 6**).

The 15,388 SNPs were identified at 19.2 Mb−25.8 Mb associated with AIR2 expression, and $-\log_{10}(p)$ for *cis*-eQTLs ranged from 5 to 8.84 (**Figure 5C**). The *cis*-eQTLs were clustered into three groups, and 75% *japonica* and all *indica* varieties were included in Groups 1 and 2, respectively. AIR2 expression levels were 1.89, 1.85, and 1.67 in each group, and there was a significant difference ($p < 0.001$) in Group 1 and Group 3 (**Table 2**). The 591 *cis*-eQTL genes for AIR2 were detected, and it was confirmed that a number of genes are related to leucine-rich repeat domains (e.g., Os11g0568800), serine/threonine kinase activity (Os11g0569300), heat shock (Os11g0578500), and RNA binding (Os11g0579900). In addition, the AIR2 was detected as a *cis*-eQTLs gene (**Supplementary File 6**).

QTLs Identification for As Content Through GWAS

To identify additional As-related genes other than known major genes and to investigate genetic factors associated with arsenic in plants by varying growing conditions under the same stress environment, we performed GWAS of the rice core collections, grown under contaminated soil conditions in 2016 and 2017. The contaminated soil was maintained under the flooded condition in 2016 and was maintained under the intermittently flooded condition in 2017. Under the flooded condition, As-associated QTLs (As-QTLs) were identified at 8.0 and 8.32 Mb on chromosome 6 and *p*-values were significant as 5.12 and 5.36, respectively (**Figure 6A**). As a result of LD (linkage disequilibrium) for As-QTLs, it was confirmed that the linkage equilibrium was weakly linked ($R^2 < 0.5$) within 8.32 Mb and strongly linked ($R^2 \leq 1$) between arsenic-related SNPs within 8.0 Mb (**Supplementary File 5**). Ten candidate genes were identified within ±50 kb based on 8.0 Mb, and a number of candidate genes were identified that respond to abiotic stress (**Figure 6C**). In particular, Os06g0254200 for ion channel activity and potassium transport, Os06g0254300 for calcium ion binding, and Os06g0255100, an iron-dependent oxygenase protein, were detected as As-related candidate genes (**Supplementary File 7**).

Under the intermittently flooded condition, As-QTLs were identified from 30.7 to 30.8 Mb and 9.6 to 14.2 Mb and 27.1 Mb on chromosomes 1 and 5, respectively (**Figure 6B**). The *p*-values of As-QTLs ranged from 5 to 6.7, and linkage equilibrium was strongly linked with $R^2 \leq 1$ (**Supplementary File 5**). The 150 candidate genes on chromosomes 1 and 5 were identified within ±50 kb based on each As-QTL. Candidate genes involved in biosynthetic process, membrane, and plastid were detected under the intermittently flooded condition, unlike candidate genes identified under the flooded condition (**Figure 6D**). In addition, Os01g0733001 (NRAMP3), which transports metal ions, and Os05g0279400 (RFP), which regulates innate immunity and disease resistance, were detected at As-QTLs (**Supplementary File 7**). GWAS results for other functional ions are not presented in the main text and are given in **Supplementary File 8**.

Selection of Low-Arsenic Rice Varieties

AIR2 showed a significant difference in expression between *japonica* and *indica* at the *cis*-eQTL ($p < 0.001$), so it was

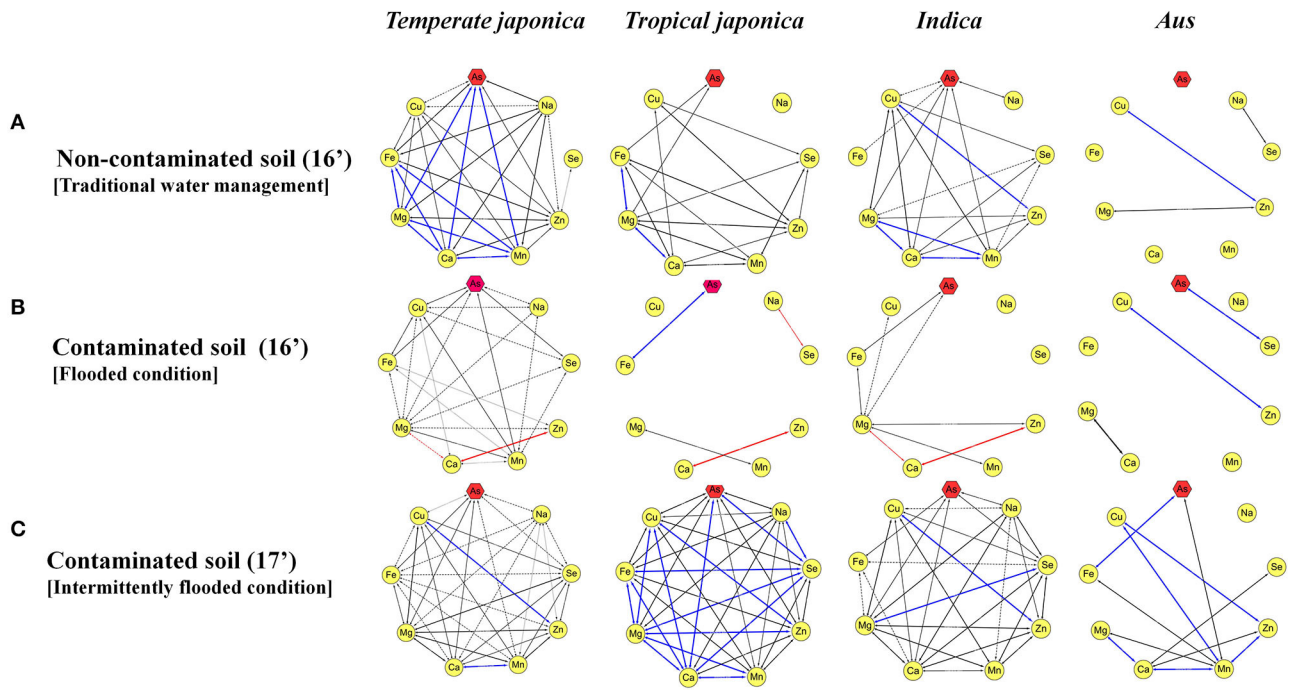


FIGURE 3 | Interaction between arsenic and other inorganic components. The correlation between ionomes was based on the *Pearson coefficient value*. Blue line: positive correlation ($0.8 < r < 1.0$), red line: negative correlation ($-1.0 < r < -0.8$), black line: positive correlation ($0.4 < r < 0.8$), and dotted line: positive correlation ($0.2 < r < 0.4$). **(A)** The interaction between ionomes of 273 rice grains in non-contaminated soil. **(B)** The interaction between ionomes of 273 rice grains under the flooded condition in contaminated soil. **(C)** The interaction between ionomes of 273 rice grains under the intermittently flooded condition in contaminated soil.

verified whether the arsenic content differs by rice subspecies in genomic sequences of AIR2. The haplotype group for AIR2 was classified into three groups, and SNP variations between groups were confirmed in nonsynonymous SNPs ($Ala^{134} \rightarrow Thr^{134}$, $Gly^{130} \rightarrow Ser^{130}$, $Ile^{101} \rightarrow Thr^{101}$). The 104 *japonica* varieties belonged to Hap1, and the average As content under flooded and intermittently flooded conditions was 0.295 and 0.161 mg kg^{-1}, respectively. Hap2 included 54 *japonica* and 18 *indica* varieties, and the average As content under flooded and intermittently flooded conditions was 0.261 and 0.143 mg kg^{-1}, respectively. Hap3 included 19 *japonica* and 25 *indica* varieties, and the average As content under flooded and intermittently flooded conditions was 0.251 and 0.134 mg kg^{-1}, respectively, which were significantly lower ($p < 0.01$) than those of Hap1 and Hap2.

The haplotype groups for AIR2 were similar to the groups clustered in the *cis*-eQTLs of AIR2. Based on the results of these analyses, rice varieties with low expression levels and low As content were selected from *japonica* and *indica*, respectively (**Figure 7**).

DISCUSSION

Variation and Interaction of Ionomes in Rice Core Collection

Partitioning the core collection based on rice subspecies and estimation of the interaction between As and other functional minerals demonstrated that environmental and genetic factors contribute to these interactions. The ionome accumulation of 273 rice grains cultivated in contaminated paddy soil was generally increased compared with unpolluted paddy soil conditions in the same year. Na, Mn, Fe, Cu, and Mg including arsenic increased, while Se, Ca, and Zn decreased in the contaminated soil (flooded condition). Nicula et al. (2013) investigated whether the properties of metals and the particularities of each species affect the uptake capacity in *Phaseolus vulgaris* and *Zea mays* and found that Fe, Cu, and Zn content of the two species was increased in contaminated soil compared with unpolluted soil. In addition, Williams et al. (2009) confirmed that Zn and Se significantly decreased in rice as the arsenic concentration increased. The results of these previous studies are consistent with the variations in ionome contents for the 273 rice accessions.

The correlations between ionomes were decreased in contaminated soil (flooded condition) compared with unpolluted soil, and it was confirmed that there was a difference by rice subspecies. The findings also revealed that the correlations between ionomes under the intermittently flooded condition were increased compared with those under the flooded condition. Therefore, it can be inferred that the contents and interaction between As and functional ions vary depending on the rice subspecies and water management. Although the genetic basis for the correlation between trace elements or toxic elements has not yet been established (Tan et al., 2020), studies have demonstrated the influence of subspecies on As, Hg, Pb, and Cd

TABLE 1 | Genes involved in As transport, detoxification, or stress response in rice.

Gene names	Gene symbols	Descriptions
OS02G0745100	LSI1	Aquaporin NIP III subfamily protein, arsenic species (As) uptake, arsenite efflux (Os02t0745100-01)
OS01G0368900	OsGRX4	Glutaredoxin (Grx) family protein, arsenic (As) stress response, drought tolerance (Os01t0368900-01)
OS02G0102300	OsHAC1;1	Arsenate reductase, regulation of arsenic accumulation (Os02t0102300-01)
OS02G0157600	STR8	Arsenate [As(V)] reductase, As(V) tolerance, control of arsenic (As) accumulation (Os02t0157600-01)
OS02G0618100	GRX9	Glutaredoxin (Grx) family protein, arsenic (As) stress response, drought tolerance (Os02t0618100-01)
OS03G0107300	LSI2	Anion transporter, silicon efflux transporter, arsenic species (As) uptake (Os03t0107300-01)
OS03G0108000	ACR2.2	Dual-specificity tyrosine phosphatase CDC25, arsenic metabolism (Os03t0108000-01)
OS03G0195800	OsSultr1;1	Similar to sulfate transporter (fragment) (Os03t0195800-01)
OS03G0747800	OsRCS3	O-Acetylserine(thiol) lyase, cysteine biosynthesis, arsenic detoxification (Os03t0747800-01)
OS04G0249600	OsHAC1;2	Arsenate reductase, sulfurtransferase/rhodanese-like protein, regulation of arsenic accumulation (Os04t0249600-02)
OS04G0620000	OsABCC1	Arsenic (As) detoxification, reduction of As in grains (Os04t0620000-01)
OS05G0497600	OsAIR1	Arsenic-induced RING E3 ligase, abiotic stress response (Os05t0497600-01)
OS05G0554000	MATE2	Arsenic stress response, regulation of plant growth and development, disease resistance (Os05t0554000-03)
OS06G0102300	OsPCS	Phytochelatin synthase 2, cadmium (Cd), and arsenic (As) tolerance (Os06t0102300-03)
OS10G0545700	ACR2.1	Dual-specificity tyrosine phosphatase CDC25, arsenic metabolism (Os10t0545700-04)
OS11G0572500	OsAIR2	Similar to zinc finger, RING type (Os11t0572500-01)
OS02G0232900	OsNIP1;1	A member of the nodulin26-like intrinsic protein (NIP) family, arsenite transporter (Os02t0232900-01)
OS02G0822100	SIET3	Similar to arsenite transport subunit B (Os02t0822100-02)
OS05G0442400	MYB-1	R-R-type MYB-like transcription factor (Os05t0442400-01)
OS08G0152000	NIP 3;2	Nodulin 26-like intrinsic membrane protein, arsenite [As(III)] uptake by lateral roots (Os08t0152000-01)
OS08G0152100	NIP 3;3	A member of the nodulin26-like intrinsic protein (NIP) family, arsenite transporter (Os08t0152100-01)
OS09G0521500	Get3	Similar to arsenical pump-driving ATPase (EC 3.6.3.16) (Os09t0521500-01)
OS01G0955700	OsCLT1	CRT-like transporter, glutathione homeostasis, arsenic tolerance (Os01t0955700-01)

contents in *japonica* and *indica* rice (Meharg et al., 2009; Jiang et al., 2012; Norton et al., 2012). In addition, Wang et al. (2020) reported that the variation in accumulation and translocation of As among 74 main rice cultivars, including 66 *japonica* and 8 *indica* cultivars, was influenced by Si, P, Fe, and Mn contents.

Water management is an effective way to reduce the absorption of toxic ions, such as As, in crops. Rahaman and Ashim (2013) reported that arsenic content by approximately 24% is reduced and Fe and Zn are also reduced to lower levels under the intermittently flooded condition. Although there are environmental factors for the annual As content of 273 rice varieties, it is consistent with previous studies in that As content was decreased by 46.8% under the intermittently flooded condition compared with the flooded condition, and Na, Fe, and Zn also were decreased by 43.8 and 32.5%, respectively. However, the intermittently flooded condition induces oxidative stress and facilitates Cd accumulation, particularly in crops, so the rice core collections were classified into subspecies, and eQTL analyses were performed on 23 genes known as As induction or resistance to identify the genetic factors.

Co-Expression of As-Related Genes

STR5, STR6, STR8, MYB1, MATE, NIP1.1, and NIP3.2 were identified as essential proteins in the expression networks of 23 As genes of the *japonica* population, and the clustering coefficient of these genes was 0.5, which was larger than that of other genes. The clustering coefficient of GET3 was the smallest at 0.23. STR5, STR8, GRX4, MRP1, ACR2.1, RCS, and AIR2 were identified as essential genes with clustering coefficients ranging from 0.45 to 0.5 in the expression network of the *indica* population. The clustering coefficients of STR6 and MATE2 were the smallest at 0.23. The average clustering coefficient of *japonica* population was 4.28, which was larger than that of *indica* population, but the shortest paths accounted for 30% of the total path, which was less than that of *indica*. Here, the clustering coefficient describes the relatedness of two proteins and measures the importance of edges in a protein–protein interaction network, thus facilitating identification of key proteins in populations (Wang et al., 2012). That is, although many essential arsenic genes are present in the *japonica* population, the distance between As-induced and As-resistant genes is longer than that of *indica*, so it is likely to be easily exposed to the risk of As. This hypothesis was indirectly verified through the results of the As content of *japonica* and *indica* in the rice core collections grown in contaminated soil. On the contrary, the expression levels of STR5, STR8, and AIR2 among the 14 essential genes were $-\log_{10}(p) > 5$, which had high associations with the SNPs of core collections. Reportedly, sulfur assimilation and GSH are upregulated in response to oxidative stress in plants exposed to heavy metals (Na and Salt, 2011). Sulfur assimilation plays an important role in minimizing As absorption and transfers to crops and arsenic detoxification (Dixit et al., 2015). *STR5, STR6,* and *STR8* are genes associated with sulfur assimilation and are known to regulate the As accumulation (Most and Papenbrock, 2015).

AIR2 (As-induced RING E3 ligase 2) encodes RING E3 ubiquitin ligase. Heterogeneous overexpression of OsAIR2 has been reported to positively regulate a plant growth in Arabidopsis or rice in response to As(V) stress. In addition, OsKAT1

TABLE 2 | eQTLs associated with STR5, STR8, and AIR2 expression levels.

Arsenic gene	Chromosome	Expression quantitative traits loci (eQTLs)	Haplotype group	Subspecies (number of varieties)	Average expression levels	Differences in gene expression between haplotype groups
STR5	1	*trans*-eQTLs 25.4–33.6 Mb	Group 1	*japonica* (160) *indica* (0) Others (0)	0.23[a]	
			Group 2	*japonica* (17) *indica* (1) Others (3)	0.35[a]	*F*-value 2.06 *p*-value 0.138
			Group 3	*japonica* (0) *indica* (42) Others (10)	0.30[a]	
STR8	2	*cis-eQTLs* 3.0–3.1 Mb	Group 1	*japonica* (177) *indica* (27) Others (3)	0.22[a]	
			Group 2	*japonica* (0) *indica* (5) Others (0)	0.33[a]	*F*-value 1.63 *p*-value 0.220
			Group 3	*japonica* (0) *indica* (11) Others (10)	0.21[a]	
AIR2	11	*cis-eQTLs* 19.2–25.8 Mb	Group1	*japonica* (132) *indica* (0) Others (3)	1.89[a]	
			Group 2	*japonica* (41) *indica* (0) Others (0)	1.85[a]	*F*-value 13.5 $p < 0.001$
			Group 3	*japonica* (4) *indica* (43) Others (10)	1.67[b]	

a and b indicates whether there is a statistically significant difference between groups.

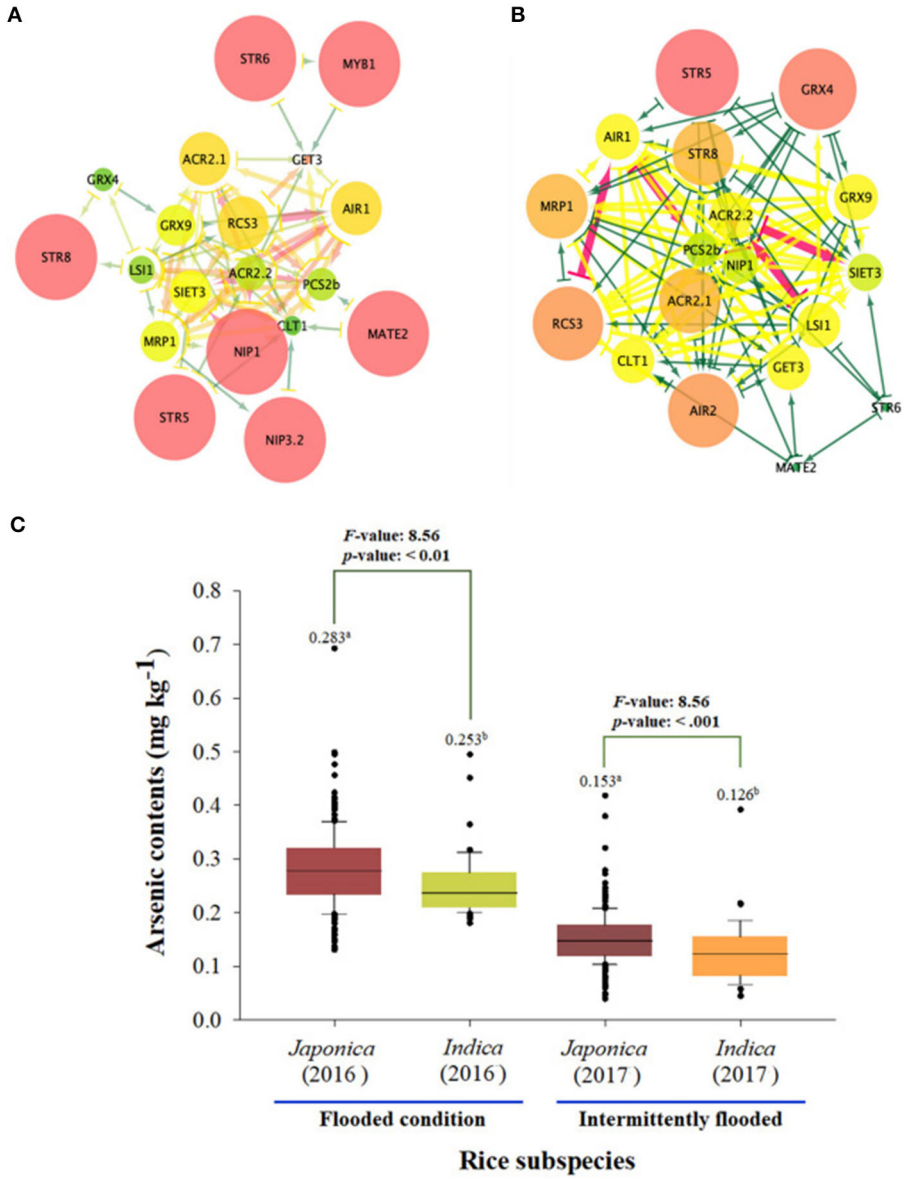


FIGURE 4 | Expression network analysis of As genes between *japonica* and *indica* population. **(A)** The expression network of As-related genes within the *japonica* population. **(B)** The expression network of As-related genes within the *indica* population. **(C)** As content of *japonica* and *indica* under flooded and intermittently flooded conditions. a and b indicates whether there is a statistically significant difference between groups.

(3-ketoacyl-CoA thiolase protein), the physical interaction partner of OsAIR2, is degraded and ubiquitinated by OsAIR2 through the 26S proteasome degradation pathway (Lim et al., 2014; Hwang et al., 2016, 2017). The eQTLs of STR5 and STR8 were not directly related to the gene expression, but the expression level of AIR2 showed a significant difference between groups at the *cis*-eQTLs. In particular, these *cis*-eQTLs include AIR2 genomic sequences, and homogeneous and heterogeneous SNPs were observed in Groups 1 (75% *japonica*) and 3 (100% *indica*), respectively. These findings suggest that SNP variations at the *cis*-eQTL associated with OsAIR2 regulate resistance to As stress, and it can be inferred that *indica* is less likely to accumulate As than *japonica*.

A number of *trans* and *cis*-eQTL genes were identified, such as nodulin 20, ABC transporters, leucine-rich repeat domains, and transcription factors at the eQTLs associated with the expression of STR5, STR8, and AIR2. Nodulin 20 (EN20) is a signaling mediator that activates plant defenses under stress, and OsABCC1, a type C ATP-binding cassette transporter (OsABCC) family, is known to detoxify As in grains (Wu et al., 2011; Song et al., 2014). In addition, the leucine-rich repeat domain (LRR) is involved in negative regulator-programmed cell death, tolerance

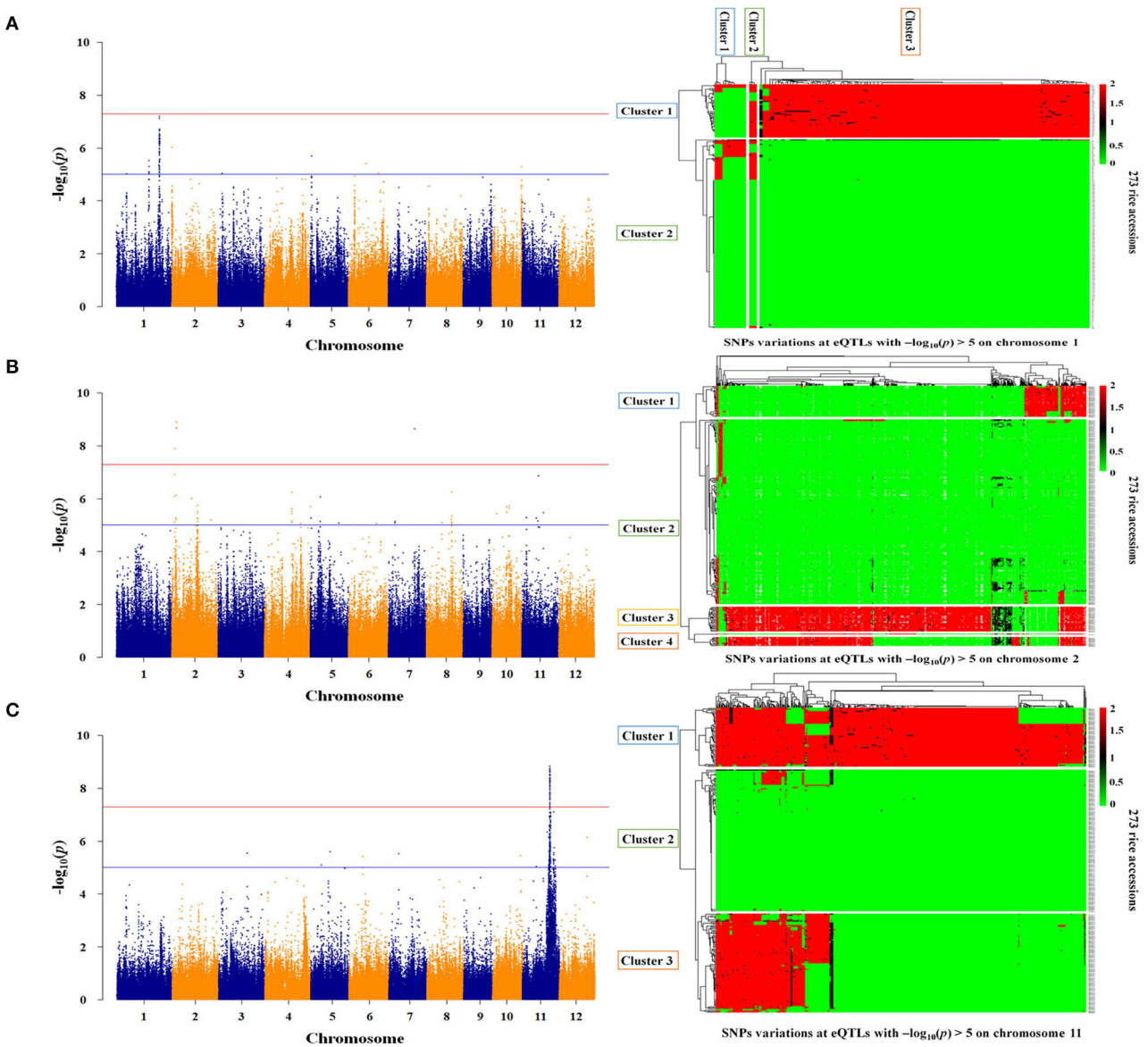


FIGURE 5 | Regulatory factors associated with As gene expression. **(A)** Genome-scale profiling for STR5 expression levels, and SNPs variations at *trans*-eQTLs (25.4–33.6 Mb) in the *japonica and indica* populations. **(B)** Genome-scale profiling for STR8 expression levels and SNPs variations at *cis*-eQTLs (3.0–3.1 Mb and 19.4–20.0 Mb) in the *japonica and indica* populations. **(C)** Genome-scale profiling for AIR2 expression levels and SNPs variations at *cis*-eQTLs (19.2 Mb−25.8 Mb) in the *japonica and indica* populations. The major allele was indicated as 0 (green), the minor allele as 2 (red), and the heterologous allele as 1 (black).

to oxidative stress, and salt stress, but the biological function of most of the LRR gene has not been clearly elucidated in the plant genome (De Lorenzo et al., 2009; Oh et al., 2010; Pitorre et al., 2010; Hwang et al., 2011).

Identification of QTLs and Candidate Genes for as Content

In general, GWAS is the most powerful statistical method that provides a genetic basis for complex traits and has been applied to major agricultural traits, including toxic minerals, yield, flowering time, and disease resistance, among others, in rice (Huang et al., 2010, 2012; Famoso et al., 2011; McCouch et al., 2016). In rice, the alleles with large effects have been retained and fixed by evolution, human selection, and inbreeding, and therefore, GWAS is widely used to investigate these complex alleles in rice to uncover the genetic basis and effects associated with various traits.

In this study, GWAS was performed by applying LMM to arsenic content of the rice core collection under different water management conditions in the contaminated soil to uncover the genetic factors associated with As. GWAS results applied with

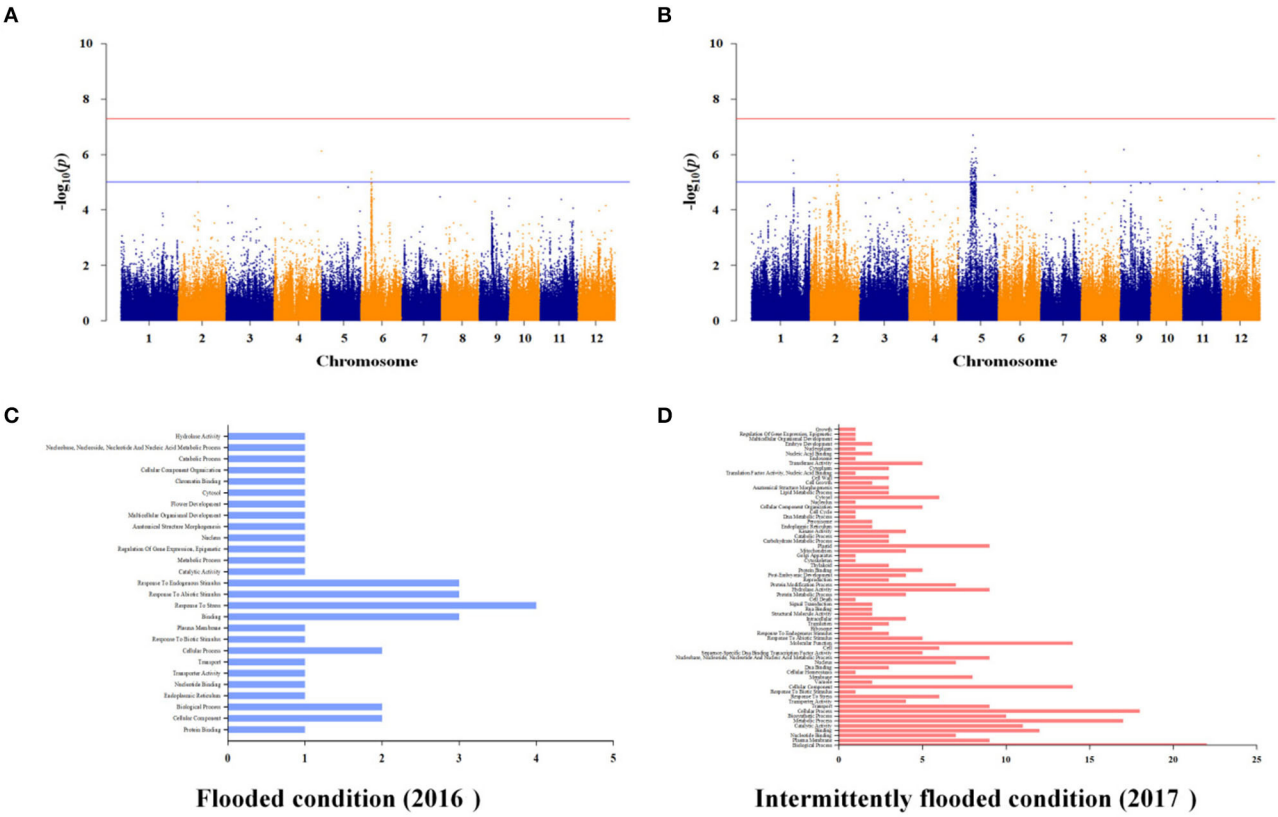


FIGURE 6 | Whole genome-scale profiling for As under flooded and intermittently flooded conditions in contaminated soil. **(A)** Significant SNPs with -$\log_{10}(p)$ > 5 were identified on chromosome 6 under the flooded condition. **(B)** Significant SNPs with -$\log_{10}(p)$ > 5 identified on chromosomes 1 and 5 under the intermittently flooded condition. **(C)** Ten candidate genes under the flooded condition were identified within strongly linked SNPs ($R^2 \leq 1$), and a number of candidate genes were involved in abiotic stress response. **(D)** The 150 candidate genes under the intermittently flooded condition were identified within strongly linked SNPs ($R^2 \leq 1$), and a number of candidate genes were involved in biosynthetic process, membrane, and plastid.

LMM were verified by comparison with a general linear model (GLM). A lambda value ranging from 0 to 1 is a measure of asymmetric association indicating the strength of the relationship between the independent and dependent variables and is used to validate the *p*-value from the GWAS results (Hartwig, 1973; Kim et al., 2019).

As-associated SNPs were mapped on chromosomes 6 and 5 under flooded conditions and intermittently flooded conditions, respectively. The association of arsenic with the genome in crops may vary depending on growing conditions. Norton et al. (2019) reported various QTLs associated with arsenic through annual comparisons by water management systems. Moreover, it is known that microorganisms directly or indirectly affect the mobility and biological availability of arsenic in rice paddies, depending on the water management system; then, arsenic uptake and translocation can be decreased or increased by altered expression of transporter genes in plants (Kumarathilaka et al., 2018b). Concordant with the findings of previous studies, this study confirmed that As-related QTLs showed different patterns under water management conditions and that there were several candidate genes involved in physiological responses under intermittently flooded conditions. In a follow-up study, it is necessary not only to verify these candidate genes based on the GWAS results for As, but also to further prove the genetic relationship between arsenic and functional ions by exploring SNPs co-localized with other functional ions.

Selection of Low-Arsenic Rice Varieties

The 273 rice accessions were classified according to rice subspecies in genomic sequences and *cis*-eQTLs of AIR2. The groups clustered in *cis*-eQTLs mostly coincided with the haplotype groups of AIR2, and *indica* varieties had lower AIR2 expression and As content compared with *japonica*. These results suggest that *indica* is relatively less likely to be exposed to As risk compared with *japonica* and also means that low expression of AIR2 can reduce As accumulation in rice. Therefore, rice varieties (eight indica, two temperate *japonica*, and two tropical *japonica*) with low AIR2 expression and low As content were finally selected based on these results.

CONCLUSION

The correlations between arsenic and functional ions were affected by rice subspecies and growing condition through the

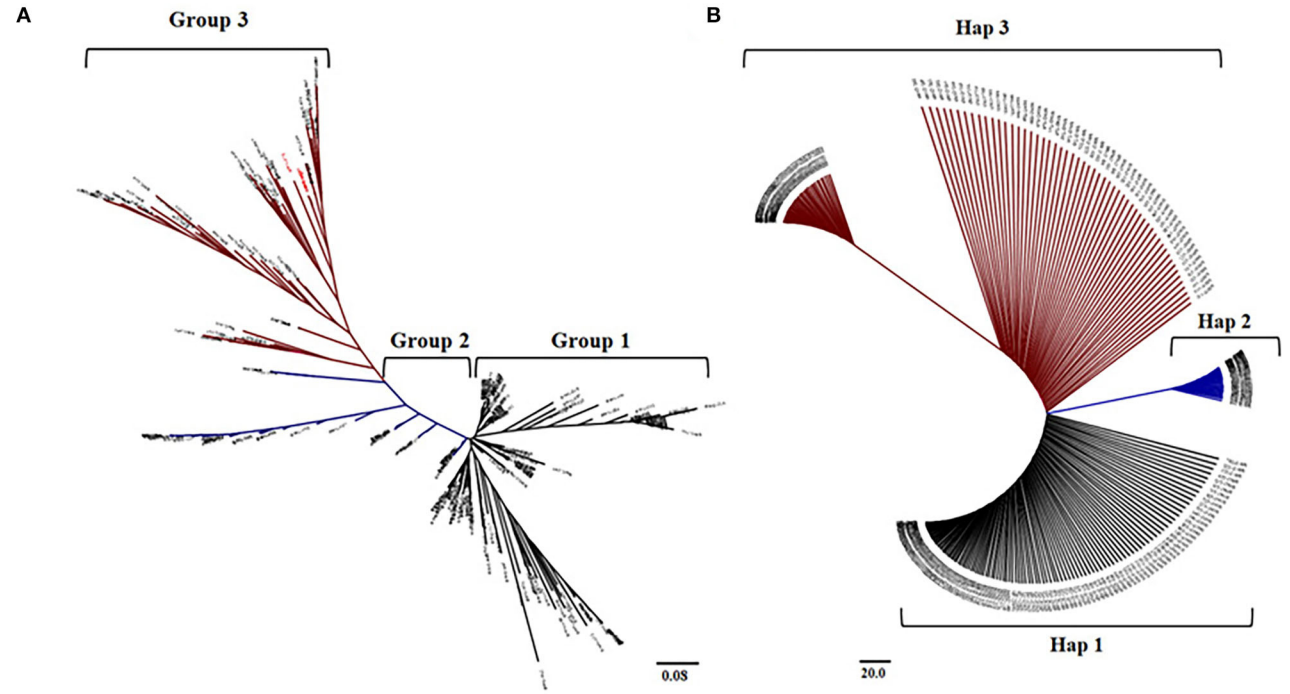


FIGURE 7 | Comparison between haplotype group and cis-eQTLs of AIR2. **(A)** Neighbor-joining tree and clustering groups for 273 accessions at the cis-eQTLs associated with AIR2 expression levels. **(B)** Neighboring binding trees and haplotype groups of 273 accessions to the genomic sequence of AIR2.

water management. As was reduced under the intermittently flooded condition compared with the flooded condition, but As reduction through the water management has limitations in terms of the rice-growing condition. Therefore, it was verified that the arsenic accumulation was lower in *indica* than in *japonica* through the association studies between transcriptome and genome data of rice core collections. AIR2 (arsenic-induced RING finger protein) expression for *indica* was relatively lower than that of *japonica* in its *cis*-eQTLs, and it was confirmed that the arsenic content of the group including a number of *indica* was also low in haplotype group for AIR2. Finally, rice varieties with low AIR2 expression and low arsenic content were selected. Therefore, the selected rice varieties are valuable as breeding materials, so verification through follow-up studies is required. In addition, it is necessary to develop a low-arsenic marker that can be applied to all rice varieties by converting three nonsynonymous SNPs into KASP markers in future through the SNP mutation information of AIR2.

DATA AVAILABILITY STATEMENT

The whole genome resequencing information for the rice core collection (273 rice accessions) is available in "Rice evolution analysis based on the chloroplast genome" by NCBI (https://www.ncbi.nlm.nih.gov/).

AUTHOR CONTRIBUTIONS

S-BL and S-WP designed the research. S-BL performed the experiments and wrote the initial manuscript. J-DS provided technical advice. G-JK and G-HC analyzed arsenic contents of rice core collection. S-WP and Y-JP supervised the study. Y-JP preprocessed rice whole-genome resequencing information and whole transcriptome data used in this study. S-KP, J-DS, and WC reviewed the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was funded by the Next-Biogreen 21 Programme, under grant agreement no. PJ013405, Studies on the metabolic behavior of harmful substances, such as soil persistence pesticides and crop uptake no. PJ01594403, and Development of low-allergy soybean breeding materials and molecular marker to improve added value no. PJ01678402.

ACKNOWLEDGMENTS

We thank Dool-Yi Kim, Mi-Suk Seo, Jung-Kyung Moon, and Yuna Kim at the National Institute of Crop Science (NICS) and Kyu-Won Kim, Sang-Ho Chu, and Myeong-Hyeon Min at the Kongju National University for their assistance in carrying out the project. We also thank Kyeong Ae Son and Ji-Hyock Ryu

at the National Institute of Agricultural Sciences (NIAS) for providing technical advice and support during fieldwork.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.905842/full#supplementary-material

Supplementary File 1 | Regional climate information for 2016 and 2017 in Yesan-gun, Chungcheongnam-do, Republic of Korea.

Supplementary File 2 | DNA sequencing and rice core collection information for 279 accessions.

Supplementary File 3 | Expression levels (FPKM) of arsenic genes for 234 rice accessions.

Supplementary File 4 | Differences between genetic models applied to GWAS for arsenic content in brown rice. It was verified that the *p*-value of the GWAS result was inflated using the lambda values (a value between 0 and 1). **(A)** GWAS results of applying the general linear model (GLM) and the linear mixed model (LMM) to the arsenic content under the flooded condition. **(B)** GWAS results of applying the general linear model (GLM) and the linear mixed model (LMM) to the arsenic content under the intermittently flooded condition.

Supplementary File 5 | Linkage disequilibrium for As-related QTLs. **(A,B)** Linkage disequilibrium for arsenic-associated QTLs on chromosome 6 under the flooded condition. **(C–E)** Linkage disequilibrium for arsenic-associated QTLs on chromosome 5 under the intermittently flooded condition. **(F,G)** Linkage disequilibrium for arsenic-associated QTLs on chromosome 1 under the intermittently flooded condition.

Supplementary File 6 | eQTLs candidate genes associated with STR5, STR8, and AIR2 expression levels.

Supplementary File 7 | Candidate genes identified from GWAS for arsenic contents under flooded and intermittently flooded conditions.

Supplementary File 8 | Genome-wide scale profiling for functional ions. **(A)** QTLs associated with functional ions in brown rice under the flooded condition. **(B)** QTLs associated with functional ions in brown rice under the intermittently flooded condition.

REFERENCES

Baxter, I. (2009). Ionomics: studying the social network of mineral nutrients. *Curr. Opin. Plant Biol.* 12, 381–386. doi: 10.1016/j.pbi.2009.05.002

Baxter, I. (2010). Ionomics: The functional genomics of elements. *Brief. Funct. Genomics* 9, 149–156. doi: 10.1093/bfgp/elp055

Bienert, G. P., Thorsen, M., Schüssler, M. D., Nilsson, H. R., Wagner, A., Tamás, M. J., et al. (2008). A subgroup of plant aquaporins facilitate the bi-directional diffusion of $As(OH)_3$ and $Sb(OH)_3$ across membranes. *BMC Biol.* 6, 1–15. doi: 10.1186/1741-7007-6-26

Browning, B. L., and Browning, S. R. (2007). Efficient multilocus association testing for whole genome association studies using localized haplotype clustering. *Genet. Epidemiol.* 31, 365–375. doi: 10.1002/gepi.20216

Chen, W. Q., Shi, Y. L., Wu, S. L., and Zhu, Y. G. (2016). Anthropogenic arsenic cycles: a research framework and features. *J. Clean. Prod*. 139, 328–336 doi: 10.1016/j.jclepro.2016.08.050

Danku, J. M. C., Lahner, B., Yakubova, E., and Salt, D. E. (2013). *Large-Scale Plant Ionomics*. *Plant Mineral Nutrients*. Totowa, NJ: Humana Press.

Das, C. K., Bastia, D., Swain, S. C., and Mahapatra, S. S. (2018a). Computational analysis of genes encoding for molecular determinants of arsenic tolerance in rice (Oryza sativa L.) to engineer low arsenic content varieties. *ORYZA-An Int. J. Rice* 55, 248–259. doi: 10.5958/2249-5266.2018.00031.0

Das, N., Bhattacharya, S., Bhattacharyya, S., and Maiti, M. K. (2017). Identification of alternatively spliced transcripts of rice phytochelatin synthase 2 gene OsPCS2 involved in mitigation of cadmium and arsenic stresses. *Plant Mol. Biol*. 94, 167–183. doi: 10.1007/s11103-017-0600-1

Das, N., Bhattacharya, S., Bhattacharyya, S., and Maiti, M. K. (2018b). Expression of rice MATE family transporter OsMATE2 modulates arsenic accumulation in tobacco and rice. *Plant Mol. Biol*. 98, 101–120. doi: 10.1007/s11103-018-0766-1

De Lorenzo, L., Merchan, F., Laporte, P., Thompson, R., Clarke, J., Sousa, C., et al. (2009). A novel plant leucine-rich repeat receptor kinase regulates the response of Medicago truncatula roots to salt stress. *Plant Cell* 21, 668–680. doi: 10.1105/tpc.108.059576

Dixit, G., Singh, A. P., Kumar, A., Dwivedi, S., Deeba, F., Kumar, S., et al. (2015). Sulfur alleviates arsenic toxicity by reducing its accumulation and modulating proteome, amino acids and thiol metabolism in rice leaves. *Sci. Rep.* 5, 1–16. doi: 10.1038/srep16205

Duxbury, J. M., Mayer, A. B., Lauren, J. G., and Hassan, N. (2003). Food Chain aspects of arsenic contamination in bangladesh: effects on quality and productivity of rice. *J. Environ. Sci. Health A* 38, 61–69 doi: 10.1081/ESE-120016881

Famoso, A. N., Zhao, K., Clark, R. T., Tung, C. W., Wright, M. H., Bustamante, C., et al. (2011). Genetic architecture of aluminum tolerance in rice (*Oryza sativa*) determined through genome-wide association analysis and QTL mapping. *PLoS Genet.* 7, e1002221. doi: 10.1371/journal.pgen.1002221

Finnegan, P. M., and Chen, W. (2012). Arsenic toxicity: the effects on plant metabolism. *Front. Physiol*. 3, 182. doi: 10.3389/fphys.2012.00182

Garcia-Oliveira, A. L., Tan, L., Fu, Y., and Sun, C. (2009). Genetic identification of quantitative trait loci for contents of mineral nutrients in rice grain. *J. Integr. Plant Biol.* 51, 84–92. doi: 10.1111/j.1744-7909.2008.00730.x

Hartwig, F. (1973). Statistical significance of the lambda coefficients. *Behav. Sci.* 18, 307–310 doi: 10.1002/bs.3830180409

Heo, E. B., Yoo, J. M., Lee, W. D., Chu, S. H., Kim, K. W., Cho, Y. H., et al. (2017). Integrated genome-wide association studies to dissect natural variation for magnesium ion contents in rice germplasm. *Korean J. Breed. Sci*. 49, 141–149. doi: 10.9787/KJBS.2017.49.3.141

Hu, P., Huang, J., Ouyang, Y., Wu, L., Song, J., Wang, S., et al. (2013). Water management affects arsenic and cadmium accumulation in different rice cultivars. *Environ. Geochem. Health.* 35, 767–778. doi: 10.1007/s10653-013-9533-z

Huang, X., Wei, X., Sang, T., Zhao, Q., Feng, Q., Zhao, Y., et al. (2010). Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat. Genet*. 42, 961–967. doi: 10.1038/ng.695

Huang, X., Zhao, Y., Wei, X., Li, C., Wang, A., Zhao, Q., et al. (2012). Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat. Genet*. 44, 32–39. doi: 10.1038/ng.1018

Huang, X. Y., and Salt, D. E. (2016). Plant ionomics: from elemental profiling to environmental adaptation. *Mol. Plant.* 9, 787–797. doi: 10.1016/j.molp.2016.05.003

Hwang, S. G., Chapagain, S., Han, A. R., Park, Y. C., Park, H. M., Kim, Y. H., et al. (2017). Molecular characterization of rice arsenic-induced RING finger E3 ligase 2 (OsAIR2) and its heterogeneous overexpression in Arabidopsis thaliana. *Physiol. Plant*. 161, 372–384. doi: 10.1111/ppl.12607

Hwang, S. G., Kim, D. S., and Jang, C. S. (2011). Comparative analysis of evolutionary dynamics of genes encoding leucine-rich repeat receptor-like kinase between rice and *Arabidopsis*. *Genetica* 139, 1023–1032. doi: 10.1007/s10709-011-9604-y

Hwang, S. G., Park, H. M., Han, A. R., and Jang, C. S. (2016). Molecular characterization of Oryza sativa arsenic-induced RING E3 ligase 1 (OsAIR1): expression patterns, localization, functional interaction, and heterogeneous overexpression. *J. Plant Physiol.* 191, 140–148. doi: 10.1016/j.jplph.2015.12.010

Isayenkov, S. V., and Matouis, F. J. M. (2008). The Arabidopsis thaliana aquaglyceroporin AtNIP7;1 is a pathway for arsenite uptake. *FEBS Lett.* 582, 1625–1628. doi: 10.1016/j.febslet.2008.04.022

Jiang, S., Shi, C., and Wu, J. (2012). Genotypic differences in arsenic, mercury, lead and cadmium in milled rice (*Oryza sativa* L.). *Int. J. Food Sci. Nutr*. 63, 468–475. doi: 10.3109/09637486.2011.636343

Kang, D. W., Kim, D. Y., Yoo, J. H., Park, S. W., Oh, K. S., Kwon, O. K., et al. (2018). Effect of soil amendments on arsenic reduction of brown rice in paddy fields. *Korean J. Soil Sci. Fertilizer.* 51, 101–110. doi: 10.7745/KJSSF.2018.51.2.101

Kawahara, Y., de la Bastide, M., Hamilton, J. P., Kanamori, H., McCombie, W. R., Ouyang, S., et al. (2013). Improvement of the Oryza sativa Nipponbare reference genome using next generation sequence and optical map data. *Rice* 6, 1–10. doi: 10.1186/1939-8433-6-4

Kim, B., Dai, X., Zhang, W., Zhuang, Z., Sanchez, D. L., Lübberstedt, T., et al. (2019). GWASpro: a high-performance genome-wide association analysis server. *Bioinformatics* 35, 2512–2514. doi: 10.1093/bioinformatics/bty989

Kim, K. W., Chung, H. K., Cho, G. T., Ma, K. H., Chandrabalan, D., Gwag, J. G., et al. (2007). PowerCore: a program applying the advanced M strategy with a heuristic search for establishing core sets. *Bioinformatics* 23, 2155–2162. doi: 10.1093/bioinformatics/btm313

Kirk, C., Jensen, M., Kjaer, C. N., Smedskjaer, M. M., Larsen, K. L., Wimmer, R., et al. (2009). Aqueous batch rebinding and selectivity studies on sucrose imprinted polymers. *Biosensors Bioelectr*. 25, 623–628. doi: 10.1016/j.bios.2009.01.021

Kumarathilaka, P., Seneweera, S., Meharg, A., and Bundschuh, J. (2018a). Arsenic speciation dynamics in paddy rice soil-water environment: sources, physico-chemical, and biological factors-a review. *Water Res*. 140, 403–414. doi: 10.1016/j.watres.2018.04.034

Kumarathilaka, P., Seneweera, S., Meharg, A., and Bundschuh, J. (2018b). Arsenic accumulation in rice (*Oryza sativa* L.) is influenced by environment and genetic factors. *Sci. Total Environ.* 642, 485–496. doi: 10.1016/j.scitotenv.2018.06.030

Lahner, B., Gong, J., Mahmoudian, M., Smith, E. L., Abid, K. B., Rogers, E. E., et al. (2003). Genomic scale profiling of nutrient and trace elements in *Arabidopsis thaliana*. *Nat. Biotechnol*. 21, 1215–1221. doi: 10.1038/nbt865

Latowski, D., Kowalczyk, A., Nawieśniak, K., and Listwan, S. (2018). "Arsenic uptake and transportation in plants," in *Mechanisms of Arsenic Toxicity and Tolerance in Plants* (Singapore: Springer), 1–26.

Le Sech, L., and Christian, H. (1991). Diallel analysis in white lupin: consequences for breeding. *Agronomie* 11, 719–726. doi: 10.1051/agro:19910902

Li, L., Ye, H., Guo, H., and Yin, Y. (2010). Arabidopsis IWS1 interacts with transcription factor BES1 and is involved in plant steroid hormone brassinosteroid regulated gene expression. *Proc. Natl. Acad. Sci. U.S.A*. 107, 3918–3923. doi: 10.1073/pnas.0909198107

Liao, J. L., Zhou, H. W., Peng, Q., Zhong, P. A., Zhang, H. Y., He, C., et al. (2015). Transcriptome changes in rice (*Oryza sativa* L.) in response to high night temperature stress at the early milky stage. *BMC Genomics* 16, 18. doi: 10.1186/s12864-015-1222-0

Liao, V. H. C., and Ou, K. L. (2005). Development and testing of a green fluorescent protein-based bacterial biosensor for measuring bioavailable arsenic in contaminated groundwater samples. *Environ. Toxicol. Chem*. 24, 1624–1631. doi: 10.1897/04-500R.1

Lim, S. D., Hwang, J. G., Han, A. R., Park, Y. C., Lee, C., Ok, Y. S., et al. (2014). Positive regulation of rice RING E3 ligase OsHIR1 in arsenic and cadmium uptakes. *Plant Mol. Biol.* 85, 365–379. doi: 10.1007/s11103-014-0190-0

Lu, F. H., and Park, Y. J. (2012a). Sequence variations in OsAGPase significantly associated with amylose content and viscosity properties in rice (*Oryza sativa* L.). *Genet. Res*. 94, 179–189. doi: 10.1017/S0016672312000390

Lu, F. H., and Park, Y. J. (2012b). An SNP downstream of the OsBEIIb gene is significantly associated with amylose content and viscosity properties in rice (*Oryza sativa* L.). *J. Cereal Sci*. 56, 706–712. doi: 10.1016/j.jcs.2012.08.007

Ma, J. F., Yamaji, N., Mitani, N., Xu, X. Y., Su, Y. H., McGrath, S. P., et al. (2008). Transporters of arsenite in rice and their role in arsenic accumulation in rice grain. *Proc. Natl. Acad. Sci. U.S.A*. 105, 9931–9935. doi: 10.1073/pnas.0802361105

McCouch, S. R., Wright, M. H., Tung, C. W., Maron, L. G., McNally, K. L., Fitzgerald, M., et al. (2016). Open access resources for genome-wide association mapping in rice. *Nature* 7, 1–14. doi: 10.1038/ncomms10532

Meharg, A. A., Williams, P. N., Adomako, A., Lawgali, Y. Y., Deacon, C., Villada, A., et al. (2009). Geographical variation in total and inorganic arsenic content of polished (White) rice. *Environ. Sci. Technol*. 43, 1612–1617. doi: 10.1021/es802612a

Most, P., and Papenbrock, J. (2015). Possible roles of plant sulfurtransferases in detoxification of cyanide, reactive oxygen species, selected heavy metals and arsenate. *Molecules* 20, 1410–1423. doi: 10.3390/molecules20011410

Muehe, E. M., Wang, T., Kerl, C. F., Planer-Friedrich, B., and Fendorf, S. (2019). Rice production threatened by coupled stresses of climate and soil arsenic. *Nat. Commun.* 10, 1–10. doi: 10.1038/s41467-019-12946-4

Na, G. N., and Salt, D. E. (2011). The role of sulfur assimilation and sulfur-containing compounds in trace element homeostasis in plants. *Environ. Exp. Bot*. 72, 18–25. doi: 10.1016/j.envexpbot.2010.04.004

NAAS (2010). *Analysis methods for soil chemical properties*. National Academy of Agriculture Science, Republic of Korea Publication No. 11-1390802-000282-01.

Nguyen, Q. T. T., Huang, T. L., and Huang, H. J. (2014). Identification of genes related to arsenic detoxification in rice roots using microarray analysis. *Int. J. Biosci. Biochem. Bioinform*. 4, 22. doi: 10.7763/IJBBB.2014.V4.304

Nicula, C., Peter, A., Mihaly-Cozmuta, L., and Mihaly-Cozmuta, A. (2013). The uptake of heavy metals in phaseolus vulgaris and zea mays seeds harvested from polluted and unpolluted areas. *Carpathian. J. Food Sci. Technol*. 5, 1–8. Available online at: http://stiinte.ubm.ro/multimedia/Chimie/index.html

Norton, G. J., Pinson, S. R. M., Alexander, J., Mckay, S., Hansen, H., Duan, G. L., et al. (2012). Variation in grain arsenic assessed in a diverse panel of rice (*Oryza sativa*) grown in multiple sites. *New Phytol.* 193, 650–664. doi: 10.1111/j.1469-8137.2011.03983.x

Norton, G. J., Travis, A. J., Talukdar, P., Hossain, M., Islam, M. R., Douglas, A., et al. (2019). Genetic loci regulating arsenic content in rice grains when grown flooded or under alternative wetting and drying irrigation. *Rice* 12, 1–15. doi: 10.1186/s12284-019-0307-9

Oh, M. H., Wang, X., Wu, X., Zhao, Y., Clouse, S. D., and Huber, S. C. (2010). Autophosphorylation of Tyr-610 in the receptor kinase BAK1 plays a role in brassinosteroid signaling and basal defense gene expression. *Proc. Natl. Acad. Sci. U.S.A*. 107, 17827–17832. doi: 10.1073/pnas.0915064107

Paszkowski, U., Kroken, S., Roux, C., and Briggs, S. P. (2002). Rice phosphate transporters include an evolutionarily divergent gene specifically activated in arbuscular mycorrhizal symbiosis. *Proc. Natl. Acad. Sci. U.S.A.* 99, 13324–13329. doi: 10.1073/pnas.202474599

Piepho, H. P., Möhring, J., Melchinger, A. E., and Büchse, A. (2008). BLUP for phenotypic selection in plant breeding and variety testing. *Euphytica* 161, 209–228. doi: 10.1007/s10681-007-9449-8

Pitorre, D., Llauro, C., Jobet, E., Guilleminot, J., Brizard, J. P., Delseny, M., et al. (2010). RLK7, a leucine-rich repeat receptor-like kinase, is required for proper germination speed and tolerance to oxidative stress in *Arabidopsis thaliana*. *Planta* 232, 1339–1353. doi: 10.1007/s00425-010-1260-4

QTLmax Global. (2022). *QTLMax 3.0: An Analytical Toolset for Enterprises in Agribusiness*. Katy, TX. Available online at: https://www.qtlmax.com

Rahaman, S., and Ashim, C. S. (2013). Water regimes: an approach of mitigation arsenic in summer rice (*Oryza sativa* L.) under different topo sequences on arsenic-contaminated soils of Bengal delta. *Paddy Water Environ*. 11, 397–410. doi: 10.1007/s10333-012-0331-5

Sakai, H., Lee, S. S., Tanaka, T., Numa, H., Kim, J., Kawahara, Y., et al. (2013). Rice annotation project database (RAP-DB): an integrative and interactive database for rice genomics. *Plant Cell Physiol.* 54, e6–e6. doi: 10.1093/pcp/pcs183

Salt, D. E. (2017). Would the real arsenate reductase please stand up?. *New Phytol.* 215, 926–928. doi: 10.1111/nph.14691

Salt, D. E., Baxter, I., and Lahner, B. (2008). Ionomics and the study of the plant ionome. *Annu. Rev. Plant Biol.* 59, 709–733. doi: 10.1146/annurev.arplant.59.032607.092942

Shi, S., Wang, T., Chen, Z., Tang, Z., Wu, Z., Salt, D. E., et al. (2016). OsHAC1; 1 and OsHAC1; 2 function as arsenate reductases and regulate arsenic accumulation. *Plant Physiol*. 172, 1708–1719. doi: 10.1104/pp.16.01332

Singh, P., Khan, A., and Srivastava, A. (2021). "Biological means of arsenic minimization with special reference to siderophore," in *Arsenic Toxicity: Challenges and Solutions* (Singapore: Springer), 253–278.

Song, W. Y., Yamaki, T., Yamaji, N., Ko, D., Jung, K. H., Fujii-Kashino, M., et al. (2014). A rice ABC transporter, OsABCC1, reduces arsenic accumulation in the grain. *Proc. Natl. Acad. Sci. U.S.A.* 111, 15699–15704. doi: 10.1073/pnas.1414968111

Sun, S. K., Chen, Y., Che, J., Konishi, N., Tang, Z., Miller, A. J., et al. (2018). Decreasing arsenic accumulation in rice by overexpressing Os NIP 1; 1 and Os NIP 3; 3 through disrupting arsenite radial transport in roots. *New Phytol*. 219, 641–653. doi: 10.1111/nph.15190

Tan, Y., Sun, L., Song, Q., Mao, D., Zhou, J., Jiang, Y., et al. (2020). Genetic architecture of subspecies divergence in trace mineral accumulation and elemental correlations in the rice grain. *Theor. Appl. Genet*. 133, 529–545. doi: 10.1007/s00122-019-03485-z

Tiwari, P., Indoliya, Y., Chauhan, A. S., Singh, P., Singh, P. K., Singh, P. C., et al. (2020). Auxin-salicylic acid cross-talk ameliorates OsMYB–R1 mediated defense towards heavy metal, drought and fungal stress. *J. Hazard. Mater*. 399, 122811. doi: 10.1016/j.jhazmat.2020.122811

Ullrich-eberius, C. I., Sanz, A., and Novacky, A. J. (1989). Evaluation of arsenate- and vanadate-associated changes of electrical membrane potential and phosphate transport in lemna gibba G1. *J.Exp. Bot.* 40, 119–128. doi: 10.1093/jxb/40.1.119

Wang, C., Tang, Z., and Zhao, F. J. (2019). "OsRCS3 functions as a cytosolic O-acetylserine (thiol) lyase and regulates arsenic accumulation in rice," in *Environmental Arsenic in a Changing World: Proceedings of the 7th International Congress and Exhibition on Arsenic in the Environment (AS 2018), July 1-6, 2018, Beijing, PR China* (Beijing: CRC Press), 343.

Wang, J., Li, M., Wang, H., and Pan, Y. (2012). Identification of essential proteins based on edge clustering coefficient. *IEEE/ACM Trans. Comput. Biol. Bioinform*. 9, 1070–1080. doi: 10.1109/TCBB.2011.147

Wang, Y., Lv, K., Shi, C., Shi, C., Li, Y., Chen, X., et al. (2020). Variation in arsenic accumulation and translocation among 74 main rice cultivars in Jiangsu Province, China. *Environ. Sci. Pollut. Res.* 27, 26249–26261. doi: 10.1007/s11356-020-08994-9

Williams, P. N., Islam, S., Islam, R., Jahiruddin, M., Adomako, E., Soliaman, A. R. M., et al. (2009). Arsenic limits trace mineral nutrition (selenium, zinc, and nickel) in Bangladesh rice grain. *Environ. Sci. Technol.* 43, 8430–8436. doi: 10.1021/es901825t

Williams, P. N., Villada, A., Deacon,. C., Raab, A., Figuerola, J., Green, A. J., et al. (2007). Greatly enhanced arsenic shoot assimilation in rice leads to elevated grain levels compared to wheat and barley. *Environ. Sci. Technol*. 41, 6854–6859. doi: 10.1021/es070627i

Wu, H., Shen, Y., Hu, Y., Tan, S., and Lin, Z. (2011). A phytocyanin-related early nodulin-like gene, BcBCP1, cloned from Boea crassifolia enhances osmotic tolerance in transgenic tobacco. *J. Plant Physiol*. 168, 935–943. doi: 10.1016/j.jplph.2010.09.019

Yamaguchi, N., Ohkura, T., Takahashi, Y., Yuji Maejima, Y., and Arao, T. (2014). Arsenic distribution and speciation near rice roots influenced by iron plaques and redox conditions of the soil matrix. *Environ. Sci. Technol*. 48, 1549–1556. doi: 10.1021/es402739a

Yamaji, N., Sakurai, G., Mitani-Ueno, N., and Ma, J. F. (2015). Orchestration of three transporters and distinct vascular structures in node for intervascular transfer of silicon in rice. *Proc. Natl. Acad. Sci. U.S.A*. 112, 11401–11406. doi: 10.1073/pnas.1508987112

Yang, H. C., Fu, H. L., Lin, Y. F., and Rosen, B. P. (2012). Pathways of arsenic uptake and efflux. *Curr. Top. Membranes* 69, 325–358. doi: 10.1016/B978-0-12-394390-3.00012-4

Yang, M., Lu, K., Zhao, F. J., Xie, W., Ramakrishna, P., Wang, G., et al. (2018). Genome-wide association studies reveal the genetic basis of ionomic variation in rice. *Plant Cell* 30, 2720–2740. doi: 10.1105/tpc.18.00375

Zhang, J., Zhu, Y. G., Zeng, D. L., Cheng, W. D., Qian, Q., and Duan, G. L. (2008). Mapping quantitative trait loci associated with arsenic accumulation in rice (*Oryza sativa*). *New Phytol*. 17RFPR7, 350–356. doi: 10.1111/j.1469-8137.2007.02267.x

Zhao, W., Cho, G. T., Ma, K. H., Chung, J. W., Gwag, J. G., and Park, Y. J. (2010). Development of an allele-mining set in rice using a heuristic algorithm and SSR genotype data with least redundancy for the post-genomic era. *Mol. Breed*. 26, 639–651. doi: 10.1007/s11032-010-9400-x

Zhao, W. G., Chung, J. W., Kwon, S. W., Lee, J. H., Ma, K. H., and Park, Y. J. (2013). Association analysis of physicochemical traits on eating quality in rice (*Oryza sativa* L.). *Euphytica* 191, 9–21. doi: 10.1007/s10681-012-0820-z

Conflict of Interest: J-DS was employed by Bio-Technology of Multidisciplinary Sciences Co.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 06 September 2022
DOI 10.3389/fpls.2022.951737

OPEN ACCESS

EDITED BY
Prasanta Kumar Subudhi,
Louisiana State University,
United States

REVIEWED BY
Boon Chin Tan,
University of Malaya, Malaysia
Sung-Hwan Cho,
University of Missouri, United States

*CORRESPONDENCE
Dong Pei
pei.dong@caf.ac.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

RECEIVED 24 May 2022
ACCEPTED 29 July 2022
PUBLISHED 06 September 2022

CITATION
Chang Y, Song X, Li M, Zhang Q, Zhang P, Lei X and Pei D (2022) Characterization of walnut *JrWOX11* and its overexpression provide insights into adventitious root formation and development and abiotic stress tolerance.
Front. Plant Sci. 13:951737.
doi: 10.3389/fpls.2022.951737

Characterization of walnut *JrWOX11* and its overexpression provide insights into adventitious root formation and development and abiotic stress tolerance

Yingying Chang[1,2†], Xiaobo Song[2†], Mingjun Li[1], Qixiang Zhang[3], Pu Zhang[2], Xiashuo Lei[2] and Dong Pei[2*]

[1]Engineering Laboratory of Green Medicinal Material Biotechnology of Henan Province, Engineering Technology Research Center of Nursing and Utilization of Genuine Chinese Crude Drugs of Henan Province, College of Life Science, Henan Normal University, Xinxiang, China, [2]State Key Laboratory of Tree Genetics and Breeding, Key Laboratory of Tree Breeding and Cultivation of the State Forestry and Grassland Administration, Research Institute of Forestry, Chinese Academy of Forestry, Beijing, China, [3]The Nurturing Station for the State Key Laboratory of Subtropical Silviculture, School of Forestry and Biotechnology, Zhejiang A&F University, Lin'an, China

The well-developed root system enables plant survival under various environmental stresses. *WUSCHEL-RELATED HOMEOBOX GENE 11* (*WOX11*) plays a critical role in adventitious root formation and development in rice, *Arabidopsis*, and easy-to-root tree poplar. However, in difficult-to-root trees, the knowledge of *WOX11* during adventitious root formation and development remains scarce. In this study, the *JrWOX11* gene was isolated from a difficult-to-root tree walnut and heterologously expressed in the "84K" poplar. The results showed that *JrWOX11* contained a similar structure and sequence to the homologous genes in rice, *Arabidopsis*, and poplar, but had different numbers and types of motifs and *cis*-elements. JrWOX11 lacked the motif GGAIQY compared to that in easy-to-root trees. In addition, *JrWOX11* expression was induced by ABA, PEG, and NaCl treatments. Overexpression of *JrWOX11* in poplar promoted root initiation and significantly increased adventitious root (ARs) number, lateral roots (LRs) number, and root hair (RH) length. Furthermore, the aboveground biomass was notably increased under NaCl and PEG treatments in transgenic plants. When NaCl and PEG were removed, the survival rate, aerial shoot development, and *de novo* root organogenesis were also markedly enhanced in transgenic shoot cuttings.

The study provides valuable information on the differences between *JrWOX11* and the homologous genes in rice, *Arabidopsis*, and poplar, and supports the critical role of *JrWOX11* in the formation of AR and tolerance to salt and osmotic stresses.

KEYWORDS

***JrWOX11*, overexpression, adventitious root formation, abiotic stress tolerance, motif-"GGAIQY"**

Key message

- WOX11 in walnut lacks the motif GGAIQY compared to that in easy-to-root trees.
- *JrWOX11* overexpression enhanced stress tolerance by promoting the rooting capacity and regulating root system architecture.

Introduction

Walnut (*Juglans regia* L.) is the most widespread nut tree species globally and is thus economically important (Zhang et al., 2020). It is also a difficult-to-root tree (Pei and Gu, 2002). In difficult-to-root trees, such as some *Eucalyptus*, *Pinus*, *Juglans*, *Carya*, *Quercus,* and *Castanea* (Basheer-Salimia, 2007; Amissah et al., 2008; Liu H. et al., 2018), the capacity of cuttings to form AR is relatively low, especially when the mature plant organs, such as shoots, were used as explants (Pijut et al., 2011). This is a significant limitation in the clonal propagation of commercial germplasm in difficult-to-root trees.

To overcome the problems associated with the loss of, or reduced, the competence of difficult-to-root trees to form AR, extensive studies have been conducted (Klerk, 2002; Gonin et al., 2019). For easy-to-root plants, a high auxin concentration above the cut site (due to auxin biosynthesis and auxin flow induced by wounding) is generally enough for adventitious root formation (Da Costa et al., 2013; Zhang et al., 2019). The detached leaves of *Arabidopsis thaliana* placed on a B5 medium initiate adventitious root formation at the cut site without any exogenous plant hormones (Chen et al., 2014). By contrast, in difficult-to-root trees, rejuvenation treatment and exogenous auxin are two prerequisites for rooting (Pei et al., 2004). Some previous studies have suggested that the rejuvenation treatment could reduce the thickness and density of sclerenchyma between the cortex and phloem, which is beneficial to rooting (Maynard and Bassuk, 1990; Liu H. et al., 2018). Rejuvenation treatment also induced the expression of the *WOX11* gene in walnut (Chang et al., 2020). Given the role *WOX11/12* plays in the process of adventitious root formation in difficult-to-root trees, it is reasonable to hypothesize that there might be some key differences in *WOX11/12* between easy-to-root plants and difficult-to-root trees.

The *WOX11/12* gene is one member of the intermediate clade in the *WUSCHEL*-related homeobox (*WOX*) family (Liu and Xu, 2018). In *Arabidopsis*, the expression of *AtWOX11* and *AtWOX12* can be induced by auxin, and AtWOX11 acts redundantly with AtWOX12 to stimulate cell fate transition to root founder cells (Liu J. et al., 2014). In rice, *OsWOX11*, *OsWOX12A*, and *OsWOX12B* could respond rapidly to auxin, cytokinin, and abiotic stress stimuli of drought, salt, and cold (Cheng et al., 2014; Jiang et al., 2017). Overexpression of *OsWOX11* in rice improves crown root emergence and growth, with the well-developed root system enabling rice survival under drought stress (Zhao et al., 2009; Cheng et al., 2016). In easy-to-root tree species poplar, such as *P. tomentosa*, "Nan-lin895" (*P. deltoides* × *P. euramericana*) and "84K" poplar (*P. alba* × *P. glandulosa*), *WOX11/12s* not only promoted adventitious root formation and increased the number of AR on the cuttings but also induced ectopic roots in the aerial parts of transgenic poplars (Liu B. et al., 2014; Xu et al., 2015; Wang et al., 2020). A new study has also reported that *PagWOX11/12a* enhanced the plant response to salt stress through the control of redox metabolic pathways (Wang et al., 2021). Despite the functional importance of WOX11/12 protein in adventitious root formation and development and stress resistance in easy-to-root plants, the current knowledge regarding the *WOX11/12* gene in difficult-to-root trees is scarce.

In this work, the *JrWOX11* gene was cloned from hybrid walnut "Zhongningsheng" (*J. hindsii* × *J. regia,* "ZNS"). The *JrWOX11* gene contained a similar structure and conserved domain sequence to the homologous genes in rice, *Arabidopsis*, and poplar, but had different numbers and types of *cis*-elements. Expression analysis showed that the *JrWOX11* transcript was expressed at significantly higher levels in roots than in other tissues and was induced by abscisic acid (ABA), salt (NaCl), and polyethylene glycol (PEG). Additionally, the overexpression of *JrWOX11* promoted adventitious root formation in poplar. The AR number, LR number, LR length, and RH length were also increased. Moreover, the tolerance to

salt and osmotic stress (NaCl and PEG) in transgenic poplar (*P. alba* × *P. glandulosa* clone "84K") plants was enhanced. Our results demonstrated that *JrWOX11* plays a crucial role in adventitious root formation and confers salt and osmotic stress tolerance in walnut.

Materials and methods

Plant materials

Hybrid poplar (*P. alba* × *P. glandulosa*) clone "84K" was used for transformation. The sterile seedlings of "84K" were cultured on half MS medium supplemented with 0.05 mg·L^{-1} of indole-3-butyric acid (IBA) and 0.02 mg·L^{-1} of naphthylacetic acid (NAA) in the tissue culture lab of the China Academy of Forestry, Beijing, China. The growth conditions were kept at 23∼25°C with 16 h light/8 h dark photoperiod. The shoots were subcultured on the same amount of fresh medium every 4 weeks.

Tobacco plants (*Nicotiana benthamiana*) were grown in pots under an 8 h day/16 h night photoperiod at 23°C. After ∼5 weeks, when 7∼10 tobacco leaves became expanded, the top leaves (third∼fifth) were used for transient expression.

The 1-year-old seedlings of hybrid walnut "ZNS," a walnut stock cultivar characterized by high yields, good adaptability, and high graft compatibility, were cultured in the greenhouse of the China Academy of Forestry, Beijing, China. The growth condition was kept at ∼25°C with a humidity of 50∼ 60%.

RNA isolation, DNA isolation, gene cloning, and vector construction

An RNeasy Plant Mini Kit and an RNase-free DNase I kit (Qiagen, Hilden, Germany) were used to extract total RNA from the rejuvenated stem of "ZNS." Complementary DNA (cDNA) synthesized by the SuperScript II reverse transcriptase (Thermo Fisher Scientific, Vilnius, Lithuania) was used as the template for cloning the coding sequence (CDS) of the *JrWOX11* gene. The putative sequence of *WOX11* online (GenBank accession no. XM_018977839) was used for primers design. The PCR product was cloned into pDONR222 and then recombined into the pMDC32 vector and pEarleyGate104 vector (ABRC stock DB3-686) to produce *35S:JrWOX11* and *35S:YFP-JrWOX11* constructs, respectively. Both of the above constructs were introduced into *Agrobacterium tumefaciens* strain GV3101 by electroporation.

Genomic DNA was extracted from the leaves of 1-year-old seedlings of "ZNS" walnut *via* a Plant DNA Isolation Reagent kit (Takara, Kyoto, Japan) and was used as the template for cloning *JrWOX11* promoter (-2,000 bp to + 206 bp). The PCR product was cloned into the pCAMBI1031 vector. All the above gene-specific primers are shown in **Supplementary Table 1**.

The analysis of the *WOX11* gene

The protein sequences of *WOX11/12* homologous genes in monocotyledon *Oryza sativa*, dicotyledonous herb *Arabidopsis thaliana*, easy-to-root trees *Populus trichocarpa*, *P. tomentosa*, "Nan-lin895" (*P. deltoides* × *P. euramericana*), "84K" poplar (*P. alba* × *P. glandulosa*), *Salix purpurea*, *Malus domestica*, *Morus notabilis*, *Durio zibethinus*, and difficult-to-root trees *J. regia*, *Carya illinoinensis*, *Quercus suber*, *Quercus lobata*, and *Castanea mollissima* were downloaded from NCBI and phytozome website. The motif enrichment and sequence distances compared to JrWOX11 protein were analyzed by MegAlign software (DNASTAR. Madison, WI, United States) and MEME v4.9.0 online software.[1] The gff3 files of *JrWOX11* (refer to XM_018977839) in walnut and *OsWOX11*, *OsWOX12a*, *OsWOX12b*, *AtWOX11*, *AtWOX12*, *PrtWOX11/12a*, and *PrtWOX11/12b* in model plants were used to analyze gene structure by Gene Structure Display Server.[2]

The upstream regions (2.0 kb) of the translation initiation sites (ATG) of *WOX11/12* homologous genes in *O. sativa*, *A. thaliana*, and *P. trichocarpa* were used as promoter fragments for the *cis*-elements analysis compared with that of *JrWOX11* gene using the program PlantCARE online.[3]

Subcellular localization of *JrWOX11*

To evaluate *JrWOX11* localization within cells, we utilized a *35S:YFP-JrWOX11* construct, such that the *JrWOX11* coding region was fused to the YFP C-terminus. The *A. tumefaciens* GV3101 colony carrying the *35S:YFP* construct, RFP-H2B (a nucleus localization signal marker tagged with RFP) and OsSP1-RFP (a plasma membrane localization marker) were individually cultured to OD_{600} = 0.4–0.6, re-suspended with the 10 mM EMS, 10 mM $MgCl_2$ (pH 5.8) and 150 μM acetosyringone to OD_{600} = 0.2 (1:1:1 v/v/v), and co-infiltrated into *N. benthamiana* leaves as a control. The *A. tumefaciens* GV3101 cells containing the *35S:YFP-JrWOX11* construct and RFP-H2B were individually cultured, re-suspended, mixed, and co-infiltrated into *N. benthamiana* leaves as described above to clarify the subcellular localization of *JrWOX11*. After about 48∼56 h incubation at 23°C under an 8 h day/16 h night photoperiod, *N. benthamiana* leaves were collected to observe the excitation of YFP and RFP using an LSM 510 confocal laser scanning microscope (Carl Zeiss AG, Oberkochen, Germany) with fluorescence imaging at 488 and 561 nm, respectively, and emission being captured between 500 and 550 nm.

1 http://meme-suite.org/tools/meme
2 http://gsds.cbi.pku.edu.cn/
3 http://bioinformatics.psb.ugent.be/webtools/plantcare/html/

Tissue-specific and inducible expression of *JrWOX11* treated with abscisic acid and salt and osmotic stress

To reveal expression profiles of *JrWOX11* gene in different organs, including roots (R), stem (S), leaf (L), female flower (FF), immature fruit (IF), and zygotic embryo (ZE), they were collected from 23-year-old "ZNS" trees from Luoning county, Henan Province. Six samples of each tissue from the same tree were mixed as one biological replicate, and three replicate trees were sampled. All samples were snap-frozen in liquid nitrogen and stored at −80°C for the expression analysis of *JrWOX11* as described in the previous study (Chang et al., 2020).

To investigate the expression profiles of *JrWOX11* under salt and osmotic stress, the 1-year-old "ZNS" seedlings were irrigated with 200 mM NaCl and 10% (w/v) PEG 6000 solution, and the roots were collected after seven treatment durations (0, 1, 3, 6, 9, 12, and 24 h). For ABA treatment, the top fourth–sixth walnut leaves were sprayed with 100 μM ABA and collected after 0, 1, 3, 6, 9, 12, and 24 h. Ten leaf sections were used for each treatment, and experiments were repeated three times. All samples were snap-frozen in liquid nitrogen and stored at −80°C for the expression analysis of *JrWOX11*.

Total RNA was isolated using the method described above (Section "RNA isolation, DNA isolation, gene cloning, and vector construction"). The synthesized cDNA was diluted 10-fold with ddH_2O to serve as a template for qRT-PCR on a Roche LightCycler 480 (Roche Applied Science, Penzberg, Upper Bavaria, Germany). Expression levels were normalized relative to the control (*GAPDH*) using the $2-\triangle\triangle Ct$ method. The qRT-PCR primers are shown in **Supplementary Table 1**.

Stable expression of *JrWOX11* in transgenic hybrid poplar "84K"

An *A. tumefaciens* strain GV3101 containing the *35S:JrWOX11* construct was used for the transformation of poplar "84K" as described previously (Liu B. et al., 2014; Li et al., 2017). The discs from the top fourth–fifth leaves from 4-week-old "84K" were notched and incubated with *A. tumefaciens* GV3101 cells containing the *35S:JrWOX11* construct for about 3 days. The hygromycin-resistant shoots were obtained from the leaf explants after 2–4 weeks on the MS medium supplemented with 0.5 $mg{\cdot}L^{-1}$ of 6-BA, 0.05 $mg{\cdot}L^{-1}$ of NAA, 3 $mg{\cdot}L^{-1}$ of hygromycin, and 200 $mg{\cdot}L^{-1}$ of timentin. Roots were induced on root induction medium (RIM: 1/2 MS + 0.05 $mg{\cdot}L^{-1}$ of IBA + 0.02 $mg{\cdot}L^{-1}$ of NAA) + 3 $mg{\cdot}L^{-1}$ of hygromycin. The resistant plants were detected by qRT-PCR of the *JrWOX11* gene expression. The plants with high expression levels of the *JrWOX11* gene were subcultured for later experiments.

Plant phenotype determination

To determine the role of the *JrWOX11* gene in root formation, a stereo microscope (Leica, Wetzlar, Germany) was used to observe the stem samples from the control and transgenic (OE6#) "84K" after 0, 1, 2, 3, 4, 5, and 6 days on RIM. Cross-sections (30 μm thick) of the stem base were prepared with a VT-1000S vibrating blade microtome (VT-1000S, Leica, Wetzlar, Germany) and stained with 2% w/v toluidine blue O solution for visualizing root primordia under a BX51 digital microscope (Olympus, Tokyo, Japan). Five seedlings of each line per replicate were used, and two independent experiments were performed.

To analyze the rooting index, including the AR number, LR number, AR length, LR length, and root hair (RH) length, the shoots (about 2.5 cm high) were cut from 4-week-old non-transgenic "84K" (CK) and transgenic seedlings (OE#3 and OE#6) and moved to RIM under 16 h light/8 h dark for about 2 and 4 weeks, respectively. Ten individuals per replicate were used for each line, and three independent experiments were performed.

Analysis of salt and osmotic stress tolerance

To evaluate salt and osmotic stress tolerance, 2-week-old non-transgenic "84K" (CK) and transgenic (OE3# and OE6#) surface-sterilized plants were grown on RIM supplemented with 0 (control) or 200 mM NaCl (salt treatment) or 5% w/v PEG 6000. Three weeks later, the plant height and fresh weight (FW) were measured, and the dry weight (DW) was determined after oven drying at 80°C for 72 h. Six biological replicates of each line per treatment were used, and two independent experiments were performed.

To examine the rooting capacity under salt and osmotic stress, shoot segments (about 2.5 cm high) of 4-week-old non-transgenic "84K" (CK) and transgenic (OE3# and OE6#) surface-sterilized plants were cultivated on RIM supplemented with 0 (control) or 200 mM NaCl (salt treatment) or 5% w/v PEG 6000 for 4 weeks and then transferred to RIM for 2 weeks. The phenotypes of the root systems were photographed every week, and the survival rate, the number of axillary buds regenerated, and the number of ARs were quantified. Six biological replicates of each line per treatment were used, and two independent experiments were performed.

Statistical analysis

Significant differences between means of CK and transgenic lines were determined using the SPSS statistical package (version 16.0; SPSS Inc., Chicago, IL, United States) at a significance level of $^{*}P < 0.05$ and $^{**}P < 0.01$ (Student's *t*-test).

Results

Identification and characterization of *JrWOX11* gene

In "ZNS" walnut, only one *WOX11/12* gene was isolated and submitted to NCBI's GenBank (*JrWOX11*, NCBI Accession No: ON979687) (**Figure 1A**). The full-length sequence of *JrWOX11* cDNA is 747 nucleotides and encodes a deduced protein of 248 amino acids (aa) residues, with a predicted molecular mass of 27.05 kDa and a theoretical pI of 5.42 (**Supplementary Table 2**). However, there were 3, 2, and 2 *WOX11/12* genes in rice, *Arabidopsis*, and *P. trichocarpa*, respectively (**Figure 1A**) (Zhang et al., 2010; Cheng et al., 2014). Twenty-six WOX11/12 protein sequences from 15 species, namely, two monocotyledons (maize and rice), a dicotyledonous herb (*A. thaliana*), eight easy-to-root trees (*P. trichocarpa*, *P. tomentosa*, Nan-lin895 poplar, "84K" poplar, *S. purpurea*, *Malus domestica, Morus notabilis*, and *D. zibethinus*), and four difficult-to-root trees (*Carya illinoinensis*, *Q. suber*, *Q. lobata*, and *Castanea mollissima*), were downloaded from NCBI and characterized. The protein length ranged from 236 to 356 aa, the pIs from 5.42 to 8.67, and the molecular mass from 25.92 to 36.47 kDa. Multialignment of WOX11/12 protein sequences (by MegAlign software) showed that CiWOX11 (96.37%) in *Carya illinoinensis*, QlWOX11 (82.57%) in *Q. lobata*, QsWOX11 (81.82%) in *Q. suber*, and CmWOX11 (82.23%) in *Castanea mollissima* shared high sequence identity with JrWOX11 in "ZNS" walnut. The WOX11/12 in *Populus*, *S. purpurea*, *Malus domestica*, *Morus notabilis*, and *D. zibethinus* had relatively moderate sequence identities with JrWOX11, ranging from 61.48 to 74.44%. The sequence of WOX11/12 in *Z. mays*, *O. sativa* and *A. thaliana* had relatively low sequence identities with JrWOX11, ranging from 45.22 to 57.01% (**Supplementary Table 2**). The simplified neighbor-joining (NJ) phylogenetic tree of the related species also showed that JrWOX11 was similar to PtrWOX11/12a and PtrWOX11/12b (**Figure 1B**).

Gene structure analysis revealed that the homologous *WOX11* genes had the same gene structure, such as three exons and two introns (**Figure 1C**). Motif enrichment analysis revealed five prominent motifs (**Figure 1D**). Among them, motif1, motif3, and motif6 were related to the homeodomain, whereas motif2 and motif4 were associated with the C-terminus domain of intermediate clade WOXs (Zhang et al., 2010). Motif7 and motif8 were observed only in dicotyledons, and motif9 was conserved only in the easy-to-root trees. Even though AtWOX11 and AtWOX12 have a similar function in adventitious root formation, motif5 was missing in AtWOX12. Furthermore, motif10 was only found in easy-to-root trees, such as *P. trichocarpa*, *Ptr*; *S. purpurea*, *Sp*; *Malus domestica*, *Md*; *Morus notabilis*, *Mn*; and *D. zibethinus*, *Dz*; but was not present in difficult-to-root trees, such as *J. regia*, *Jr*; *Carya illinoinensis*, *Ci*; *Quercus suber*, *Qs*; *Quercus lobata*, *Ql,* and *Castanea mollissima*, *Cm*. The sequence of motif10 in PtrWOX11/12a was "QQAQASGGAIQYKGCNTSIGF," among which "GGAIQY" was extremely conserved (**Figures 1A,D** and **Supplementary Figure 2**). To sum up, there might be differences in the number of WOX11/12 proteins and the motifs between easy-to-root trees and difficult-to-root trees.

Analysis of *JrWOX11* promoter sequence

The 2 kb sequence upstream of *WOX11/12s* CDS was used to identify *cis*-acting elements. The results showed that the identified *cis*-elements could be divided into four categories: (1) growth- and development-responsive elements, (2) plant hormone-responsive, (3) abiotic stress-responsive elements, and (4) WOX-consensus motif (**Figure 2** and **Supplementary Table 3**). There were 3, 1, 2, 5, 5, 5, 3, and 2 root-specific elements in the promoter region of *OsWOX11*, *OsWOX12a*, *OsWOX12b*, *AtWOX11*, *AtWOX12*, *PtrWOX11/12a*, *PtrWOX11/12b,* and *JrWOX11*, respectively. Plant hormone-related elements were also identified, such as 2, 5, 4, 5, 3, 2, 3, and 10 auxin-responsive elements (ARFAT and TGA elements), 4, 1, 5, 8, 9, 8, 8, and 5 cytokinin-responsive element (ARR1AT), 1, 1, 0, 1, 3, 1, 0, and 2 gibberellin-responsive element (GARE1OSREP1), and 2, 3, 0, 0, 3, 1, 2, and 7 abscisic acid-responsive element (ABRE), in the promoter region of *OsWOX11*, *OsWOX12a*, *OsWOX12b*, *AtWOX11*, *AtWOX12*, *PtrWOX11/12a*, *PtrWOX11/12b,* and *JrWOX11*, respectively. A few elements responsive to various abiotic stresses were also identified, mainly including DRE, MBS, G-box, and W-box, and the total numbers were 0, 3, 3, 4, 1, 3, 4, and 1 in the above sequence of *WOX* genes. Compared to rice, *Arabidopsis*, and poplar, the promoter region of *WOX11* in walnut contained fewer root motifs and stress-responsive elements but contained more abscisic acid-responsive elements and AuxRE.

WOX11 is shown to specifically bind to the WOX-consensus "TTAATGG/C" motif (Lohmann et al., 2001; Zhao et al., 2009, 2015; Liu B. et al., 2018). Within the *OsWOX12*, *AtWOX12*, and *PtrWOX11/12b* genes, we identified one or two WOX-consensus motifs common to their promoters. However, walnut did not harbor any *WOX11/WOX12b* genes, suggesting that WOX11-binding genes might differ in different plants (**Figure 2** and **Supplementary Table 3**).

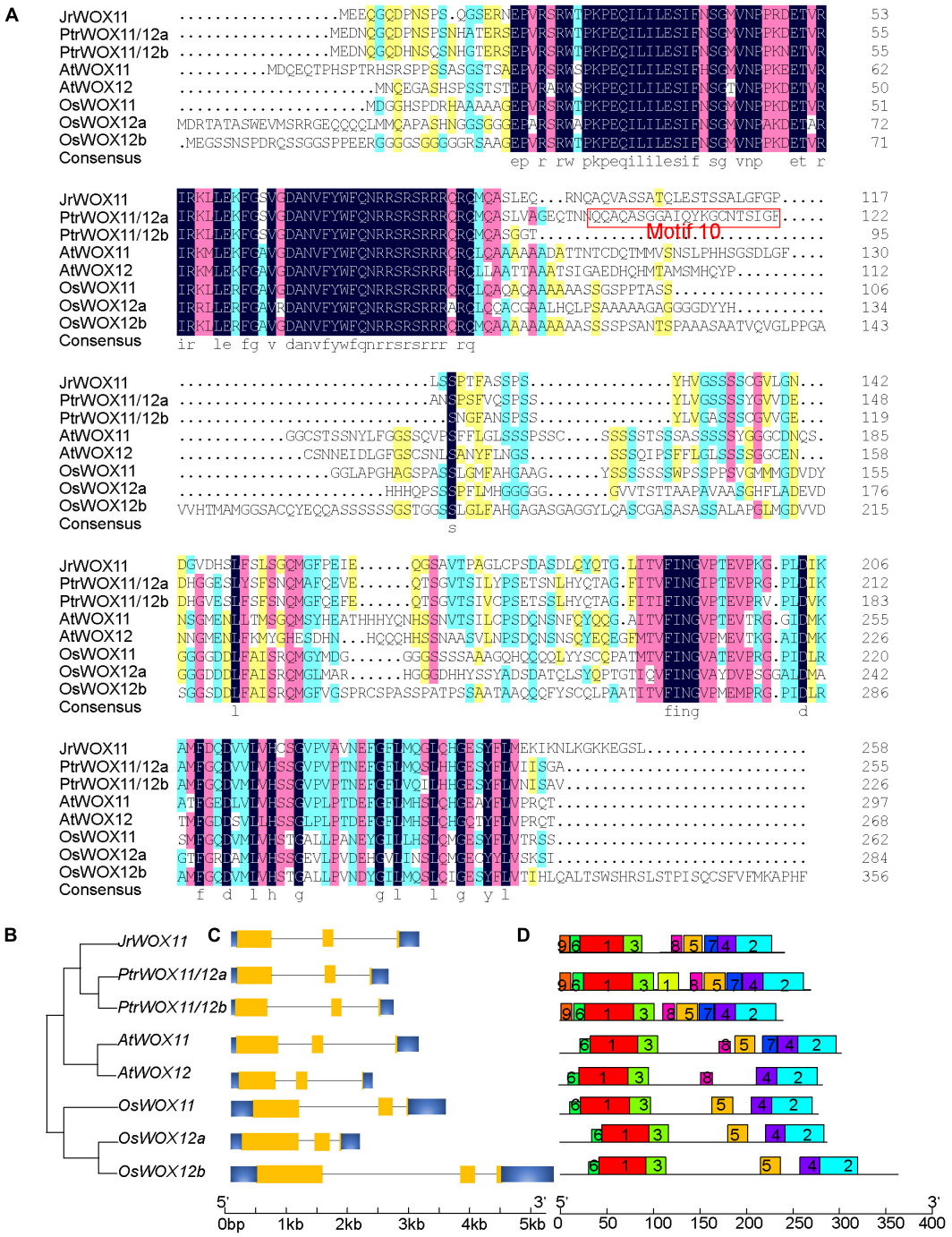


FIGURE 1
Identification and characterization of cloning of the *JrWOX11* gene. **(A)** Sequence alignment of homologous WOX11/12 proteins. Protein sequences were obtained from NCBI and are shown in **Supplementary Table 2**. **(B)** Phylogenetic tree based on full-length coding sequences of *WOX11/12* genes. **(C)** Gene structure analysis of *WOX11/12* genes. **(D)** Motifs analysis of WOX11 proteins. MEME software was used to search motifs, and then redraw the map with the software TBtools (Chen et al., 2020). Motif10 was marked with a red box in panel **(A)**.

Subcellular localization of *JrWOX11*

In our previous research, JrWOX11 was predicted to be localized in the nucleus (Chang et al., 2020). To validate this prediction, the *35S:YFP-JrWOX11* construct was transiently co-expressed with the nuclear marker in tobacco leaves. The *35S:YFP* construct mixed with the nucleus and the membrane markers was transformed as a control. The nucleus and

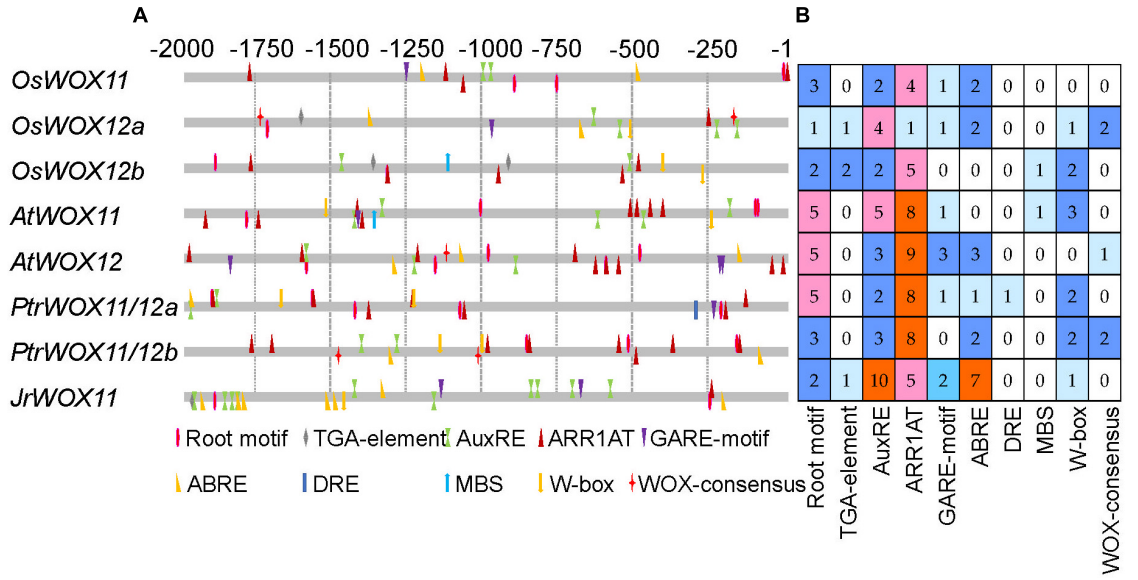


FIGURE 2
The *cis*-element distribution and number analysis of the promoter region of *WOX11/12* genes in *Oryza sativa*, *Arabidopsis thaliana*, *Populus trichocarpa*, and *Juglans hindsii* × *Juglans regia* cv. "ZNS." **(A)** The position of the *cis*-element in the upstream 2,000 bp region of *WOX11/12* genes. **(B)** The number of each *cis*-element.

plasma membrane were marked by RFP staining. As shown in **Figures 3e–h**, the 35S:YFP-JrWOX11 fluorescent signals and nucleus marker were observed in the same location, indicating that JrWOX11 is localized in the nucleus. This was consistent with the prediction. By contrast, in cells transformed with the positive control *35S:YFP*, the fusion protein was distributed throughout the cell, including the nucleus and cytoplasm, which were marked by the nuclear and plasma membrane RFP marker staining (**Figures 3a–d**).

Tissue-specific and abiotic stress response in the expression of the *JrWOX11* gene

The spatial and temporal expression patterns of the *JrWOX11* gene in different tissues were analyzed using qRT-PCR. As shown in **Figure 4A** and our previous study (Chang et al., 2020), the *JrWOX11* showed relatively high expression in roots (R), but relatively low expression in stems (S). Previous studies reported that the *WOX11* gene was involved in drought and salt resistance (Cheng et al., 2016; Wang et al., 2018, 2020). Here, the transcription levels of *JrWOX11* in response to ABA, NaCl, and PEG were analyzed. As shown in **Figures 4B–D**, *JrWOX11* transcription was rapidly induced after treatment with ABA, NaCl, and PEG. The transcription levels increased with prolonged treatments, peaking at 9, 6, and 6 h, respectively, then slightly decreased after that. It could be concluded that the expression of *JrWOX11* is responsive to ABA, NaCl, and PEG.

Overexpression of *JrWOX11* promoted adventitious root formation in "84K" poplar

The *JrWOX11* had the highest expression in roots, which may play a crucial role in root growth and development. To verify this, *JrWOX11* was transformed into "84K" poplar, resulting in 14 transgenic lines (**Supplementary Figure 1**). Lines OE3# and OE6# with high expression of JrWOX11 were used in the work reported here. The phenotypes of the stem base of control (CK) and the OE6# line were analyzed by stereo microscopy and sectioning (**Figures 5A,B**). The results showed that on day 5, small protrusions from the stem base were observed for the first time in control stems, which might have represented a bump of AR primordium. On day 6, the skin of the original protrusion at the base of the control stem ruptured and ARs appeared. By contrast, ARs in OE lines occurred on day 3, 2 days earlier than in the control. Therefore, the *JrWOX11* gene is considered to promote the emergence of adventitious roots.

Overexpression of *JrWOX11* influenced the development of the adventitious root system

To further verify whether JrWOX11 was involved in the development of the AR system, the root phenotypes of

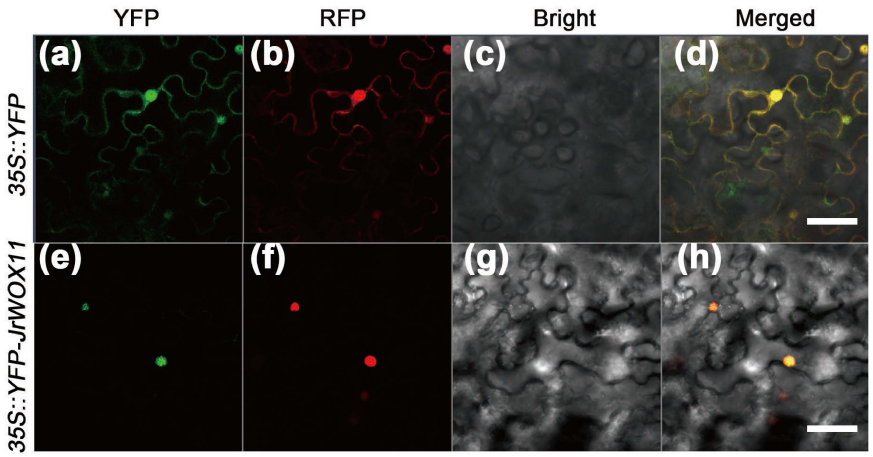


FIGURE 3
Subcellular localization of JrWO11 protein. 35S::YFP and 35S::YFP-JrWO11 fusion proteins were expressed transiently into *Nicotiana benthamiana* leaf epidermal cells 48∼56 h after vacuum injection. The YFP channel **(a,e)**, the RFP channel **(b,f)**, the bright field **(c,g)**, and the merged images **(d,h)**. Scale bar: 50 μm.

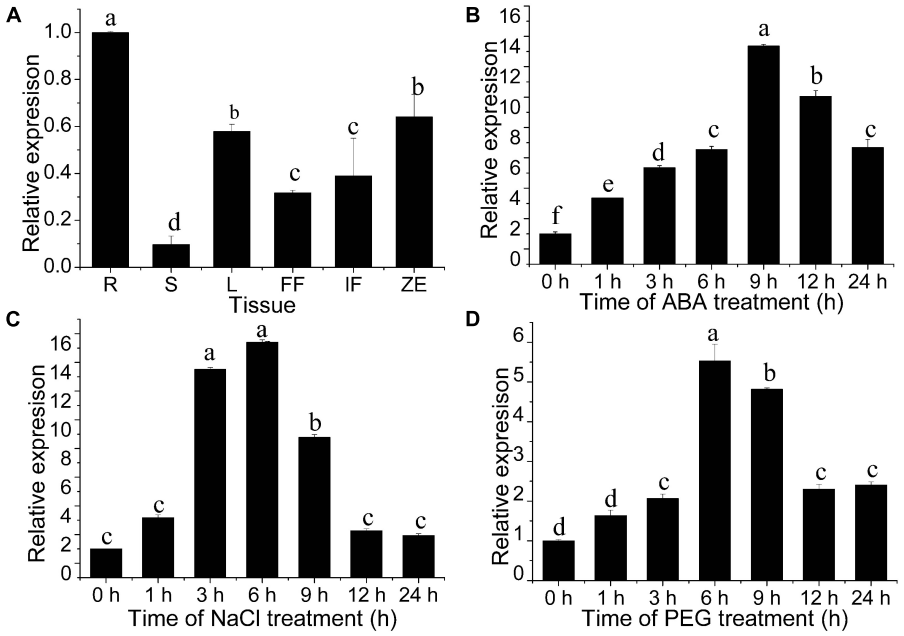


FIGURE 4
Expression patterns of *JrWOX11* as assessed by qRT-PCR. **(A)** Spatial and temporal expression of *JrWOX11* in different tissues. **(B–D)** Relative expression of *JrWOX11* under 100 μM ABA, salt (200 mM NaCl) and 10% (w/v) PEG 6000 solution. Error bars = SEM. Statistical differences were analyzed using Student's *t*-test, $n = 3$. Different letters labeled indicate significant differences.

two OE lines (3# and 6#) were analyzed (Figures 6A–C). In 2-week-old seedlings, the AR number, length, and LR length were significantly higher in *JrWOX11*-transgenic lines (OE3# and OE6#) than controls (CK1 and CK2) (Figures 6H,I). Similarly, these root indices in 4-week-old seedlings of OE3# and OE6# lines were also significantly increased compared with controls (Figures 6E,F). Interestingly, transgenic plants overexpressing *JrWOX11* had significantly increased root hair length (Figures 6D,G). These results suggested that *JrWOX11* might regulate root development and root hair growth.

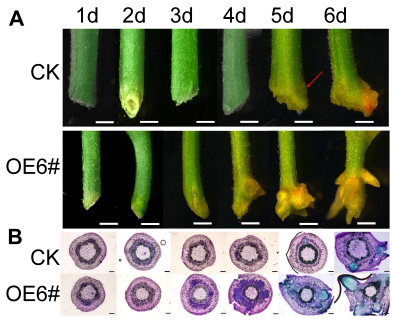


FIGURE 5
Overexpression of *JrWOX11* promotes root formation. **(A)** The phenotype investigation of the poplar stem base during rooting 1–6 days. Bar = 1 mm. **(B)** The histological investigation of the poplar stem base during rooting 1–6 days. Bar = 200 μm.

JrWOX11 increases tolerance to salt and osmotic stress

To validate the effectiveness of root system architecture in NaCl and osmotic stress tolerance, 2-week-old OE (3# and 6#) and control (CK) seedlings were subjected to no-stress treatment (Control), 200 mM NaCl, and 5% w/v PEG6000 treatment in sterile seed tubes (**Figures 7A,B**). Three weeks later, the plant height, the aboveground fresh weight, and dry weight of OE3# and OE6# were superior to CK in the control treatment, suggesting that the overexpression of the *JrWOX11* gene improved the growth of transgenic plants. Under salt and osmotic stress, the aboveground fresh weight and dry weight were significantly increased in *JrWOX11*-overexpressing plants compared with CK plants, indicating that *JrWOX11* improved tolerance to NaCl and osmotic stress by promoting root growth.

Adventitious root formation is a *de novo* organogenesis process that enables plants to cope with environmental stresses. To evaluate the growth and root regeneration response of *JrWOX11*-overexpressing plants and control type (CK) to NaCl and osmotic stress, shoot cuttings were subjected to 200 mM NaCl and 5% w/v PEG6000 stress for 4 weeks. The root and aboveground biomass were significantly improved in *JrWOX11*-overexpressing plants compared with CK shoot cuttings under no-stress conditions (**Figure 7C** and **Supplementray Figure 3**). Additionally, the transgenic shoot cuttings showed significantly improved salt and osmotic tolerance compared to CK plants (**Figure 7C**). The survival rate, the number of axillary buds regenerated, and the number of ARs were increased substantially 2 weeks after plants were transferred to media without 200 mM NaCl and 5% w/v PEG6000 (**Figure 7D**). These findings suggested that *JrWOX11* enhanced tolerance to salt and osmotic stress by improving the *de novo* root organogenesis and shoot development (Liu J. et al., 2014, Liu B. et al., 2018; Hu and Xu, 2016; Zhou et al., 2017).

Discussion

Recalcitrance to adventitious root formation from stem cuttings is a major limitation for clonal propagation in difficult-to-root trees (Macdonald, 1990). Given the energy and time required for maintaining difficult-to-root trees in a juvenile state in various environments, the development of rootstock varieties that are easy-to-root and tolerant to environmental stresses through molecular breeding is crucial. However, there is little knowledge about the genes conferring rooting capacity and abiotic stress tolerance. Previous studies have shown that the WOX11/12 gene, encoding a transcription factor, can be induced by salt and drought stress and that it participates in root formation and development in rice, *Arabidopsis*, and poplar (Zhao et al., 2009, 2015; Liu B. et al., 2014; Liu J. et al., 2014; Hu and Xu, 2016). Here, we isolated and characterized a *WOX11* gene in "ZNS" walnut, a difficult-to-root tree. In addition, in transgenic "84K" poplar overexpressing *JrWOX11*, the capacity to form adventitious roots and to develop the root system has been improved. Transgenic "84K" poplar overexpressing *JrWOX11* also showed increased NaCl and osmotic tolerance.

Previous studies have shown that *WOX11* is a crucial factor that controls normal cells shifting to root primordium cells (Liu J. et al., 2014). This process mainly depends on auxin biosynthesis and auxin transport (Sarkar et al., 2007; Chen et al., 2016; Wei et al., 2019). In *Arabidopsis*, AtWOX11 binds directly to the promoter of a LATERAL ORGAN BOUNDARIES DOMAIN transcription factor (AtLBD16) to result in the formation of LR root primordium cells (Wu et al., 2018). In the present study, overexpression of *JrWOX11* in "84K" poplar promoted the formation of root primordium cells 2s day earlier than in the control plants (**Figure 5**). It could be speculated that the *JrWOX11* gene regulated adventitious root formation *via* the auxin-dependent pathway.

Overexpression of *WOX11/12* genes, such as *OsWOX11*, *PtoWOX11*, *PeWOX11a*, *PeWOX11b,* and *PagWOX11/12*a resulted in increased numbers of adventitious and ectopic roots, which contributed to plant resistance to drought (Cheng et al., 2016; Wang et al., 2020). Furthermore, Wang et al. (2021) reported that *PagWOX11/12a* enhanced plant response to salt stress by controlling the redox metabolic pathways. Consistent with the previous studies in rice and poplar, *JrWOX11* not only increased the numbers of ARs, LRs, ectopic roots, and the length of RH but also improved the NaCl and osmotic tolerance in overexpressing "84K" poplar (**Figure 7** and **Supplementary Figure 3**). Notwithstanding its limitations, the study suggests that *JrWOX11* increases "84K" poplar salt and

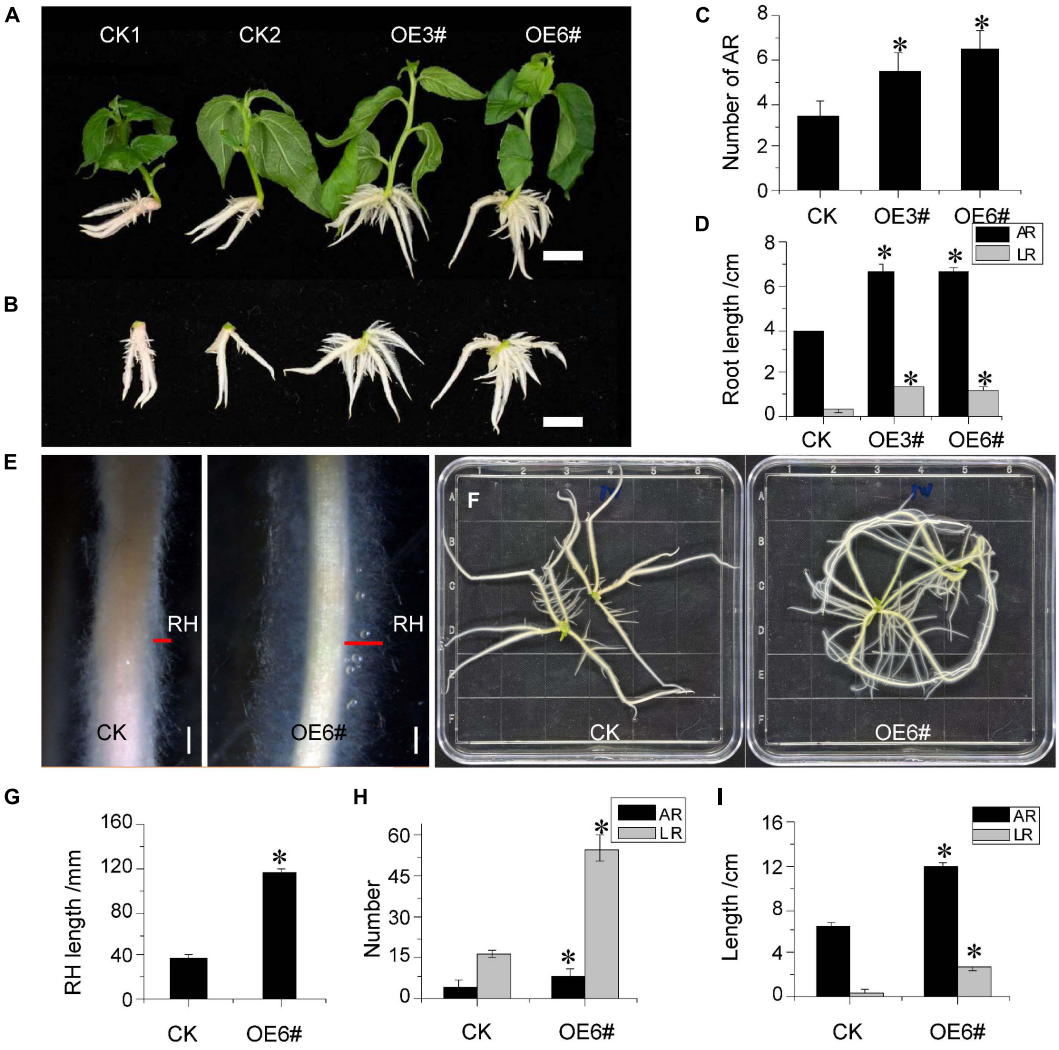


FIGURE 6
Overexpression of *JrWOX11* regulated root system architecture. **(A)** Comparison of AR number between CK and OE 3# and OE 3# 6# (2 weeks old). **(B)** The vertical view of ARs of CK and overexpressing lines 3# and 6# (2 weeks old). Bar = 1 cm. **(C,D)** Average AR length, AR number, and LR length in Control and the overexpressing lines (2 weeks old). **(E)** Root hairs of 4-week-old seedlings. Bars = 2 mm. **(F)** Root system phenotypes of the control group (CK) and WOX11 transgenic plants (OE#6) (4 weeks old). Bar = 1 cm. **(G–I)** Length of RH, average root length, and root number in Control and the overexpressing line (OE#6) (4 weeks old). Asterisk indicates a significant difference (*$P < 0.05$) between transgenic lines and WT plants.

osmotic tolerance by controlling adventitious root formation and root system development.

To investigate further the potential differences in *WOX11/12* between easy-to-root plants and difficult-to-root trees, we compared JrWOX11 to 26 WOX11/12 proteins from 15 species. In rice, *Arabidopsis,* and some easy-to-root trees, more than one WOX11/12 protein was found (Figure 1). However, "ZNS" walnut contained only one WOX11 protein, JrWOX11. In terms of protein distance matrix to JrWOX11, the descending order of similarity was: difficult-to-root trees (96.37 to 81.21%) > easy-to-root trees (74.44 to 61.48%) > *Arabidopsis* (57.01 and 54.67%) > rice (55.04 to 52.36%) > maize (49.78 and 45.22%). These results were consistent with the evolutionary relationships of the 27 WOX11/12 proteins analyzed. Furthermore, the MEME result of WOX11/12 pointed out another difference between easy-to-root trees and difficult-to-root trees, with motif10 structure (GGAIQY motif) lacking in difficult-to-root trees (Figure 1 and Supplementary Figure 2). Given that *JrWOX11* improves the capacity for adventitious root formation in transgenic "84K" poplar, a lack of the GGAIQY motif was not a determining factor for rooting difficulty in difficult-to-root trees. However, when the expression of *JrWOX11* was suppressed, the rooting capacity was

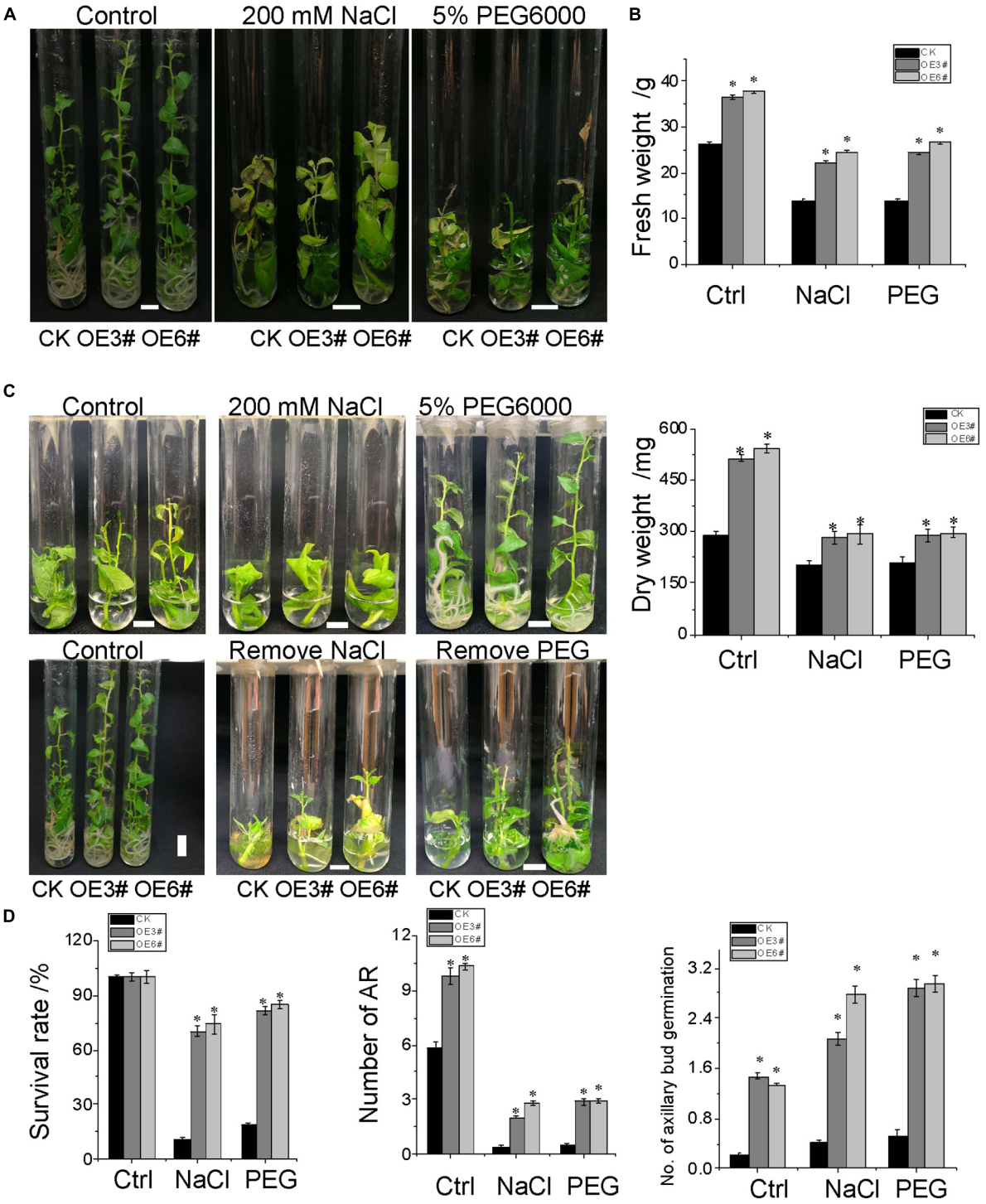


FIGURE 7
Stress response of *JrWOX11*-transgenic "84K" plants. **(A)** The growth and phenotype of *JrWOX11*-transgenic and WT poplar plants. **(B)** Effects of salt and osmotic stress on fresh weight and dry weight. Four-week-old seedlings were cultured vertically on the liquid RIM medium (1/2MS + 0.01 mg·L^{-1} of IBA + 30 g·L^{-1} of sucrose) supplemented with 200 mM NaCl or 5% PEG6000 for 2 weeks. **(C)** The photographs were taken, and the fresh weight and dry weight of each plant were measured. **(D)** The survival rate and regeneration response were observed when NaCl and PEG were removed. About 2.5 cm shoot cuttings were cultured vertically on liquid RIM medium supplemented with 200 mM NaCl or 5% PEG6000 for 4 weeks, then were transferred to liquid RIM medium without NaCl and PEG6000 for 2 weeks. The photographs were taken, the survival rate and the number of axillary bud regeneration, and the number of ARs of each plant were measured. All the above data were represented as the mean ± SD of at least three independent replications. Asterisk indicates a significant difference (*$P < 0.05$) between transgenic lines and WT plants.

significantly diminished (Chang et al., 2022). Age and associated sclerenchyma in difficult-to-root trees should also be considered (Sun and Zhu, 2020). It was reported that *WOX11/WOX12* could be suppressed by age-regulated *ETHYLENE INSENSITIVE 3* (*EIN3*) in *Arabidopsis* (Li et al., 2020). In walnut, rejuvenation diminished the extent of sclerenchyma, which is beneficial to auxin transport and rooting (Liu H. et al., 2018).

We also analyzed the differences in the *cis*-element of *WOX11/12* genes between easy-to-root trees and difficult-to-root trees. Compared with OsWOX11, OsWOX12a, OsWOX12b, AtWOX11, AtWOX12, PtrWOX11/12a, and PtrWOX11/12b, the promoter of *JrWOX11* in "ZNS" walnut and *J. regia* (Chang et al., 2020) harbored more AuxRE and ABREs but fewer MBS, G-box, W-box, and DRE motifs (**Figure 2** and **Supplementary Figure 4**), which are generally known as elements that respond to abiotic stresses (Shi et al., 2011; Lee et al., 2013; Baruah et al., 2021). The result might imply that *JrWOX11* enhanced the stress tolerance of transgenic "84K" poplar *via* the ABA-dependent pathways (Putterill et al., 2021). Further work is needed to test this suggestion.

Conclusion

In the present study, we successfully cloned full-length cDNA and genomic sequences of *JrWOX11* from "ZNS" walnut; its expression was induced by ABA, NaCl, and PEG treatment. Compared to easy-to-root plants, walnut harbored fewer *WOX11/12* genes, the JrWOX11 protein sequence lacked the "GGAIKY" motif, and the promoter of JrWOX11 harbored more AuxRE and ABREs but fewer MBS, G-box, W-box, and DRE motifs. Additionally, the overexpression of *JrWOX11* significantly improved the adventitious root formation capacity, increased the number of AR and LR, and elongated the LR and RH lengths in "84K" poplar overexpressing *JrWOX11*. Moreover, the tolerance to NaCl and osmotic stress in transgenic "84K" poplar plants and shoot cuttings was also enhanced. This likely reflects the fact that *JrWOX11* promotes *de novo* root organogenesis and root system development, thereby regulating NaCl and osmotic tolerance. The differences in WOX11/12 between easy-to-root plants and difficult-to-root trees provide the foundation for the molecular breeding of difficult-to-root trees.

Data availability statement

The data presented in this study are deposited in the GenBank repository, accession number: ON979687.

Author contributions

DP and QZ contributed substantially to the experimental design, conceived the study, and approved the final manuscript. XS contributed to funding, interpreted the data, authored or reviewed the draft, and revised the manuscript. ML gave substantial suggestions to the manuscript. YC, PZ, and XL carried out mainly experiments, finished the draft of the manuscript, and comprehensively analyzed the data from all experimental results. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by the National Natural Science Foundation of China (Grant No. 32101479), the State Key Laboratory of Tree Genetics and Breeding program (Grant No. CAFYBB2019ZY001), and the Postdoctoral Initiative Foundation of Henan Normal University (Grant No. 5101049470215).

Acknowledgments

We are grateful to the following investigators for helping with the tissue culture: Ms. Jiaqi Zhang and Bi Liang at Zhejiang A& F University. We would like to express our gratitude to EditSprings (https://www.editsprings.cn) for the expert linguistic services provided.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.951737/full#supplementary-material

SUPPLEMENTARY FIGURE 1
The detection of transgenic 84K poplar. **(a)** DNA detection of *HPT* gene in WT and transgenic line OE1#∼OE14 #. **(b)** The expression of the *JrWOX11* gene in stems of WT and transgenic line OE1#∼OE14 #5 microshoots.

SUPPLEMENTARY FIGURE 2
Sequence alignment of homologous WOX11/12 proteins. **(a)** Phylogenetic tree based on full-length coding sequences of WOX11/12 genes using the NJ method. **(b)** Motifs analysis of WOX11/12 proteins. MEME software was used to search motifs, and then redraw the map with the software TBtools **(c)**. *Zm*, *Zea mays*; *Os*, *Oryza sativa*; *At*, *Arabidopsis thaliana*; *Ptr*, *Populus trichocarpa*; *Pto*, *Populus tomentosa*; *Pe*, *P. deltoides* × *P. euramericana* cv. "Nan-lin895," *Pag*, *P. alba* × *P. glandulosa* cv. "84K," *Sp*, *Salix purpurea; Md*, *Malus domestica; Mn*, *Morus notabilis; Dz*, *Durio zibethinus; Jr*, *Juglans hindsii* × *J. regia* cv. "ZNS," *Ci*, *Carya illinoinensis; Qs*, *Quercus suber; Ql*, *Quercus lobuta; Cm*, *Castanea mollissima*.

SUPPLEMENTARY FIGURE 3
Plant phenotype determination during NaCl and PEG treatment. **(a)** Shoot cuttings of CK, OE3#, and OE6 # were cultured vertically on the liquid RIM medium (1/2MS + 0.01 mg·L^{-1} of IBA + 30 g·L^{-1} of sucrose) as Control. **(b,c)** Shoot cuttings of CK, OE3#, and OE6 were cultured vertically on liquid RIM medium supplemented with 100 and 200 mM NaCl. **(d,e)** Shoot cuttings of CK, OE3#, and OE6 were cultured vertically on liquid RIM medium supplemented with 5 and 10% PEG6000. Photographic pictures were made once a week. 1∼4 w: 1∼4 weeks after the shoot cuttings were cultured on the corresponding medium. R1W∼R2W: 1∼4 weeks after the corresponding stress is removed.

SUPPLEMENTARY FIGURE 4
Promoter sequence alignment of *WOX11* gene in "ZNS" walnut and *Juglans regia*.

References

Amissah, J. N., Paolillo, D. J., and Bassuk, N. (2008). Adventitious root formation in stem cuttings of *Quercus bicolor* and *Quercus macrocarpa* and its relationship to stem anatomy. *J. Am. Soc. Hortic. Sci.* 133, 479–486. doi: 10.1117/12.665960

Baruah, P. M., Kashyap, P., Krishnatreya, D. B., Bordoloi, K. S., and Agarwala, N. (2021). Identification and functional analysis of drought responsive lncRNAs in tea plant. *Plant Gene* 27:100311. doi: 10.1016/j.plgene.2021.100311

Basheer-Salimia, R. (2007). Juvenility, Maturity, and rejuvenation in woody plants. *Hebron Univ. Res. J.* 3, 17–43.

Chang, Y., Song, X., Zhang, Q., Liu, H., Bai, Y., Lei, X., et al. (2020). Genome-wide identification of *WOX* gene family and expression analysis during rejuvenational rhizogenesis in walnut (*Juglans regia* L.). *Forests* 11:16. doi: 10.3390/f11010016

Chang, Y., Song, X., Zhang, Q., Zhang, P., Lei, X., and Pei, D. (2022). Robust CRISPR/Cas9 mediated gene editing of *JrWOX11* manipulated adventitious rooting and vegetative growth in a nut tree species of walnut. *Sci. Hortic.* 303:111199. doi: 10.1016/j.scienta.2022.111199

Chen, C., Chen, H., Zhang, Y., Thomas, H., Frank, M., He, Y., et al. (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant.* 13, 1194–1202. doi: 10.1016/j.molp.2020.06

Chen, L., Tong, J., Xiao, L., Ruan, Y., Liu, J., Zeng, M., et al. (2016). YUCCA-mediated auxin biogenesis is required for cell fate transition occurring during *de novo* root organogenesis in *Arabidopsis*. *J. Exp. Bot.* 67, 4273–4284. doi: 10.1093/jxb/erw213

Chen, X., Qu, Y., Sheng, L., Liu, J., Huang, H., and Xu, L. (2014). A simple method suitable to study *de novo* root organogenesis. *Front. Plant Sci.* 5:208. doi: 10.3389/fpls.2014.00208

Cheng, S., Huang, Y., Zhu, N., and Zhao, Y. (2014). The rice *WUSCHEL-related homeobox* genes are involved in reproductive organ development, hormone signaling and abiotic stress response. *Gene* 549, 266–274. doi: 10.1016/j.gene.2014.08.003

Cheng, S., Zhou, D.-X., and Zhao, Y. (2016). *WUSCHEL -related homeobox* gene *WOX11* increases rice drought resistance by controlling root hair formation and root system development. *Plant Signal. Behav.* 11:e1130198. doi: 10.1080/15592324.2015.1130198

Da Costa, C., de Almeida, M., Ruedell, C. M., Schwambach, J., Maraschin, F. S., and Fett-Neto, A. G. (2013). When stress and development go hand in hand: main hormonal controls of adventitious rooting in cuttings. *Front. Plant Sci.* 4:133. doi: 10.3389/fpls.2013.00133

Gonin, M., Bergougnoux, V., Nguyen, T. D., Gantet, P., and Champion, A. (2019). What makes adventitious roots? *Plants* 8:240. doi: 10.3390/plants8070240

Hu, X., and Xu, L. (2016). Transcription factors WOX11/12 directly activate WOX5/7 to promote root primordia initiation and organogenesis. *Plant Physiol.* 172, 2363–2373. doi: 10.1104/pp.16.01067

Jiang, W., Zhou, S., Zhang, Q., Song, H., Zhou, D.-X., and Zhao, Y. (2017). Transcriptional regulatory network of *WOX11* is involved in the control of crown root development, cytokinin signals, and redox in rice. *J. Exp. Bot.* 68, 2787–2798. doi: 10.1093/jxb/erx153

Klerk, G. (2002). Rooting of microcuttings:theory and practice. *Vitr. Cell. Dev. Biol. Plant* 38, 415–422. doi: 10.1079/IVP2002335

Lee, S. C., Kim, S. H., and Kim, S. R. (2013). Drought inducible *OsDhn1* promoter is activated by *OsDREB1A* and *OsDREB1D*. *J. Plant Biol.* 56, 115–121. doi: 10.1007/s12374-012-0377-3

Li, H., Yao, L., Sun, L., and Zhu, Z. (2020). ETHYLENE INSENSITIVE 3 suppresses plant *de novo* root regeneration from leaf explants and mediates age-regulated regeneration decline. *Development* 147:dev179457. doi: 10.1242/dev.179457

Li, J., Zhang, J., Jia, H., Liu, B., Sun, P., Hu, J., et al. (2017). The *WUSCHEL-related homeobox 5a* (*PtoWOX5a*) is involved in adventitious root development in poplar. *Tree Physiol.* 38, 139–153. doi: 10.1093/treephys/tpx118

Liu, B., Wang, L., Zhang, J., Li, J., Zheng, H., Chen, J., et al. (2014). *WUSCHEL-related homeobox* genes in *Populus tomentosa*: diversified expression patterns and a functional similarity in adventitious root formation. *BMC Genomics* 15:296. doi: 10.1186/1471-2164-15-296

Liu, J., Sheng, L., Xu, Y., Li, J., Yang, Z., Huang, H., et al. (2014). WOX11 and 12 are involved in the first-step cell fate transition during *de novo* root organogenesis in *Arabidopsis*. *Plant Cell* 26, 1081–1093. doi: 10.1105/tpc.114.122887

Liu, B., Zhang, J., Yang, Z., Matsui, A., Seki, M., Li, S., et al. (2018). *PtWOX11* acts as master regulator conducting the expression of key transcription factors to induce *de novo* shoot organogenesis in poplar. *Plant Mol. Biol.* 98, 389–406. doi: 10.1007/s11103-018-0786-x

Liu, H., Gao, Y., Song, X. B., Ma, Q. G., Zhang, J. P., and Pei, D. (2018). A novel rejuvenation approach to induce endohormones and improve rhizogenesis in mature *Juglans* tree. *Plant Methods* 14:13. doi: 10.1186/s13007-018-0280-0

Liu, J., Hu, X., Qin, P., Prasad, K., Hu, Y., and Xu, L. (2018). The WOX11 - LBD16 pathway promotes pluripotency acquisition in callus cells during *de novo* shoot regeneration in tissue culture. *Plant Cell Physiol.* 59, 734–743. doi: 10.1093/pcp/pcy010

Liu, W., and Xu, L. (2018). Recruitment of *IC-WOX* genes in root evolution. *Trends Plant Sci.* 23, 490–496. doi: 10.1016/j.tplants.2018.03.011

Lohmann, J. U., Hong, R. L., Hobe, M., Busch, M. A., Parcy, F., Simon, R., et al. (2001). A molecular link between stem cell regulation and floral patterning in *Arabidopsis*. *Cell* 105, 793–803. doi: 10.1016/S0092-8674(01)00384-1

Macdonald, B. (1990). *Practical Woody Plant Propagation for Nursery Growers*. Portland, OR: Timber Press.

Maynard, B. K., and Bassuk, N. L. (1990). Stockplant etiolation, shading, and stem banding effects on the auxin dose-response of rooting in softwood stem

cuttings of *Carpinus betulus* 'fastigiata'. *J. Am. Soc. Hortic. Sci.* 25:1126a. doi: 10.21273/HORTSCI.25.9.1126a

Pei, D., and Gu, R. S. (2002). Shoot rooting *in vitro* for walnut cultivars. *Sci. Silvae Sin.* 38, 32–37. doi: 10.3321/j.issn:1001-7488.2002.02.007

Pei, D., Zhang, J. P., and Xu, H. Z. (2004). *Cutting propagation method for Juglans: China, CN ZL200410009201.4.*

Pijut, P. M., Woeste, K. E., and Michler, C. H. (2011). Promotion of adventitious root formation of difficult-to-root aardwood tree species. *Hortic. Rev.* 38, 213–251. doi: 10.1002/9780470872376.ch6

Putterill, J., Robson, F., Lee, K., Simon, R., and Coupland, G. (2021). The *CONSTANS* gene of arabidopsis promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. *Cell* 80, 847–857. doi: 10.1016/0092-8674(95)90288-0

Sarkar, A. K., Luijten, M., Miyashima, S., Lenhard, M., Hashimoto, T., Nakajima, K., et al. (2007). Conserved factors regulate signalling in *Arabidopsis thaliana* shoot and root stem cell organizers. *Nature* 446, 811–814. doi: 10.1038/nature05703

Shi, J., Zhang, L., An, H., Wu, C., and Guo, X. (2011). *GhMPK16*, a novel stress-responsive group D *MAPK* gene from cotton, is involved in disease resistance and drought sensitivity. *BMC Mol. Biol.* 12:22. doi: 10.1186/1471-2199-12-22

Sun, L., and Zhu, Z. (2020). The Molecular basis of age-modulated plant *de novo* root regeneration decline in *Arabidopsis thaliana*. *Plant Cell Physiol.* 62, 3–7. doi: 10.1093/pcp/pcaa134

Wang, L., Zhang, C., Wang, Y., Wang, Y., Yang, C., Lu, M., et al. (2018). Tamarix hispida aquaporin *ThPIP2;5* confers salt and osmotic stress tolerance to transgenic *Tamarix* and *Arabidopsis*. *Environ. Exp. Bot.* 152, 158–166. doi: 10.1016/j.envexpbot.2017.05.018

Wang, L. Q., Li, Z., Wen, S. S., Wang, J. N., Zhao, S. T., and Lu, M. Z. (2020). *WUSCHEL-related homeobox* gene *PagWOX11/12a* responds to drought stress by enhancing root elongation and biomass growth in poplar. *J. Exp. Bot.* 71, 1503–1513. doi: 10.1093/jxb/erz490

Wang, L. Q., Wen, S. S., Wang, R., Wang, C., Gao, B., and Lu, M. Z. (2021). *PagWOX11/12a* activates *PagCYP736A12* gene that facilitates salt tolerance in poplar. *Plant Biotechnol. J.* 19, 2249–2260. doi: 10.1111/pbi.13653

Wei, K., Ruan, L., Wang, L., and Cheng, H. (2019). Auxin-induced adventitious root formation in nodal cuttings of *Camellia sinensis*. *Int. J. Mol. Sci.* 20:4817. doi: 10.3390/ijms20194817

Wu, L., Jie, Y., Yachao, G., Peng, Q., and Lin, X. (2018). Pivotal role of *LBD16* in root and root-like organ initiation. *Cell. Mol. Life Sci.* 75, 3329–3338. doi: 10.1007/s00018-018-2861-5

Xu, M., Xie, W., and Huang, M. (2015). Two *WUSCHEL-related homeobox* genes, *PeWOX11a* and *PeWOX11b*, are involved in adventitious root formation of poplar. *Physiol. Plant.* 155, 446–456. doi: 10.1111/ppl.12349

Zhang, G., Zhao, F., Chen, L., Pan, Y., Sun, L., Bao, N., et al. (2019). Jasmonate-mediated wound signalling promotes plant regeneration. *Nat. Plants* 5, 491–497. doi: 10.1038/s41477-019-0408-x

Zhang, J., Zhang, W., Ji, F., Qiu, J., Song, X., Bu, D., et al. (2020). A high-quality walnut genome assembly reveals extensive gene expression divergences after whole-genome duplication. *Plant Biotechnol. J.* 18, 1848–1850. doi: 10.1111/pbi.13350

Zhang, X., Zong, J., Liu, J., Yin, J., and Zhang, D. (2010). Genome-wide analysis of *WOX* gene family in rice, sorghum, maize, *Arabidopsis* and Poplar. *J. Integr. Plant Biol.* 52, 1016–1026. doi: 10.1111/j.1744-7909.2010.00982.x

Zhao, Y., Cheng, S., Song, Y., Huang, Y., Zhou, S., Liu, X., et al. (2015). The interaction between rice *ERF3* and *WOX11* promotes crown root development by regulating gene expression involved in cytokinin signaling. *Plant Cell* 27, 2469–2483. doi: 10.1105/tpc.15.00227

Zhao, Y., Hu, Y., Dai, M., Huang, L., and Zhou, D.-X. (2009). The *WUSCHEL-related homeobox* gene *WOX11* is required to activate shoot-borne crown root development in rice. *Plant Cell* 21, 736–748. doi: 10.1105/tpc.108.061655

Zhou, S., Jiang, W., Long, F., Cheng, S., Yang, W., Zhao, Y., et al. (2017). Rice homeodomain protein WOX11 recruits a histone acetyltransferase complex to establish programs of cell proliferation of crown root meristem. *Plant Cell* 29, 1088–1104. doi: 10.1105/tpc.16.00908

frontiers | Frontiers in Plant Science

TYPE Review
PUBLISHED 07 September 2022
DOI 10.3389/fpls.2022.847166

OPEN ACCESS

EDITED BY
Mukesh Jain,
Jawaharlal Nehru University, India

REVIEWED BY
Fuai Sun,
University of California, Davis, United States
Ali Raza,
Fujian Agriculture and Forestry University, China
Rohit Joshi,
Institute of Himalayan Bioresource Technology (CSIR), India

*CORRESPONDENCE
Tianzuo Wang
tzwang@ibcas.ac.cn

SPECIALTY SECTION
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

RECEIVED 01 January 2022
ACCEPTED 15 August 2022
PUBLISHED 07 September 2022

CITATION
Zhang X, Sun Y, Qiu X, Lu H, Hwang I and Wang T (2022) Tolerant mechanism of model legume plant *Medicago truncatula* to drought, salt, and cold stresses.
Front. Plant Sci. 13:847166.
doi: 10.3389/fpls.2022.847166

Tolerant mechanism of model legume plant *Medicago truncatula* to drought, salt, and cold stresses

Xiuxiu Zhang[1,2], Yu Sun[3], Xiao Qiu[4], Hai Lu[1], Inhwan Hwang[5] and Tianzuo Wang[2,6]*

[1]College of Biological Sciences and Biotechnology, Beijing Forestry University, Beijing, China, [2]State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciencess, Beijing, China, [3]Key Laboratory of Mollisols Agroecology, Northeast Institute of Geography and Agroecology, Chinese Academy of Sciencess, Changchun, China, [4]Inner Mongolia Academy of Agricultural and Animal Husbandry Sciences, Hohhot, China, [5]Department of Life Sciences, Pohang University of Science and Technology, Pohang, South Korea, [6]College of Resources and Environment, University of Chinese Academy of Sciences, Beijing, China

Legume plants produce one-third of the total yield of primary crops and are important food sources for both humans and animals worldwide. Frequent exposure to abiotic stresses, such as drought, salt, and cold, greatly limits the production of legume crops. Several morphological, physiological, and molecular studies have been conducted to characterize the response and adaptation mechanism to abiotic stresses. The tolerant mechanisms of the model legume plant *Medicago truncatula* to abiotic stresses have been extensively studied. Although many potential genes and integrated networks underlying the *M. truncatula* in responding to abiotic stresses have been identified and described, a comprehensive summary of the tolerant mechanism is lacking. In this review, we provide a comprehensive summary of the adaptive mechanism by which *M. truncatula* responds to drought, salt, and cold stress. We also discuss future research that need to be explored to improve the abiotic tolerance of legume plants.

KEYWORDS

***Medicago truncatula*, abiotic stresses, morphological regulation, physiological regulation, functional genes, transcription factors**

Introduction

The climate change products a series of environmental factors which show negative effects to plants (Farooq et al., 2022). Among the environmental constraints, drought, salt, and cold are the main abiotic stresses that influence plants' physiological and biochemical processes, ultimately reducing crop production (Rhaman et al., 2021; Farooq et al., 2022). Up to 45% of the world's farmland faces frequent water scarcity

(Rhaman et al., 2021), and 20–50% of irrigated lands are affected by salinity (Munns and Tester, 2008). Approximately 57 and 26% of the world's land and rural areas are affected by cold stress, respectively (Cramer et al., 2011). Much work has been devoted to explore the mechanism by which plants respond and adapt to abiotic stresses, and these findings have meaningful implications for improving crop production.

Legume plants are particularly important sources of food for both humans and animals. Abiotic stresses affect their growth and development. There is an eager need to clarify the mechanism by which legumes respond to abiotic stresses, and such research will aid the breeding of climate-resilient varieties. The legume model plant *M. truncatula* has small genome, short life cycle, self-pollination ability, and high genetic transformation efficiency (Tang et al., 2014). So, *M. truncatula* has been widely used in genomic, genetic, and physiological studies. Many studies focus on elucidating the mechanism by which *M. truncatula* responds and adapts to abiotic stresses. In this review, we summarize the general morphological, physiological, and molecular features by which *M. truncatula* responds and adapts to drought, salt, and cold stress. We also incorporate the crosstalk between different abiotic stresses, and discuss the implications for breeding stress-tolerant legume crops.

Drought stress

Drought stress significantly reduces leaf water potential and stomatal closure of *M. truncatula* plants, resulting in reduced photosynthesis, which in turn restricts plants' growth (Nunes et al., 2008; Luo et al., 2016). Drought also causes photooxidative damage to thylakoid membranes and reduces chlorophyll content and photosystem II activity (Luo et al., 2016). To cope with drought stress, *M. truncatula* plants have evolved various responses such as alterations in tissue architectures and expression patterns of functional genes (**Figure 1**). Indeed, 5-week-old *M. truncatula* R108 plants are still recoverable after withholding irrigation for 12 days (Luo et al., 2016).

Morphological and physiological regulation of drought tolerance

Plants' shoot architecture is affected by drought stress and can be used as an indicator of drought adaptation (Farooq et al., 2009; Nguyen et al., 2013). Branching is a key determinant of shoot architecture. In *Arabidopsis*, AtSPL9 (squamosa promoter binding protein-like) controls the initiation of cauline leaf axillary meristems (Tian et al., 2014). In transgenic *M. truncatula* plants overexpressing *MtSPL8* inhibits branching by directly suppressing axillary bud formation (Gou et al., 2018). Down-regulation *MsSPL8* increases branch density and enhances drought and salt tolerance in transgenic alfalfa (Gou et al., 2018). Interestingly, *M. truncatula* plants with up-regulated *MtRAV3* (related to ABI3/VP1 transcription factor) have the similar tolerance to drought and salt stress, but exhibiting dwarfing, late flowering, and smaller leaves (Wang S. et al., 2021). Thus, MtRAV3 and MtSPL8 might play different roles in morphological development. However, Guo et al. (2021) find that overexpression of *MtDof32* (DNA-binding one zinc finger 32) in *Arabidopsis* results in reduced branches and enlarged leaves, but still have enhanced osmotic and salt tolerance. Although both *MtDof32* and *MtRAV3* enhance drought and salt tolerance in transgenic plants, they regulate different morphological development. Indeed, *MtDof32* enhances plants' stress tolerance by regulating the rosette numbers. Thus, how to improve the shoot structure under drought condition might be a useful way to improve crops' drought tolerance.

In addition, regulation of stomatal conductance (Nunes et al., 2008) and epidermal wax accumulation (Zhang et al., 2005) is an effective way for *M. truncatula* plants to cope with drought stress. Furthermore, osmotic and oxidative regulation are also essential in the response of *M. truncatula* to drought stress (Yousfi et al., 2010; Luo et al., 2016). For instance, *M. truncatula* populations adapt to drought tolerance by accumulating solutes such as proline, soluble sugars, and K^+ (Yousfi et al., 2010). The *M. truncatula* lines overexpressing the oat arginine decarboxylase gene produce more soluble polyamines (PAs), resulting in greater drought tolerance compared to wild-type plants (Duque et al., 2016). The PAs including spermidine (Spd), spermine (Spm), and putrescine (Put) are involved in plant responses to abiotic stresses (Alcázar et al., 2010). Pagano et al. (2022) found that desiccation induces the expression of Spm synthase *MtSPMS* and Spd synthase *MtSPDS* in over-primed seeds. As for the oxidative reactions, *M. truncatula* plants that accumulate less peroxide and malondialdehyde (MDA) are more tolerant to drought stress (Luo et al., 2016; Wang S. et al., 2021).

Molecular regulation of drought tolerance

Drought-related functional genes

Zhang et al. (2014) identified many drought-responsive genes in *M. truncatula*. The genes *MtP5CS* (encoding proline synthase) and *MtProDH* (encoding proline dehydrogenase) regulate proline accumulation coordinately in response to drought stress. Indeed, overexpression of *P5CS* in *M. truncatula* results in more proline accumulation and greater drought tolerance (Verdoy et al., 2006). Heterologous expression of calcium-binding protein gene *MtCaMP1* in *Arabidopsis* induces *P5CS1* and suppresses *ProDH*, making transgenic plants more tolerant to drought stress (Wang T. Z. et al., 2013). Whereas,

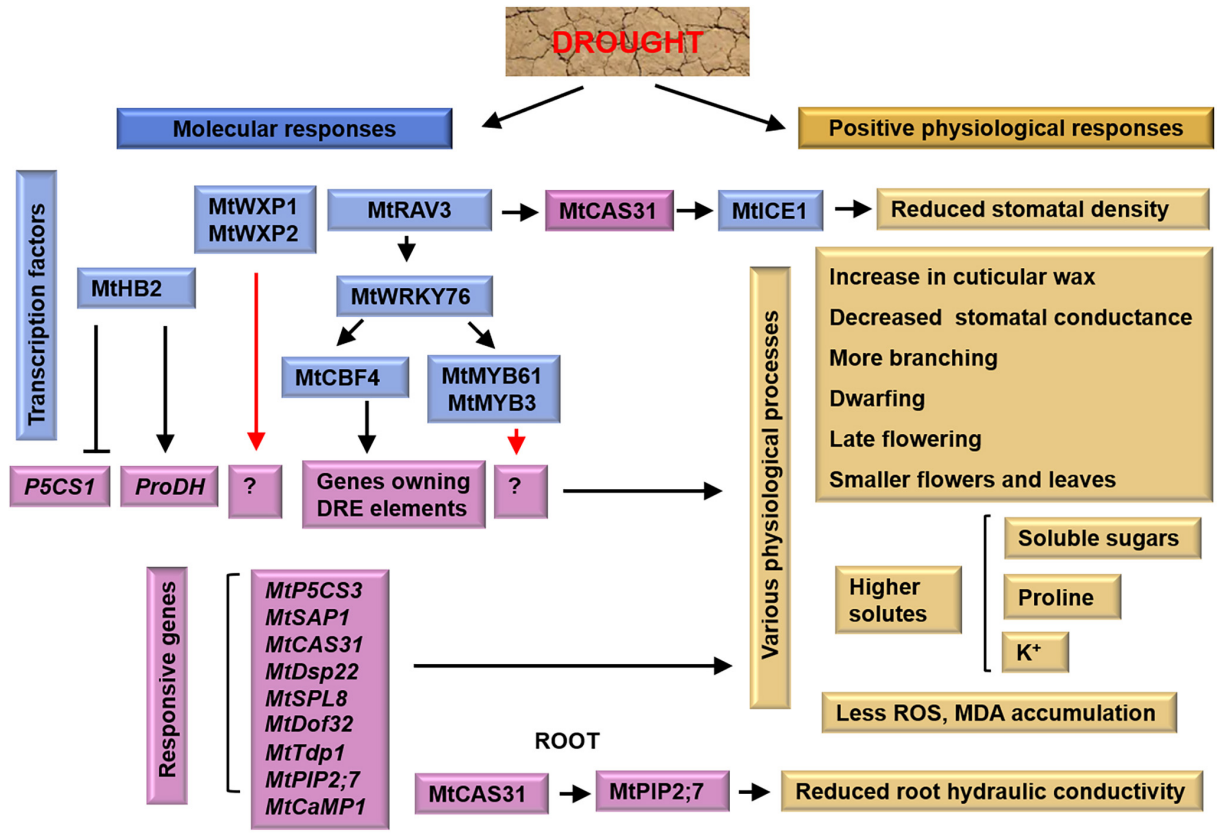


FIGURE 1
Tolerant mechanism of *M. truncatula* to drought stress. In shoots, MtRAV3 induces the expression of *MtCAS31*, whose encoded protein binds with MtICE1, resulting in reduced stomatal density. In roots, MtCAS31 facilitates the autophagic degradation of aquaporin MtPIP2; 7, and reduces root hydraulic conductivity. Blue boxes and purple boxes indicate transcription factors and responsive genes, respectively. Yellow boxes indicate physiological responses of *M. truncatula*. The black arrows represent the targets, the red arrows represent the unknown factors and the cross-lines mean suppression.

plants with *Tnt1* transposon insertion of *MtP5CS3* accumulate less proline and are sensitivity to salt and drought stresses (Nguyen et al., 2013). In addition, cold-acclimation specific protein 31 (MtCAS31), a Y_2K_4-type dehydrin, interacts with AtICE1 (inducer of CBF expression 1) to regulate stomatal development. Overexpression of *MtCAS31* in *Arabidopsis* reduces stomatal density and significantly enhances drought tolerance in transgenic plants (Xie et al., 2012). Li et al. (2018) generate the *mtcas31* mutant by transcription activator-like effector nuclease (TALEN) technology, and identify that MtCAS31 interacts with leghemoglobin MtLb120-1 to regulate drought response. Moreover, in response to drought stress, MtCAS31 promotes the autophagic degradation of the aquaporin MtPIP2; 7, thereby reducing water loss and improving drought tolerance (Li et al., 2020). Recently, 39 autophagy−related (ATG) genes are identified in *M. truncatula*. Most of them are highly induced during seed development and drought stress, indicating that autophagy plays an important role in seed development and responses to drought stress in *M. truncatula* (Yang et al., 2021).

Plants overexpressing the stress-associated protein genes *MtSAP1* accumulate more nitric oxide (NO), which is beneficial to plant growth under osmotic and salt stress (Charrier et al., 2012, 2013). In turn, NO interacts with reactive oxygen species (ROS) to affect the *SAPs*' expression (Delledonne et al., 2001; Wendehenne et al., 2004; Qiao and Fan, 2008). Of the 17 *MtSAPs*, the *MtSAP4*, *6*, *9*, *11*, *14*, and 15 are induced by drought stress (Zhou et al., 2018). In addition, Macovei et al. (2010) find that *MtTdp1*, a tyrosyl-DNA phosphodiesterase gene, is up-regulated by PEG treatment suggesting a relationship between drought response and DNA repair pathway. While, *MtTdp2α* positively regulates *M. truncatula* in salt response due to strong antioxidant effects of transgenic plants (Confalonieri et al., 2019). Recently, Pagano et al. (2022) found that desiccation treatment on over-primed seeds alters rRNA accumulation, promotes signal molecule 3′−phosphoadenosine 5′−phosphate (PAP) production, and up−regulates genes involved in ribogenesis. In addition, early light-inducible proteins (ELIPs) and ELIP-like proteins are pigment-binding components that protect against photooxidative damage (Araújo et al., 2013).

Transgenic plants overexpressing the ELIP-like gene *CpDsp22* (desiccation stress protein 22 from *Craterostigma plantagineum*) recover faster from water deficit (Araújo et al., 2013). These results provide insights into NO and nucleic acid organization in response to oxidative stress caused by drought stress in *M. truncatula*.

Drought-related transcription factors

Transcription factors (TFs) regulate the transcription of downstream genes by binding to their *cis*-elements in promoters playing important roles in response to various stresses (Oztur et al., 2002; Porto et al., 2014). For instance, the C-repeat binding factor 4 (MtCBF4), belonging to the APETALA2/EREBP (AP2-EREBP) family, binds to the dehydration responsive (DRE) element of downstream genes to regulate drought response (Li et al., 2011; **Table 1**). Overexpression of TF *MtWRKY76* in *M. truncatula* promotes the expression of *MtCAS31*, *MtCBF4*, *MtMYB61*, and *MtMYB3*, and enhances drought tolerance in transgenic plants (Liu et al., 2016). Meanwhile, TF *MtRAV3* up-regulates the expression of *MtWRKY76*, *MtMYB61*, *MtCAS31*, *MtAOX1*, and *MtERF1* (Wang S. et al., 2021). In addition, the ethylene response factor (ERF) TFs MtWXP1 and MtWXP2 mediate cuticular wax production. Overexpression of *MtWXP1* and *MtWXP2* enhances transgenic plants' drought tolerance (Zhang et al., 2005, 2007). These two wax genes are expected to have great potential for crop improvement through genetic modification. While, TF MtHB2 is a homeodomain leucine zipper (HD-Zip) protein that negatively regulates drought stress by affecting osmotic and oxidative responses (Song et al., 2012). Li et al. (2022) identifies 15 *HD-ZIP ?* genes in *M. truncatula*. In particular, *MtHB7* and *MtHB12* are positively associated with salt, osmotic stress, and abscisic acid (ABA), while *MtHB13* and *MtHB23* are negatively associated with these stresses. This genome-wide analysis of the HD-ZIP I TFs in *M. truncatula* provides valuable references for further research.

Drought-related plant growth regulators

Plant growth regulators (PGRs) such as auxin, ABA, and ethylene regulate plants in response to abiotic stresses (Rhaman et al., 2021). Both PEG and ABA treatment induces the expression of 9-*cis*-epoxycarotenoid dioxygenase gene *NCED5* leads to increased endogenous ABA content in *M. truncatula* (Planchet et al., 2011; Luo et al., 2016). Meanwhile, water deficit induces endogenous NO accumulation through an ABA-dependent pathway (Planchet et al., 2014). While, exogenous ABA addition induces asparagine and proline production contributing to osmotic adjustment under water deficit (Planchet et al., 2011). However, the modulation of proline metabolism is independent of NO production under water deficit (Planchet et al., 2014). So, exploring the central role of ABA in water-deficit tolerance could lead us to obtain more information on osmotic adjustment and nitrogen metabolism under adverse conditions.

Salt stress

Salt stress causes osmotic stress, ion toxicity, and oxidative damage to *M. truncatula* plants, resulting in reduced photosynthesis and biomass (Yousfi et al., 2010; Arraouadi et al., 2012; Luo et al., 2016; Gou et al., 2018; Zhang X. X. et al., 2019; Wang S. et al., 2021). *M. truncatula* minimizes these damages by regulating the production of osmolytes and antioxidants in cells, the extrusion of Na^+ out of cells, and the reduction of Na^+ in leaves (**Figure 2**). In fact, hydroponic *M. truncatula* R108 can tolerate 100 mM NaCl for nearly 1 week (Merchan et al., 2003; Zhang X. X. et al., 2019;). *M. truncatula* genotype TN1.11 is the most tolerant to salt stress among R108, Jemalong A17, TN6.18, and TN1.11.

Morphological and physiological regulation of salt tolerance

The root architecture of plants is affected by salt stress. de Zélicourt et al. (2012) find that shorter and less branched roots are beneficial for preventing Na^+ uptake in *M. truncatula* (Ariel et al., 2010).

Multiple studies show that *M. truncatula* plants with high drought tolerance also display strong salt tolerance, suggesting some overlapping mechanism between them (Yousfi et al., 2010; Luo et al., 2016; Gou et al., 2018; Wang S. et al., 2021). Both salt and drought responses involve osmotic and oxidative regulation. López et al. (2008) find that accumulation of amino acids and sugars in shoot cells alleviates the adverse effects of Na^+ in *M. truncatula*. Salt-tolerant *M. truncatula* genotypes accumulate more antioxidants and have strong peroxide scavenging ability (Mhadhbi et al., 2011, 2013; Amouri et al., 2018). In rice seedlings, PAs regulate cell membrane stability as ROS scavengers and antioxidants under salt stress (Ghosh et al., 2012). The sensitive *M. truncatula* cultivar TN6.18 has a lower (Spd + Spm)/Put ratio, indicating that this ratio may be related to oxidative status (Antoniou et al., 2021). Nevertheless, the PA levels are easily affected by plants' condition and environment. So, the application of (Spd + Spm)/Put ratio and PA levels to asses salt tolerance in plants needs to be determined.

Salt stress causes ion toxicity in plant cells. Greater Na^+ accumulation in *M. truncatula* leads to smaller root and shoot structures (Arraouadi et al., 2012; Zhang X. X. et al., 2019). When *M. falcata* and *M. truncatula* are subjected to salt shock, *M. falcata* shows stronger tolerance for its effective extrusion of Na^+ out of cells (Liu et al., 2015). This result is in agreement with that *M. truncatula* lines with the highest salt stress tolerance have the lowest Na^+ content in leaves (Aydi et al., 2008). In addition, legumes tend to restrict Na^+ transport toward the shoots to keep a relatively low Na^+ content in their photosynthetic organs (Winter and Läuchli, 1982). Transgenic *M. truncatula* lines overexpressing the calmodulin-like gene *MtCML40* are more

sensitive to salt stress because of the greater Na^+ accumulation in their shoots (Zhang X. X. et al., 2019).

Molecular regulation of salt tolerance

Salt-related functional genes

Several salt-responsive genes throughout leaf senescence are identified in *M. truncatula* (Dong et al., 2021). These genes are mainly related to protein and amino acid metabolism, photosynthesis, chlorophyll metabolism, and hormone signaling. Long et al. (2016, 2018) characterize the proteome-level changes associated with the salt stress response, which are consistent with previous studies (Kang et al., 2010; Dong et al., 2021). Hence, *M. truncatula* responds to salt stress by altering gene expression, biosynthesis of proteins and metabolites, and modifications in hormonal signaling, etc.

Several studies show that proline is related to the regulation of salt stress. *Arabidopsis* plants expressing *MtHB2* are susceptible to salt stress due to lower proline and soluble sugar content in the cells. This is because MtHB2 may bind to the promoters of *P5CS1* and *P5CS2* to inhibit their expression (Song et al., 2012). Besides, De Lorenzo et al. (2009) identify a salt-induced gene *Srlk* in *M. truncatula* and RNA interference (RNAi) created *Srlk* mutants accumulate less Na^+ in plants than in wild-type plants. Liu et al. (2015) find that the SOS (salt overly sensitive) system mediates cytosolic Na^+ out of cells. The protein MtCaMP1 up-regulates the vacuolar Na^+/H^+ antiporter *AtNHX1* and reduces Na^+ content in transgenic *Arabidopsis* plants (Wang T. Z. et al., 2013). The NHX transporters sequester Na^+ into vacuoles and decrease the Na^+ damage to the organelles in the cytoplasm. Four distinct NHX isoforms (AtNHX1–AtNHX4) are confirmed in Arabidopsis, and their roles in vacuolar ion and pH homeostasis have been determined (Bassil et al., 2019). In *M. truncatula*, six *MtNHX*s are identified, and *MtNHX3*, *MtNHX6*, and *MtNHX7* in roots are induced by salt stress (Sandhu et al., 2018). In addition, Du et al. (2021) identify several salt stress responsive *CBL-CIPK* genes in *M. truncatula* and *M. sativa*. Collectively, *SOS* pathway,

TABLE 1 TFs of *M. truncatula* involved in drought, salt, and cold stress.

Family	Transcription factors	Downstream genes	Stress	References
AP2/EREBP	MtCBF1	Unknown	Cold	Pennycooke et al., 2008; Zhang et al., 2011; Sun et al., 2021
	MtCBF2	Unknown	Cold	
	MtCBF3	Unknown	Cold	
	MtCBF4	*MtCAS15, MtCOR15A, MtCOR15B, MtKIN1, MtRD17, MtRD29A, MtRD29B*	Drought, salt, cold	Li et al., 2011; Zhang et al., 2016; Sun et al., 2021
	MtDREB1C	Unknown	Cold	Chen et al., 2010
	MtWXP1	Unknown	Drought, cold	Zhang et al., 2007
	MtWXP2	Unknown	Drought, cold	
MYB	MtMYB3	*MtCBF4*	Cold, drought	Zhang et al., 2016
	MtMYB61	*MtMYB3*	Cold, drought	
	MtMYBS1	*AtP5CS*	Salt	Dong et al., 2017
	MtMYB634	Unknown	Salt	Gruber et al., 2009
	MtMYB636	Unknown	Salt	
	MtMYB119	Unknown	Salt	
	MtMYB1070	Unknown	Salt	
bHLH	MtbHLH-658	Unknown	Salt	Zahaf et al., 2012
WRKY	MtWRKY76	Unknown	Drought, salt	Liu et al., 2015
TFIIIA-like	MtZpt2-1	*MtCorA1, MtFpf1, MtPrp2*	Salt	Merchan et al., 2007
	MtZpt2-2	Unknown	Salt	
HD-ZIP	MtHB1	*MtLBD1*	Salt	Ariel et al., 2010
	MtHB2	*AtP5CS1, AtProDH*	Drought, salt	Song et al., 2012
	MtHB7, MtHB12, MtHB13, MtHB23	Unknown	Drought, salt	Li et al., 2022
NAC	MtNAC969	Unknown	Salt	de Zélicourt et al., 2012
RAV	MtRAV3	*MtWRKY76, MtMYB61, MtCAS31, MtAOX1, MtERF1*	Drought, salt	Wang S. et al., 2021

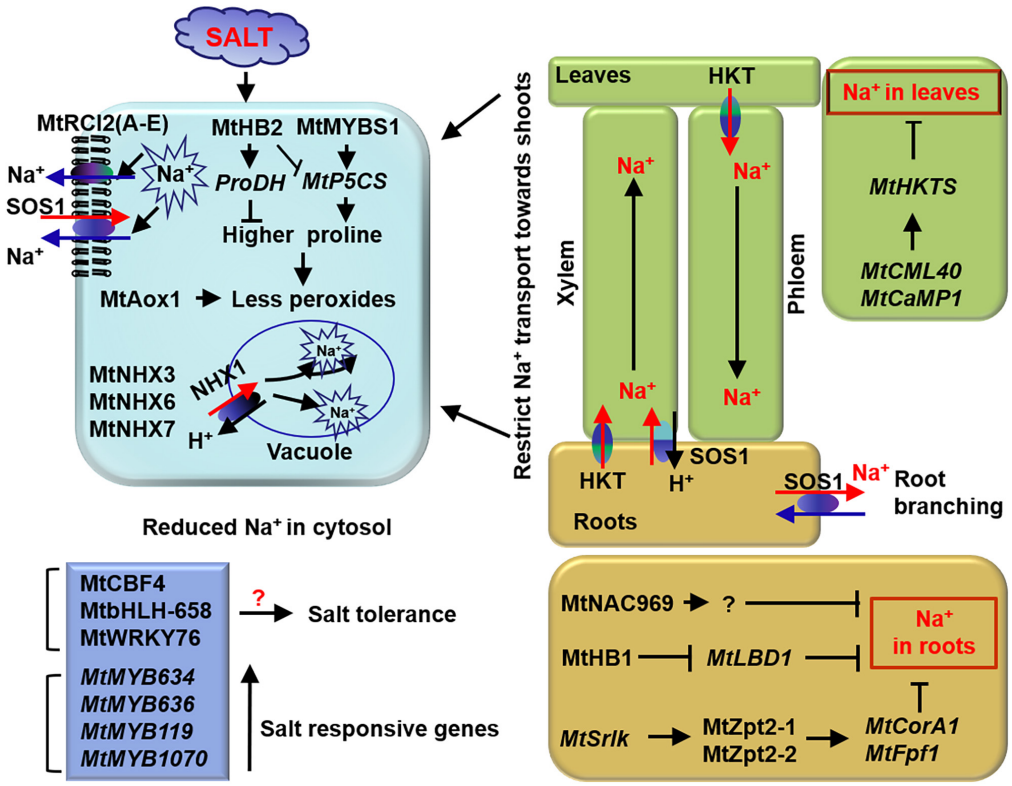


FIGURE 2
Tolerant mechanism of *M. truncatula* to salt stress. MtHB2 and MtMYBS1 and oxidase MtAox1 are involved in osmotic regulation. Membrane proteins MtRCI2 (A–E) and Na^+/H^+ exchangers MtNHX3, MtNHX6, and MtNHX7 play roles in reducing Na^+ accumulation in cells. MtNAC969, MtHB1, MtZpt2-1, and MtZpt2-1 play roles in reducing Na^+ accumulation in roots. While, Ca^{2+} sensor MtCML40 and MtCaMP1 participate in Na^+ long-distance transportation regulation between roots and shoots. Light blue box represents cells. Blue boxes represent transcription factors related to salt stress. Green and yellow boxes mean shoots and roots, respectively. The black arrows represent the targets or directions, the red arrows represent transport of Na^+ and the cross-lines mean suppression.

CBL-CIPK family genes, and *NHX* genes play crucial roles in response to salt stress.

Salt-related transcription factors

Gruber et al. (2009) identify many salt-responsive TFs in *M. truncatula* roots belonging to AP2/EREBP, HD-ZIP, and MYB families (**Table 1**). TF MtMYBS1 promotes the expression of *P5CS* and mitigates the restriction of root growth under salt stress (Dong et al., 2017). TF MtHB1 suppresses the expression of *MtLBD1* (lateral organ boundaries gene), reducing lateral roots formation and Na^+ uptake (Ariel et al., 2010). Furthermore, overexpression of *MtNAC969* induces the formation of shorter and less branched roots, whereas RNAi-mediated *MtNAC969* inactivation promotes lateral root formation. Interestingly, both root systems improved plant growth under salt stress (de Zélicourt et al., 2012). This discrepancy might because that MtNAC969 might participate in multiple pathways controlling root system adaptation to salt stress. In addition, overexpression of *MtbHLH-658*, *MtRAV3*, and *MtWRKY76* improves root growth under salt stress in transgenic plants (Zahaf et al., 2012; Liu et al., 2016; Wang S. et al., 2021). Merchan et al. (2007) identify two salt responsive IIIA-like TFs MtZpt2-1 and MtZpt2-2. Overexpression each of them significantly improves root growth under salt stress (De Lorenzo et al., 2007). There exists many stress-related *cis*-elements in *MtZpt2-1*, allowing it to respond and adapt to abiotic stresses (Wang T. Z. et al., 2014). However, the target genes for most of these TFs have not been identified.

Salt-related epigenetic regulation

Epigenetic modifications play important "switch" roles in regulating gene expression, thereby affecting plant responses to abiotic stresses (Dong et al., 2018). The epigenetics refers to alterations in gene expressions caused by DNA methylation and histone modification (Saeed et al., 2022). Yaish et al. (2018) analyze the *M. truncatula* genome-wide DNA methylation in response to salt stress and find that the whole DNA methylation level is increased, and the 5-methylcytosine nucleotide (5-mC) landscape is remodeled under salt stress. More precisely, the DNA methylation and histone modification of *MtMYBS1* are

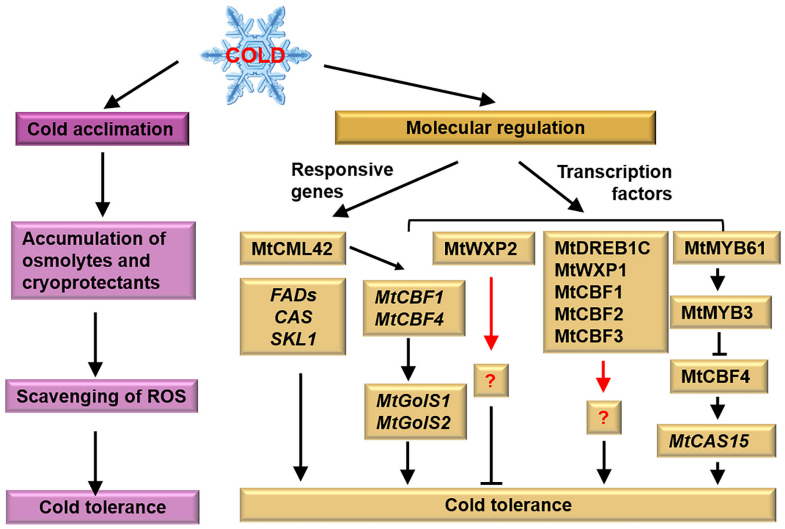


FIGURE 3
Tolerant mechanism of *M. truncatula* to cold stress. *M. truncatula* enhances its cold tolerance after cold acclimation. Meanwhile, cold responsive genes and transcription factors associated with cold tolerance have been identified. Pink and yellow boxes represent physiological and molecular responses, respectively. The black arrows represent the targets, the red arrows represent the unknown factors and the cross-lines mean suppression.

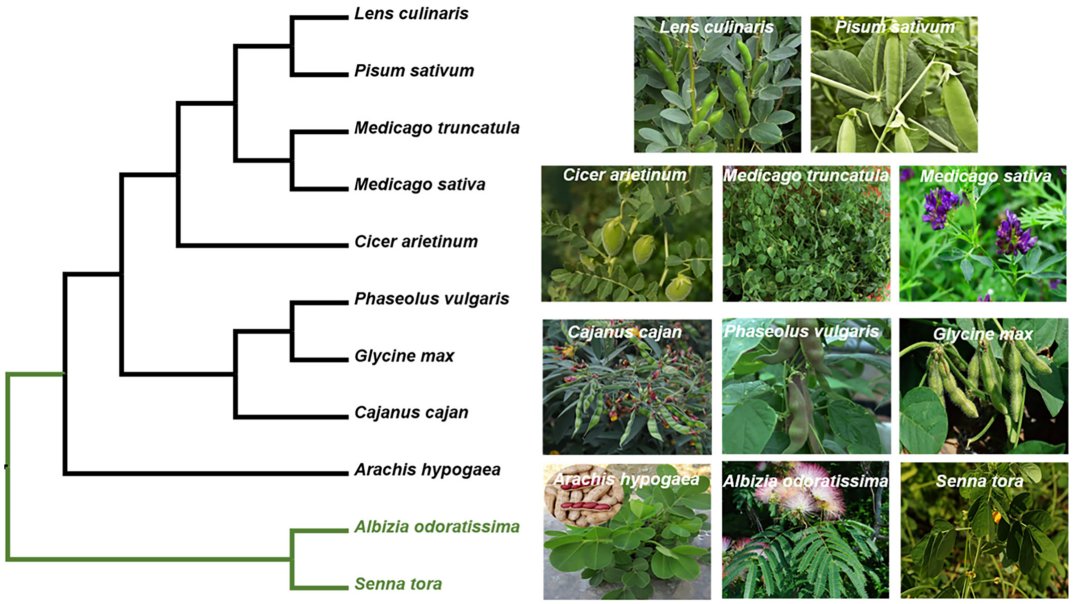


FIGURE 4
Phylogeny of *M. truncatula* with the other legumes. The phylogeny of the targeted species was reconstructed based on their plastomes. The data matrix of Azani et al. (2017) was used as a basic DNA matrix, from which we extract the sequences of the targeted species and outgroups (*Albizia odoratissima* and *Senna tora*). Then, these sequences were aligned with the complete plastomes of *M. truncatula*. The maximum likelihood (ML) phylogeny was reconstructed using RAxML version 8.2.12. Their accession number obtained from GenBank. *Albizia odoratissima*: NC_034987.1; *Arachis hypogaea*: NC_026676.1; *Cajanus cajan*: NC_031429.1; *Cicer arietinum*: NC_011163.1; *Glycine max*: NC_007942.1; *Lens culinaris*: NC_027152.1; *Medicago sativa*: KU321683.1; *Medicago truncatula*: JX512024.1; *Phaseolus vulgaris*: NC_009259.1; *Pisum sativum*: NC_014057.1; *Senna tora*: NC_030193.1.

analyzed under salt stress. Indeed, the expression of *MtMYBS1* is negatively correlated with its DNA methylation modification, and positively correlated with histone H3K9ac modification under salt stress (Dong et al., 2018). These studies provide critical theoretical guidance for further understanding of epigenetic regulation in response to salt stress in *M. truncatula*.

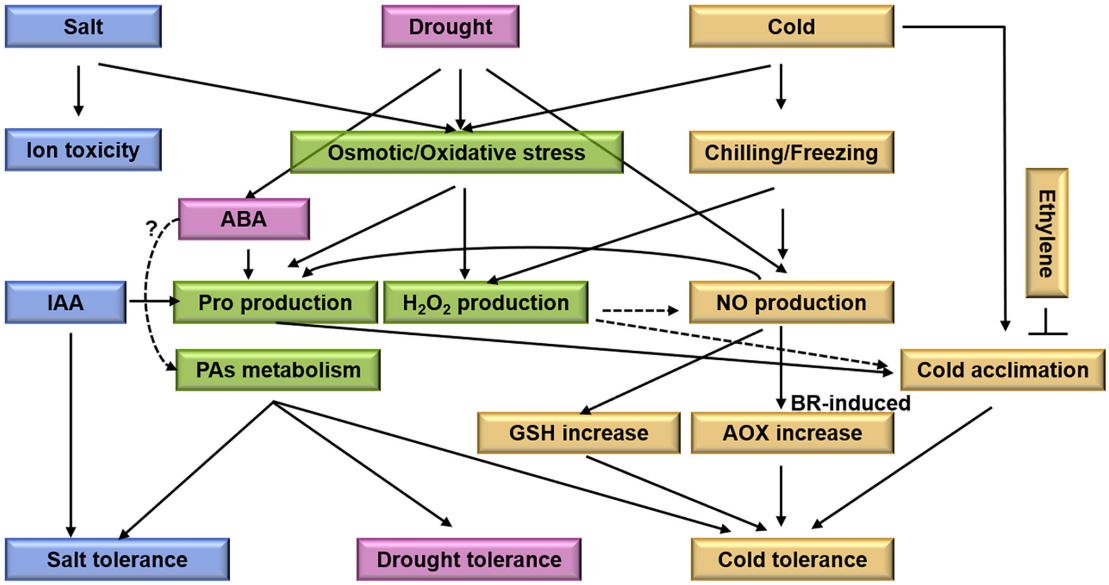


FIGURE 5
A schematic representation of cross-talks between drought, salt, and cold stress in *M. truncatula*. Drought, salt, and cold stress cause secondary stress including osmotic stress and oxidative stress. And then induce proline and H_2O_2 production. In *M. truncatula*, ABA could induce proline accumulation, contributing to osmotic adjustment under drought condition. And NO in cold acclimation through regulation on GSH synthesis and redox is associated with the differential cold tolerance. Meanwhile, ethylene reduces cold acclimation-induced freezing tolerance. IAA may have positive effects on their development and differentiation under salt stress. PAs as compound positively regulate *M. truncatula* in response to drought, salt, and cold stress. Ion toxicity and cold acclimation are specific process for salt stress and cold stress, respectively. Blue boxes indicate salt related elements. Pink and yellow boxes indicate drought and cold related elements, respectively. Green boxes indicate the central elements between drought, salt, and cold stress. The black arrows represent the targets, the dotted lines represent the possible regulatory targets and the cross-lines mean suppression.

Salt-related plant growth regulators

Bianco and Defez (2009) compared *Mt-RD64* plants noduled by *Sinorhizobium meliloti* RD64, which have higher indole-3-acetic acid (IAA) content in nodules and roots, with the control plants. The results show that *Mt-RD64* plants accumulate higher endogenous osmolyte in shoots and are more tolerant to salt stress (Bianco and Defez, 2009). Thus, exogenous IAA might be able to stimulate osmolyte production and positively affect plant development and differentiation under salt stress.

Cold stress

Cold stress includes chilling stress and freezing stress. When the temperature is low but above 0°C (i.e., chilling), membrane fluidity decreases. When it is below 0°C and is defined as freezing, ice formation might occur within tissues, resulting in membrane damage (Raza et al., 2021). The freezing tolerance of many plants is increased after exposure to low, non-freezing temperatures, which is referred to as cold acclimation (Xin and Browse, 2000; **Figure 3**). Cold acclimated *M. truncatula* A17 seedlings exposed to −10°C are still survivable (Zhang et al., 2011). In molecular terms, the cold responsive genes and CBF-dependent signaling pathways play roles to enhance the cold tolerance of *M. truncatula* (**Figure 3**).

Morphological and physiological regulation of cold tolerance

Cold-acclimated *M. truncatula* exhibits shorter stems, fewer leaves, smaller tissues, and higher root/shoot ratio compared to plants under normal condition (Thapa et al., 2008). While, the relationship between these phenotypes and cold tolerance remains vague. Pennycooke et al. (2008) find that cold acclimation does not significantly improve the survival rate of *M. truncatula* under freezing stress. However, Thapa et al. (2008) and Zhang et al. (2011) demonstrate that cold acclimation is able to improve the freezing tolerance of *M. truncatula*. The different cold acclimation regimes, 4°C in Zhang et al. (2011) and 2°C in Pennycooke et al. (2008), and different cultivars and ages of seedlings may explain this discrepancy. Thapa et al. (2008) propose that 3.5°C day/−1°C night for 1 week is the best regime for cold acclimation in *M. truncatula*. Cold acclimation induces the accumulation of sucrose and proline in *M. falcata* and *M. truncatula* (Zhang et al., 2011). Although

MtP5CS3 is induced at 4°C, it is unlikely associated with cold tolerance (Nguyen et al., 2013). These data indicate that the cold tolerance might be independent with the proline concentration but positively correlated with the soluble sugar concentration.

Molecular regulation of cold tolerance

Cold-related functional genes

Zhang et al. (2018) figure out 20 *MtFAD* genes (fatty acid desaturase genes) involved in chilling response. The *FADs* are involved in the desaturation of fatty acids affecting the function of the membrane system (Wallis and Browse, 2002). Another important agent of cold-related genes is the cold-responsive (*COR*) genes. Mohapatra et al. (1989) isolate three *COR*s specifically expressed during cold acclimation in *Medicago* referred to as *CAS* (cold acclimation specific). The *CAS18* gene of *M. falcata* is positively correlated with freezing tolerance and its expression is much higher in cold-acclimated plants than in non-acclimated plants (Wolfraim et al., 1993). Pennycooke et al. (2008) find that the *M. truncatula* genome contains a single *CAS31* gene, whereas the *M. falcata* genome contains multiple *MfCAS30* and *MfCAS31* genes. So, *M. falcata* is more tolerant to cold stress than *M. truncatula*. Zhang et al. (2011) clarify that *MfCAS17* and *MfCAS18* contribute to the stronger cold acclimation effects on *M. alfalfa* than on *M. truncatula*. Zhao et al. (2014) find that cold acclimation—induced the transcription of *MtCAS15* is suppressed in the ethylene-insensitive mutant *skl*, indicating that *MtSKL1* is required for cold acclimation.

Cold-related transcription factors

Recently, the expression profiles of *DREBs* in *M. truncatula* and *M. sativa* are identified in the cold-stress response (Shu et al., 2016; Sheng et al., 2022). A cluster of *DREB* subfamily members on *M. truncatula* chromosome 6 is induced by both cold and freezing stress (Shu et al., 2016), and 33 *MsDREBs* are significantly upregulated by cold treatment (Sheng et al., 2022). The genome-wide identification of *DREBs* in *Medicago* species provides promising molecular targets for the improvement of cold tolerance in crops. Overexpression of *MtDREB1C/MtCBF3* inhibits shoot growth and enhances the freezing tolerance of *M. truncatula* (Chen et al., 2010). Transgenic *M. truncatula* plants overexpressing *MfERF1* show enhanced tolerance to both freezing and chilling stress through promoting PA turnover, antioxidant protection, and proline accumulation (Zhuo et al., 2018). Overexpression of *WXP1* in *M. truncatula* enhances the plants' freezing tolerance without altering growth and development. However, plants overexpressing *WXP2* are more sensitive to freezing (Zhang et al., 2007). These results indicate that *WXP1* is a useful candidate gene for improving plant freezing tolerance by genetic conduction.

The TFs MtCBF1, MtCBF2, and MtCBF3 have been shown to participate in cold acclimation in *M. truncatula* (Pennycooke et al., 2008; Zhang et al., 2011). TF MtCBF4 not only positively regulates cold acclimation and freezing tolerance but also enhances drought and salt tolerance (Li et al., 2011; Zhang et al., 2016). Although the differential response of MtCBFs to cold stress is unknown, the major components involved in CBF-dependent signaling pathways are illustrated under cold stress. TF MtMYB3 binds to the *cis*-elements of *MtCBF4* promoter and represses its expression. TF MtCBF4 directly activates the transcription of *MtCAS15*. TF MtMYB61 relieve the inhibitory effect of MtMYB3 on *MtCBF4* (Zhang et al., 2016). Besides, Qu et al. (2016) indicates that the MfNAC3 plays roles in response to cold stress by regulating the expression of *MtCBF4*. Recently, Sun et al. (2021) identifies that MtCML42 positively regulates the expression of *MtCBF1* and *MtCBF4*, thereby upregulating the expression of the *COR* genes, *MtGolS1* and *MtGolS2*, and leads to raffinose accumulation and improved cold tolerance.

Cold-related epigenetic regulation

Demethylases containing Jumonji (JMJ) C domain are involved in removal of methyl groups at lysine or arginine residues (Lu et al., 2008). In *M. truncatula*, *MtJMJC5* undergoes cold-specifically induced alternative splicing, which is reversible depending on temperature (Shen et al., 2016). Previous studies show that AtJMJ30/JMJD5 is a component of the plant circadian clock (Lu et al., 2011). So, there may exist a *MtJMJC5*-dependent link between the circadian clock and ambient temperature fluctuation in *M. truncatula*.

Cold-related plant growth regulators

Zhao et al. (2009) proposes that nitrate reductase (NR)-dependent NO production plays an important role in the cold acclimation-induced increase in freezing tolerance by modulating proline accumulation in *Arabidopsis*. In addition, the role of NO in cold acclimation through the regulation of glutathione (GSH) synthesis has been studied in *M. falcata* and *M. truncatula* (Zhang P. P. et al., 2019). Exogenous application of ethylene reduces cold acclimation-induced freezing tolerance (Zhao et al., 2014). These results indicate that there may have some relationships between NO and ethylene molecules and osmotic regulation in response to cold tolerance.

Conclusion and future perspectives

Legumes are a particularly important source of food for livestock worldwide (Wang T. Z. et al., 2021). The conventional breeding of crops is time-consuming, labor-intensive, and cost-inefficient (Mishra et al., 2021). An efficient solution is to generate stress-tolerant varieties with the help of information obtained in the lab. Thus, understanding the physiological and

molecular processes of legumes in response to abiotic stresses is really important. *M. truncatula* is closely related to many legumes and forages (**Figure 4**). In this review, we summarize the mechanism by which *M. truncatula* responds and adapts to drought (**Figure 1**), salt (**Figure 2**), and cold stress (**Figure 3**) as well as crosstalk between them (**Figure 5**). These studies provide genetic resources and molecular markers that could be used in future studies.

All drought, salt and cold stress can induce osmotic and oxidative stress. Proline is involved in abiotic tolerance as a compatible osmolyte, molecular chaperone, and ROS scavenger (Szabados and Savouré, 2010). There are three *MtP5CS*s in *M. truncatula*. The *MtP5CS1* is constitutively expressed and the *MtP5CS3* participates in osmotic regulation (Armengaud et al., 2004; Kim and Nam, 2013). Both IAA and ABA induce proline accumulation but enhance the salt and drought tolerance, respectively, indicating that proline play different roles in salt and drought stress. Meanwhile, both drought and cold stress induce NO production, however, drought or cold induced NO production *via* different pathways (Planchet et al., 2014; Arfan et al., 2019). Complex crosstalk suggests that plants integrate hormones and signaling pathways to get better adaptation to abiotic stresses. With the help of modern molecular technologies, such as transgenic and CRISPR/Cas9 approaches, clarifying the functions of different factors is promising. Recently, Wang et al. (2022) using the CRISPR/Cas9 toolkit generates single and double knockout mutants in *MtDMP8* or *MtDMP9* and assesses their roles in haploid induction in *M. truncatula*. However, there is no report about tolerant mechanism to abiotic stresses using CRISPR/Cas9 technology.

Currently, many studies on *M. truncatula* are carried out in the culture room. So, more field experiments should be performed in future research. In field condition, plants often face several distinct environmental stresses simultaneously. For example, plants in arid regions often suffer from drought and heat stress (Iyer et al., 2013). In *M. truncatula*, interactive effects of ozone and drought have been well studied (Iyer et al., 2013). However, how combined occurrence of other kinds of abiotic stresses impact growth and development of *M. truncatula* is still not known yet and will be an important research topic in the future. In addition, *M. truncatula* is a cultivated species, some stress tolerance genes might have been lost during the domestication process (Wang T. Z. et al., 2021). *M. ruthenica*, a wild *Medicago* forage, retains these genes. Therefore, *M. ruthenica* provides a valuable model plant for studying the molecular mechanism of abiotic stresses tolerance in legumes.

Author contributions

XZ and TW conceived the concept of the work and wrote the manuscript. YS, XQ, HL, and IH revised the manuscript. All authors approved the final manuscript.

Funding

This research was co-funded by the National Natural Science Foundation of China (32101419 and 32070351) and the Science and Technology Program of Inner Mongolia, China (2021GG0372).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Alcázar, R., Altabella, T., Marco, F., Bortolotti, C., Reymond, M., Koncz, C., et al. (2010). Polyamines: molecules with regulatory functions in plant abiotic stress tolerance. *Planta* 231, 1237–1249.

Amouri, A. A., González, E. M., and Aoul, S. H. (2018). Physiological and biochemical characterization of rootlets response to salt stress in two *Medicago truncatula* Gaertn. ecotypes. *Plant Root* 12, 1–10.

Antoniou, C., Zarza, X., Gohari, G., Panahirad, S., Filippou, P., Tiburcio, A. F., et al. (2021). Involvement of polyamine metabolism in the response of *Medicago truncatula* genotypes to salt stress. *Plants* 10:269. doi: 10.3390/plants10020269

Araújo, S. S., Duque, A. S., Silva, J. M., Santos, D., Silva, A. B., and Fevereiro, P. (2013). Water deficit and recovery response of *Medicago truncatula* plants expressing the ELIP-like DSP22. *Biol. Plant.* 57, 159–163.

Arfan, M., Zhang, D. W., Zou, L. J., Luo, S. S., Tan, W. R., Zhu, T., et al. (2019). Hydrogen peroxide and nitric oxide crosstalk mediates brassinosteroids induced cold stress tolerance in *Medicago truncatula*. *Int. J. Mol. Sci.* 20:144. doi: 10.3390/ijms20010144

Ariel, F., Diet, A., Verdenaud, M., Gruber, V., Frugier, F., Chan, R., et al. (2010). Environmental regulation of lateral root emergence in *Medicago truncatula*

requires the HD-Zip I transcription factor HB1. *Plant Cell* 22, 2171–2183. doi: 10.1105/tpc.110.074823

Armengaud, P., Thiery, L., Buhot, N., Grenier de March, G., and Savouré, A. (2004). Transcriptional regulation of proline biosynthesis in *Medicago truncatula* reveals developmental and environmental specific features. *Physiol. Plant* 120, 442–450. doi: 10.1111/j.0031-9317.2004.00251.x

Arraouadi, S., Badri, M., Abdelly, C., Huguet, T., and Aouani, M. E. (2012). QTL mapping of physiological traits associated with salt tolerance in *Medicago truncatula* recombinant inbred lines. *Genomics* 99, 118–125. doi: 10.1016/j.ygeno.2011.11.005

Aydi, S., Sassi, S., and Abdelly, C. (2008). Growth, nitrogen fixation and ion distribution in *Medicago truncatula* subjected to salt stress. *Plant Soil* 312, 59–67.

Azani, N., Babineau, M., Donovan Bailey, C., Banks, H., Barbosa, A. R., Barbosa Pinto, R., et al. (2017). A new subfamily classification of the leguminosae based on a taxonomically comprehensive phylogeny: the Legume Phylogeny Working Group (LPWG). *Taxon* 66, 44–77.

Bassil, E., Zhang, S., Gong, H. J., Tajima, H., and Blumwald, E. (2019). Cation specificity of vacuolar NHX-type cation/H+ antiporters. *Plant Physiol.* 179, 616–629. doi: 10.1104/pp.18.01103

Bianco, C., and Defez, R. (2009). *Medicago truncatula* improves salt tolerance when nodulated by an indole-3-acetic acid-overproducing *Sinorhizobium meliloti* strain. *J. Exp. Bot.* 11, 3097–3107. doi: 10.1093/jxb/erp140

Charrier, A., Lelievre, E., Limami, A. M., and Planchet, E. (2013). *Medicago truncatula* stress associated protein 1 gene (MtSAP1) overexpression confers tolerance to abiotic stress and impacts proline accumulation in transgenic tobacco. *J. Plant Physiol.* 170, 874–877. doi: 10.1016/j.jplph.2013.01.008

Charrier, A., Planchet, E., Cerveau, D., Gimeno-Gilles, C., Verdu, I., Limami, A. M., et al. (2012). Overexpression of a *Medicago truncatula* stress-associated protein gene (MtSAP1) leads to nitric oxide accumulation and confers osmotic and salt stress tolerance in transgenic tobacco. *Planta* 236, 567–577. doi: 10.1007/s00425-012-1635-9

Chen, J. R., Lü, J. J., Liu, R., Xiong, X. Y., Wang, T. X., Chen, S. Y., et al. (2010). DREB1C from *Medicago truncatula* enhances freezing tolerance in transgenic *M. truncatula* and China Rose (*Rosa chinensis* Jacq.). *Plant Growth Regul.* 60, 199–211.

Confalonieri, M., Carelli, M., Tava, A., and Borrelli, L. (2019). Overexpression of MtTdp2α (tyrosyl-DNA phosphodiesterase 2) gene confers salt tolerance in transgenic *Medicago truncatula*. *Plant Cell Tiss. Org.* 137, 157–172.

Cramer, G. R., Urano, K., Delrot, S., Pezzotti, M., and Shinozaki, K. (2011). Effects of abiotic stress on plants: a systems biology perspective. *BMC Plant Biol.* 11:163. doi: 10.1186/1471-2229-11-163

De Lorenzo, L., Merchan, F., Blanchet, S., Megias, M., Frugier, F., Crespi, M., et al. (2007). Differential expression of the TFIIIA regulatory pathway in response to salt stress between *Medicago truncatula* genotypes. *Plant Physiol.* 145, 1521–1532. doi: 10.1104/pp.107.106146

De Lorenzo, L., Merchan, F., Laporte, P., Thompson, R., Clarke, J., Sousa, C., et al. (2009). A novel plant leucine-rich repeat receptor kinase regulates the response of *Medicago truncatula* roots to salt stress. *Plant Cell* 21, 668–680. doi: 10.1105/tpc.108.059576

de Zélicourt, A., Diet, A., Marion, J., Laffont, C., Ariel, F., Moison, M., et al. (2012). Dual involvement of a *Medicago truncatula* NAC transcription factor in root abiotic stress response and symbiotic nodule senescence. *Plant J.* 70, 220–230. doi: 10.1111/j.1365-313X.2011.04859.x

Delledonne, M., Zeier, J., Marocco, A., and Lamb, C. (2001). Signal interactions between nitric oxide and reactive oxygen intermediates in the plant hypersensitive disease resistance response. *Proc. Natl. Acad. Sci. U.S.A.* 98, 13454–13459. doi: 10.1073/pnas.231178298

Dong, S., Sang, L., Xie, H., Chai, M., and Wang, Z. Y. (2021). Comparative transcriptome analysis of salt stress-induced leaf senescence in *Medicago truncatula*. *Front. Plant Sci.* 12:666660. doi: 10.3389/fpls.2021.666660

Dong, W., Liu, X. J., Gao, T. X., and Song, Y. G. (2018). Analyses of DNA methylation and histone modification of MtMYBS1 in *Medicago truncatula* under salinity stress. *Plant Physiol. J.* 54, 1596–1604.

Dong, W., Song, Y., Zhao, Z., Qiu, N. W., Liu, X., and Guo, W. (2017). The *Medicago truncatula* R2R3-MYB transcription factor gene MtMYBS1 enhances salinity tolerance when constitutively expressed in *Arabidopsis thaliana*. *Biochem. Biophys. Res. Commun.* 490, 225–230. doi: 10.1016/j.bbrc.2017.06.025

Du, W. X., Yang, J. F., Ma, L., Su, Q., and Pang, Y. Z. (2021). Identification and characterization of abiotic stress responsive CBL-CIPK family genes in *Medicago*. *Int. J. Mol. Sci.* 22:4634. doi: 10.3390/ijms22094634

Duque, A. S., López-Gómez, M., Kráčmarová, J., Gomes, C. N., Araújo, S. S., Lluch, C., et al. (2016). Genetic engineering of polyamine metabolism changes *Medicago truncatula* responses to water deficit. *Plant Cell Tiss. Org.* 127, 681–690.

Farooq, M., Wahid, A., Kobayashi, N., Fujita, D., and Basra, S. M. A. (2009). Plant drought stress: effects, mechanism and management. *Agron. Sustain. Dev.* 29, 185–212.

Farooq, M. S., Uzair, M., Raza, A., Habib, M., Xu, Y. L., Yousuf, M., et al. (2022). Uncovering the research gaps to alleviate the negative impacts of climate change on food security: a review. *Front. Plant Sci.* 13:927535. doi: 10.3389/fpls.2022.927535

Ghosh, N., Das, S. P., Mandal, C., Gupta, S., Das, K., Dey, N., et al. (2012). Variations of antioxidative responses in two rice cultivars with polyamine treatment under salinity stress. *Physiol. Mol. Biol. Plants* 18, 301–313. doi: 10.1007/s12298-012-0124-8

Gou, J., Debnath, S., Sun, L., Flanagan, A., Tang, Y., Jiang, Q., et al. (2018). From model to crop: functional characterization of SPL8 in *M. truncatula* led to genetic improvement of biomass yield and abiotic stress tolerance in alfalfa. *Plant Biotechnol. J.* 16, 951–962. doi: 10.1111/pbi.12841

Gruber, V., Blanchet, S., Diet, A., Zahaf, O., Boualem, A., Kakar, K., et al. (2009). Identification of transcription factors involved in root apex responses to salt stress in *Medicago truncatula*. *Mol. Genet. Genomics* 281, 55–66. doi: 10.1007/s00438-008-0392-8

Guo, T., Wang, S. M., Zhang, T. J., Xu, L. X., Li, Y. R. Z., Chao, Y. H., et al. (2021). Expression of the *Medicago truncatula* MtDof32 transcription factor regulates plant growth and enhances abiotic stress tolerances in transgenic *Arabidopsis*. *Environ. Exp. Bot.* 183:10433.

Iyer, N., Tang, Y. H., and Mahalingam, R. (2013). Physiological, biochemical and molecular responses to a combination of drought and ozone in *Medicago truncatula*. *Plant Cell Environ.* 36, 706–720. doi: 10.1111/pce.12008

Kang, J. M., Xie, W. W., Sun, Y., Yang, Q. C., and Wu, M. S. (2010). Identification of genes induced by salt stress from *Medicago truncatula* L. seedlings. *Afr. J. Biotechnol.* 9, 7589–7594.

Kim, G. B., and Nam, Y. W. (2013). A novel Δ1-pyrroline-5-carboxylate synthetase gene of *Medicago truncatula* plays a predominant role in stress-induced proline accumulation during symbiotic nitrogen fixation. *J. Plant Physiol.* 170, 291–302. doi: 10.1016/j.jplph.2012.10.004

Li, D., Zhang, Y., Hu, X., Shen, X., Ma, L., Su, Z., et al. (2011). Transcriptional profiling of *Medicago truncatula* under salt stress identified a novel CBF transcription factor MtCBF4 that plays an important role in abiotic stress responses. *BMC Plant Biol.* 11:109. doi: 10.1186/1471-2229-11-109

Li, X., Feng, H., Wen, J. Q., Dong, J. L., and Wang, T. (2018). MtCAS31 aids symbiotic nitrogen fixation by protecting the leghemoglobin MtLb120-1 under drought stress in *Medicago truncatula*. *Front. Plant Sci.* 9:633. doi: 10.3389/fpls.2018.00633

Li, X., Hou, Y. Y., Zhang, F., Li, M. N., Yi, F. Y., Kang, J. M., et al. (2022). Identification and characterization of stress responsive homeodomain leucine zipper transcription factors in *Medicago truncatula*. *Mol. Biol. Rep.* 49, 3569–3581. doi: 10.1007/s11033-022-07197-4

Li, X., Liu, Q., Feng, H., Deng, J., Zhang, R., Wen, J., et al. (2020). Dehydrin MtCAS31 promotes autophagic degradation under drought stress. *Autophagy* 16, 862–877. doi: 10.1080/15548627.2019.1643656

Liu, L., Zhang, Z., Dong, J., and Wang, T. (2016). Overexpression of MtWRKY76 increases both salt and drought tolerance in *Medicago truncatula*. *Environ. Exp. Bot.* 123, 50–58.

Liu, M., Wang, T. Z., and Zhang, W. H. (2015). Sodium extrusion associated with enhanced expression of SOS1 underlies different salt tolerance between *Medicago falcata* and *Medicago truncatula* seedlings. *Environ. Exp. Bot.* 110, 46–55.

Long, R., Gao, Y., Sun, H., Zhang, T., Li, X., Li, M., et al. (2018). Quantitative proteomic analysis using iTRAQ to identify salt-responsive proteins during the germination stage of two *Medicago* species. *Sci. Rep.* 8:9553. doi: 10.1038/s41598-018-27935-8

Long, R., Li, M., Zhang, T., Kang, J., Sun, Y., Cong, L., et al. (2016). Comparative proteomic analysis reveals differential root proteins in *Medicago sativa* and *Medicago truncatula* in response to salt stress. *Front. Plant Sci.* 7:424. doi: 10.3389/fpls.2016.00424

López, M., Herrera-Cervera, J. A., Iribarne, C., Tejera, N. A., and Lluch, C. (2008). Growth and nitrogen fixation in *Lotus japonicus* and *Medicago truncatula* under NaCl stress: nodule carbon metabolism. *J. Plant Physiol.* 165, 641–650. doi: 10.1016/j.jplph.2007.05.009

Lu, F. L., Li, G. L., Cui, X., Liu, C. Y., Wang, X. J., and Cao, X. F. (2008). Comparative analysis of JmjC domain-containing proteins reveals the potential histone demethylases in *Arabidopsis* and rice. *J. Integr. Plant Biol.* 50, 886–896. doi: 10.1111/j.1744-7909.2008.00692.x

Lu, S. X., Knowles, S. M., Webb, C. J., Celaya, R. B., Cha, C., Siu, J. P., et al. (2011). The Jumonji C domain-containing protein JMJ30 regulates period length in the *Arabidopsis circadian* clock. *Plant Physiol.* 155, 906–915. doi: 10.1104/pp.110.167015

Luo, S. S., Sun, Y. N., Zhou, X., Zhu, T., Zhu, L. S., Arfan, M., et al. (2016). *Medicago truncatula* genotypes Jemalong A17 and R108 show contrasting variations under drought stress. *Plant Physiol. Biochem.* 109, 190–198. doi: 10.1016/j.plaphy.2016.09.019

Macovei, A., Balestrazzi, A., Confalonieri, M., and Carbonera, D. (2010). The tyrosyl-DNA phosphodiesterase gene family in *Medicago truncatula* Gaertn.: bioinformatic investigation and expression profiles in response to copper- and PEG-mediated stress. *Planta* 232, 393–407. doi: 10.1007/s00425-010-1179-9

Merchan, F., Breda, C., Hormaeche, J. P., Sousa, C., Kondorosi, A., Aguilar, O., et al. (2003). A Krüppel-like transcription factor gene is involved in salt stress responses in *Medicago* spp. *Plant Soil* 257, 1–9.

Merchan, F., De Lorenzo, L., Rizzo, S. G., Niebel, A., Manyani, H., Frugier, F., et al. (2007). Identification of regulatory pathways involved in the reacquisition of root growth after salt stress in *Medicago truncatula*. *Plant J.* 1, 1–17. doi: 10.1111/j.1365-313X.2007.03117.x

Mhadhbi, H., Fotopoulos, V., Mylona, P. V., Jebara, M., Aouani, M. E., and Polidoros, A. N. (2011). Antioxidant gene-enzyme responses in *Medicago truncatula* genotypes with different degree of sensitivity to salinity. *Physiol. Plant.* 141, 201–214. doi: 10.1111/j.1399-3054.2010.01433.x

Mhadhbi, H., Fotopoulos, V., Mylona, P. V., Jebara, M., Aouani, M. E., and Polidoros, A. N. (2013). Alternative oxidase 1 (Aox1) gene expression in roots of *Medicago truncatula* is a genotype-specific component of salt stress tolerance. *J. Plant Physiol.* 170, 111–114. doi: 10.1016/j.jplph.2012.08.017

Mishra, D., Shekhar, S., Chakraborty, S., and Chakraborty, N. (2021). High temperature stress responses and wheat: impacts and alleviation strategies. *Environ. Exp. Bot.* 190, 1–13.

Mohapatra, S. S., Wolfraim, L., Poole, R. J., and Dhindsa, R. S. (1989). Molecular cloning and relationship to freezing tolerance of cold-acclimation-specific genes of Alfalfa. *Plant Physiol.* 89, 375–380. doi: 10.1104/pp.89.1.375

Munns, R., and Tester, M. (2008). Mechanism of salinity tolerance. *Annu. Rev. Plant Biol.* 59, 651–681.

Nguyen, M. L., Kim, G. B., Hyun, S. H., Lee, S. Y., Lee, C. Y., Choi, H. K., et al. (2013). Physiological and metabolo mic analysis of a knockout mutant suggests a critical role of MtP5CS3 gene in osmotic stress tolerance of *Medicago truncatula*. *Euphytica* 193, 101–120.

Nunes, C., de Sousa Araújo, S., da Silva, J. M., Fevereiro, M. P. S., and da Silva, A. B. (2008). Physiological responses of the legume model *Medicago truncatula* cv. Jemalong to water deficit. *Environ. Exp. Bot.* 63, 289–296.

Oztur, Z. N., Talame, V., Deyholos, M., Michalowski, C. B., Galbraith, D. W., Gozukirmizi, N., et al. (2002). Monitoring large-scale changes in transcript abundance in drought- and salt-stressed barley. *Plant Mol. Biol.* 48, 551–573. doi: 10.1023/a:1014875215580

Pagano, A., Zannino, L., Pagano, P., Doria, E., Dondi, D., Macovei, A., et al. (2022). Changes in genotoxic stress response, ribogenesis and PAP (3′-phosphoadenosine 5′-phosphate) levels are associated with loss of desiccation tolerance in overprimed *Medicago truncatula* seeds. *Plant Cell Environ.* 45, 1457–1473. doi: 10.1111/pce.14295

Pennycooke, J. C., Cheng, H., and Stockinger, E. J. (2008). Comparative genomic sequence and expression analyses of *Medicago truncatula* and alfalfa subspecies falcata COLD-ACCLIMATION-SPECIFIC genes. *Plant Physiol.* 146, 1242–1254. doi: 10.1104/pp.107.108779

Planchet, E., Rannou, O., Ricoult, C., Boutet-Mercey, S., Maia-Grondard, A., and Limami, A. M. (2011). Nitrogen metabolism responses to water deficit act through both abscisic acid (ABA)-dependent and independent pathways in *Medicago truncatula* during post-germination. *J. Exp. Bot.* 2, 605–615. doi: 10.1093/jxb/erq294

Planchet, E., Verdu, I., Delahaie, J., Cukier, C., Girard, C., Morère-Le Paven, M., et al. (2014). Abscisic acid-induced nitric oxide and proline accumulation in independent pathways under water-deficit stress during seedling establishment in *Medicago truncatula*. *J. Exp. Bot.* 8, 2161–2170. doi: 10.1093/jxb/eru088

Porto, M. S., Pinheiro, M. P., Batista, V. G., dos Santos, R. C., Filho Pde, A., and de Lima, L. M. (2014). Plant promoters: an approach of structure and function. *Mol. Biotechnol.* 56, 38–49.

Qiao, W., and Fan, L. M. (2008). Nitric oxide signaling in plant responses to abiotic stresses. *J. Int. Plant Biol.* 50, 1238–1246.

Qu, Y., Duan, M., Zhang, Z., Dong, J., and Wang, T. (2016). Overexpression of the *Medicago falcata* NAC transcription factor MfNAC3 enhances cold tolerance in *Medicago truncatula*. *Environ. Exp. Bot.* 129, 67–76.

Raza, A., Tabassum, J., Kudapa, H., and Varshney, R. K. (2021). Can omics deliver temperature resilient ready-to-grow crops? *Crit. Rev. Biotechnol.* 41, 1209–1232. doi: 10.1080/07388551.2021.1898332

Rhaman, M. S., Imran, S., Karim, M. M., Chakrobortty, J., Mahamud, M. A., Sarker, P., et al. (2021). 5-aminolevulinic acid-mediated plant adaptive responses to abiotic stress. *Plant Cell Rep.* 40, 1451–1469. doi: 10.1007/s00299-021-02690-9

Saeed, F., Chaudhry, U. K., Bakhsh, A., Raza, A., Saeed, Y., Bohra, A., et al. (2022). Moving beyond DNA sequence to improve plant stress responses. *Front. Genet.* 13:874648. doi: 10.3389/fgene.2022.874648

Sandhu, D., Pudussery, M. V., Kaundal, R., Suarez, D. L., Kaundal, A., and Sekhon, R. S. (2018). Molecular characterization and expression analysis of the Na+/H+ exchanger gene family in *Medicago truncatula*. *Funct. Integr. Genomics* 18, 141–153.

Shen, Y. F., Wu, X. P., Liu, D. M., Song, S. J., Liu, D. C., and Wang, H. Q. (2016). Cold-dependent alternative splicing of a Jumonji C domain-containing gene MtJMJC5 in *Medicago truncatula*. *Biochem. Biophys. Res. Commun.* 474, 271–276. doi: 10.1016/j.bbrc.2016.04.062

Sheng, S., Guo, X. Y., Wu, C. Z., Xiang, Y. C., Duan, S. H., Yang, W. Q., et al. (2022). Genome-wide identification and expression analysis of DREB genes in alfalfa (*Medicago sativa*) in response to cold stress. *Plant Signal. Behav.* 17:e2081420. doi: 10.1080/15592324.2022.2081420

Shu, Y., Liu, Y., Zhang, J., Song, L., and Guo, C. (2016). Genome-wide analysis of the AP2/ERF superfamily genes and their responses to abiotic stress in *Medicago truncatula*. *Front. Plant Sci.* 6:1247. doi: 10.3389/fpls.2015.01247

Song, S. Y., Chen, Y., Zhao, M. G., and Zhang, W. H. (2012). A novel *Medicago truncatula* HD-Zip gene, MtHB2, is involved in abiotic stress responses. *Environ. Exp. Bot.* 80, 1–9.

Sun, Q., Huang, R., Zhu, H. F., Sun, Y. M., and Guo, Z. F. (2021). A novel *Medicago truncatula* calmodulin-like protein (MtCML42) regulates cold tolerance and flowering time. *Plant Physiol.* 108, 1069–1082. doi: 10.1111/tpj.15494

Szabados, L., and Savouré, A. (2010). Proline: a multifunctional amino acid. *Trends Plant Sci.* 15, 89–97.

Tang, H., Krishnakumar, V., Bidwell, S., Rosen, B., Chan, A., Zhou, S., et al. (2014). An improved genome release (version Mt4.0) for the model legume *Medicago truncatula*. *BMC Genomics* 15:312. doi: 10.1186/1471-2164-15-312

Thapa, B., Arora, R., Knapp, A. D., and Brummer, E. C. (2008). Applying freezing test to quantify cold acclimation in *Medicago truncatula*. *J. Amer. Soc. Hort. Sci.* 133, 684–691.

Tian, C., Zhang, X., He, J., Yu, H., Wang, Y., Shi, B., et al. (2014). An organ boundary-enriched gene regulatory network uncovers regulatory hierarchies underlying axillary meristem initiation. *Mol. Syst. Biol.* 10:755. doi: 10.15252/msb.20145470

Verdoy, D., Coba De La Pena, T., Redondo, F. J., Lucas, M. M., and Pueyo, J. J. (2006). Transgenic *Medicago truncatula* plants that accumulate proline display nitrogen-fixing activity with enhanced tolerance to osmotic stress. *Plant Cell Environ.* 29, 1913–1923. doi: 10.1111/j.1365-3040.2006.01567.x

Wallis, J. G., and Browse, J. (2002). Mutants of *Arabidopsis* reveal many roles for membrane lipids. *Prog. Lipid Res.* 41, 254–278. doi: 10.1016/s0163-7827(01)00027-3

Wang, N., Xia, X. Z., Jiang, T., Li, L. L., Zhang, P. C., Niu, L. F., et al. (2022). In planta haploid induction by genome editing of DMP in the model legume *Medicago truncatula*. *Plant Biotechnol. J.* 20, 22–24. doi: 10.1111/pbi.13740

Wang, S., Guo, T., Shen, Y., Wang, Z., Kang, J., Zhang, J., et al. (2021). Overexpression of MtRAV3 enhances osmotic and salt tolerance and inhibits growth of *Medicago truncatula*. *Plant Physiol. Biochem.* 163, 154–165. doi: 10.1016/j.plaphy.2021.04.003

Wang, T. Z., Ren, L. F., Li, C. H., Zhang, D., Zhang, X. X., Zhou, G., et al. (2021). The genome of a wild *Medicago* species provides insights into the tolerant mechanism of legume forage to environmental stress. *BMC Biol.* 19:96. doi: 10.1186/s12915-021-01033-0

Wang, T. Z., Tian, Q. Y., Wang, B. L., Zhao, M. G., and Zhang, W. H. (2014). Genome variations account for different response to three mineral elements between *Medicago truncatula* ecotypes Jemalong A17 and R108. *BMC Plant Biol.* 14:122. doi: 10.1186/1471-2229-14-122

Wang, T. Z., Zhang, J. L., Tian, Q. Y., Zhao, M. G., Zhang, W. H., and Diane, B. (2013). A *Medicago truncatula* EF-Hand family gene, MtCaMP1, is involved in drought and salt stress tolerance. *PLoS One* 8:e58952. doi: 10.1371/journal.pone.0058952

Wendehenne, D., Durner, J., and Klessig, D. F. (2004). Nitric oxide: a new player in plant signaling and defense responses. *Curr. Opin. Plant Biol.* 7, 449–455.

Winter, E., and Läuchli, A. (1982). Salt tolerance of *Trifolium alexandrinum* L. I. comparison of the salt response of *T. alexandrinum* and *T. pratense*. *Funct. Plant Biol.* 9, 221–226.

Wolfraim, L. A., Langis, R., Tyson, H., and Dhindsa, R. S. (1993). cDNA sequence, expression, and transcript stability of a cold acclimation-specific gene, cas18, of alfalfa (*Medicago falcata*) cells. *Plant Physiol.* 101, 1275–1282. doi: 10.1104/pp.101.4.1275

Xie, C., Zhang, R., Qu, Y., Miao, Z., Zhang, Y., Shen, X., et al. (2012). Overexpression of MtCAS31 enhances drought tolerance in transgenic *Arabidopsis* by reducing stomatal density. *New Phytol.* 195, 124–135. doi: 10.1111/j.1469-8137.2012.04136.x

Xin, Z. G., and Browse, J. (2000). Cold comfort farm: the acclimation of plants to freezing temperatures. *Plant Cell Environ.* 23, 893–902.

Yaish, M. W., Al-Lawati, A., Al-Harrasi, I., and Patankar, H. V. (2018). Genome wide DNA methylation analysis in response to salinity in the model plant caliph medic (*Medicago truncatula*). *BMC Genom.* 19:78. doi: 10.1186/s12864-018-4484-5

Yang, M. K., Wang, L. P., Chen, C. M., Guo, X., Lin, C. L., Huang, W., et al. (2021). Genome-wide analysis of autophagy-related genes in *Medicago truncatula* highlights their roles in seed development and response to drought stress. *Sci. Rep.* 11:22933. doi: 10.1038/s41598-021-02239-6

Yousfi, N., Slama, I., Ghnaya, T., Savoure, A., and Abdelly, C. (2010). Effects of water deficit stress on growth, water relations and osmolyte accumulation in *Medicago truncatula* and *M. laciniata* populations. *C. R. Biol.* 333, 205–213. doi: 10.1016/j.crvi.2009.12.010

Zahaf, O., Blanchet, S., de Zélicourt, A., Alunni, B., Plet, J., Laffont, C., et al. (2012). Comparative transcriptomic analysis of salt adaptation in roots of contrasting *Medicago truncatula* genotypes. *Mol. Plant* 5, 1068–1081. doi: 10.1093/mp/sss009

Zhang, J. Y., Broeckling, C. D., Blancaflor, E. B., Sledge, M. K., Sumner, L. W., and Wang, Z. Y. (2005). Overexpression of WXP1, a putative *Medicago truncatula* AP2 domain-containing transcription factor gene, increases cuticular wax accumulation and enhances drought tolerance in transgenic alfalfa (*Medicago sativa*). *Plant J.* 42, 689–707. doi: 10.1111/j.1365-313X.2005.02405.x

Zhang, J. Y., Broeckling, C. D., Sumner, L. W., and Wang, Z. Y. (2007). Heterologous expression of two *Medicago truncatula* putative ERF transcription factor genes, WXP1 and WXP2, in *Arabidopsis* led to increased leaf wax accumulation and improved drought tolerance, but differential response in freezing tolerance. *Plant Mol. Biol.* 64, 265–278. doi: 10.1007/s11103-007-9150-2

Zhang, J. Y., Cruz, D. E. C. M. H., Torres-Jerez, I., Kang, Y., Allen, S. N., Huhman, D. V., et al. (2014). Global reprogramming of transcription and metabolism in *Medicago truncatula* during progressive drought and after rewatering. *Plant Cell Environ.* 37, 2553–2576. doi: 10.1111/pce.12328

Zhang, L. L., Zhao, M. G., Tian, Q. Y., and Zhang, W. H. (2011). Comparative studies on tolerance of *Medicago truncatula* and *Medicago falcata* to freezing. *Planta* 234, 445–457.

Zhang, X. X., Wang, T. Z., Liu, M., Sun, W., and Zhang, W. H. (2019). Calmodulin-like gene MtCML40 is involved in salt tolerance by regulating MtHKTs transporters in *Medicago truncatula*. *Environ. Exp. Bot.* 157, 79–90.

Zhang, P. P., Li, S. S., Guo, Z. F., and Lu, S. Y. (2019). Nitric oxide regulates glutathione synthesis and cold tolerance in forage legumes. *Environ. Exp. Bot.* 167, 103851–103858.

Zhang, Z., Hu, X., Zhang, Y., Miao, Z., Xie, C., Meng, X., et al. (2016). Opposing control by transcription factors MYB61 and MYB3 increases freezing tolerance by relieving C-repeat binding factor suppression. *Plant Physiol.* 172, 1306–1323. doi: 10.1104/pp.16.00051

Zhang, Z., Wei, X., Liu, W., Min, X., Jin, X., Ndayambaza, B., et al. (2018). Genome-wide identification and expression analysis of the fatty acid desaturase genes in *Medicago truncatula*. *Biochem. Biophys. Res. Commun.* 499, 361–367. doi: 10.1016/j.bbrc.2018.03.165

Zhao, M. G., Chen, L., Zhang, L. L., and Zhang, W. H. (2009). Nitric reductase-dependent nitric oxide production is involved in cold acclimation and freezing tolerance in *Arabidopsis*. *Plant Physiol.* 151, 755–767. doi: 10.1104/pp.109.140996

Zhao, M. G., Liu, W. J., Xia, X. Z., Wang, T. Z., and Zhang, W. H. (2014). Cold acclimation-induced freezing tolerance of *Medicago truncatula* seedlings is negatively regulated by ethylene. *Physiol. Plant.* 152, 115–129. doi: 10.1111/ppl.12161

Zhou, Y., Zeng, L., Chen, R., Wang, Y., and Song, J. (2018). Genome-wide identification and characterization of stress-associated protein (SAP) gene family encoding A20/AN1 zinc-finger proteins in *Medicago truncatula*. *Arch. Biol. Sci.* 70, 87–98.

Zhuo, C., Liang, L., Zhao, Y., Guo, Z., and Lu, S. (2018). A cold responsive ethylene responsive factor from *Medicago falcata* confers cold tolerance by up-regulation of polyamine turnover, antioxidant protection, and proline accumulation. *Plant Cell Environ.* 41, 2021–2032. doi: 10.1111/pce.13114

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 20 September 2022
DOI 10.3389/fpls.2022.938635

OPEN ACCESS

EDITED BY
Rajeev K. Varshney,
International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), India

REVIEWED BY
Haoqiang Yu,
Maize Research Institute of Sichuan Agricultural University, China
Neelam Mishra,
St Joseph's College (Autonomous), India

*CORRESPONDENCE
Yan Li
yanli1@njau.edu.cn

SPECIALTY SECTION
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

RECEIVED 07 May 2022
ACCEPTED 30 August 2022
PUBLISHED 20 September 2022

CITATION
Jin T, An J, Xu H, Chen J, Pan L, Zhao R, Wang N, Gai J and Li Y (2022) A soybean sodium/hydrogen exchanger GmNHX6 confers plant alkaline salt tolerance by regulating Na^+/K^+ homeostasis. *Front. Plant Sci.* 13:938635. doi: 10.3389/fpls.2022.938635

A soybean sodium/hydrogen exchanger GmNHX6 confers plant alkaline salt tolerance by regulating Na^+/K^+ homeostasis

Ting Jin, Jiaxin An, Huadong Xu, Jie Chen, Lang Pan, Ranran Zhao, Ning Wang, Junyi Gai and Yan Li*

National Key Laboratory of Crop Genetics and Germplasm Enhancement, Key Laboratory for Biology and Genetic Improvement of Soybean (General, Ministry of Agriculture), National Center for Soybean Improvement, Jiangsu Collaborative Innovation Center for Modern Crop Production, Nanjing Agricultural University, Nanjing, China

Alkaline soil has a high pH due to carbonate salts and usually causes more detrimental effects on crop growth than saline soil. Sodium hydrogen exchangers (NHXs) are pivotal regulators of cellular Na^+/K^+ and pH homeostasis, which is essential for salt tolerance; however, their role in alkaline salt tolerance is largely unknown. Therefore, in this study, we investigated the function of a soybean *NHX* gene, *GmNHX6*, in plant response to alkaline salt stress. *GmNHX6* encodes a Golgi-localized sodium/hydrogen exchanger, and its transcript abundance is more upregulated in alkaline salt tolerant soybean variety in response to $NaHCO_3$ stress. Ectopic expression of *GmNHX6* in *Arabidopsis* enhanced alkaline salt tolerance by maintaining high K^+ content and low Na^+/K^+ ratio. Overexpression of *GmNHX6* also improved soybean tolerance to alkaline salt stress. A single nucleotide polymorphism in the promoter region of *NHX6* is associated with the alkaline salt tolerance in soybean germplasm. A superior promoter of *GmNHX6* was isolated from an alkaline salt tolerant soybean variety, which showed stronger activity than the promoter from an alkaline salt sensitive soybean variety in response to alkali stress, by luciferase transient expression assays. Our results suggested soybean *NHX6* gene plays an important role in plant tolerance to alkaline salt stress.

KEYWORDS

abiotic stress, alkaline salt tolerance, natural variation, promoter, sodium bicarbonate, sodium hydrogen exchanger, soybean

Introduction

Saline-alkali soils currently account for 20% of irrigated land (Bouras et al., 2022). Land alkalization has become a major and increasingly serious problem in the world. Compared with neutral salt stress, the high pH environment in alkaline soil destroys soil structure and affects the absorption of essential elements such as phosphorus and iron by plants (Shi and Wang, 2005). Therefore, alkaline salt causes more serious effects than neutral salt stress on

plants. Soil alkalization can reduce soil osmotic potential, and cause ion imbalance, disrupt physiological processes, inhibit the growth and development of plants, leading to a serious decline in the yield and quality, and even the death of plants (Yang et al., 2011). Saline-alkali stress leads to leaf wilting and chlorosis (Sun et al., 2021), fewer pods, seeds and lower 100-seed weight (He et al., 2020), inhibited nodule development (He et al., 2021) and other phenomena (Chen et al., 2018) of soybean, resulting in soybean growth retardation, and eventually lead to a serious decline in soybean yield (He et al., 2020). Plants can effectively reduce the damage caused by salt-alkali stress through ion selective absorption, compartmentalization and scavenging reactive oxygen species, etc. (Zhu, 2001; Zhang et al., 2019a). Sodium hydrogen exchanger (NHXs) are integral membrane proteins residing in the plasma membrane, endosome (Dragwidge et al., 2019), and vacuole (Sellamuthu et al., 2020), which belong to the monovalent cation/proton antiporter family (Bassil et al., 2011b). NHXs play important roles in regulating of cellular ion homeostasis (Bassil et al., 2019), pH (Reguera et al., 2015), vesicle trafficking (Bassil et al., 2019), protein transport (Dragwidge et al., 2019), auxin transport (Zhang et al., 2020a), salt tolerance (Long et al., 2020), cell turgor and expansion (Walker et al., 1996), as well as growth and development (Bassil et al., 2019) in many species ranging from bacteria to human. In plants, NHX utilizes an H^+ electrochemical gradient established by (H^+)-ATPase (in the plasma membrane) and (H^+)-PPase (in vacuoles) to allow Na^+ or K^+ exchange for H^+, to maintain pH and ion homeostasis (Rodríguez-Rosales et al., 2009; Bassil et al., 2012).

In *Arabidopsis thaliana*, the NHX exchanger family contains eight members which are divided into three subclasses based on their subcellular localizations: vacuolar (AtNHX1-AtNHX4), endosomal (AtNHX5 and AtNHX6), and plasma membrane (AtNHX7/SOS1and AtNHX8) localized exchangers (Reguera et al., 2014; Qiu, 2016a). Plasma membrane NHXs (AtNHX7/SOS1and AtNHX8) are required for Na^+, K^+, and pH homeostasis, and play an important role in salt tolerance (An et al., 2007; Oh et al., 2010). Vacuolar membrane NHXs (AtNHX1-AtNHX4) are critical for vacuolar pH and K^+ homeostasis (Barragán et al., 2012), salt and drought stress responses (Moghaieb et al., 2014; Zhang et al., 2019b), osmotic adjustment (Quintero et al., 2000), flower development (Bassil et al., 2011b), as well as plant growth and development (Qiu, 2016a). AtNHX5 and AtNHX6 localize to the Golgi and trans-Golgi network (TGN), and are important for plant growth and response to salt stress (Bassil et al., 2011b), maintaining ion and pH homeostasis (Wang et al., 2015), protein transport (Qiu, 2016b), plant growth and development (Lv et al., 2020), and seedling growth (Zhang et al., 2020b). The *Arabidopsis nhx5 nhx6* double mutant showed reduced growth, smaller and fewer cells, smaller rosettes and shorter seedlings, late flowering, and increased sensitivity to salinity (Bassil et al., 2011a). The *nhx5 nhx6 syp22* triple mutant had short siliques and low seed setting rate, but larger seeds. In addition, the triple mutant had numerous smaller protein storage vacuoles (PSVs) and accumulated precursors of seed storage proteins, suggesting that *AtNHX5* and *AtNHX6* play important roles in seed production, protein trafficking and PSV biogenesis (Wu et al., 2016). A recent report suggests that NHX5 and NHX6 might regulate auxin transport through endoplasmic reticulum (Fan et al., 2018).

There are also few studies on NHX5 or NHX6 from other plant species in addition to *Arabidopsis*. Overexpression of *PdNHX6* from *Phoenix dactylifera* in *Arabidopsis* plants enhanced salt tolerance, retained higher chlorophyll and water content, maintained a balanced Na^+/K^+ ratio, and increased seed germination under salinity when compared to the wild-type plants (Al-Harrasi et al., 2020). *AoNHX6* from *Avicennia officinalis* showed high expression levels in the roots, and complementation with *AoNHX6* improved the tolerance of yeast mutants and *Arabidopsis* mutants to both NaCl and KCl stress (Krishnamurthy et al., 2019). The ectopic expression of endosomal-type *MnNHX6* from *Morus notabilis* in *Arabidopsis* and *nhx1* yeast mutant can greatly enhance their salt tolerance compared with vacuolar-type MnNHXs (Cao et al., 2020). Up to now, the role of *NHX* genes in soybean tolerance to salt-alkali stress is largely unknown, except that *GmNHX5* is found regulating salt tolerance in a recent study (Sun et al., 2021). Alkaline soils cause damage to plants not only through salt stress, but also through high pH (Shi and Wang, 2005), therefore, alkali stress is more serious than salt stress, while the role of *NHX* genes in alkali tolerance of soybean remains unclear.

In our previous transcriptomic study (unpublished) using an alkaline salt tolerant wild soybean variety, we identified an alkali-responsive gene (corresponding to Glyma.09G018200 in the reference genome of soybean variety Williams 82), which encodes a sodium/hydrogen exchanger, and is designated as *GmNHX6*. Here we investigated the role of *GmNHX6* in alkali stress by comparing the relative expression level of *GmNHX6* gene in alkaline salt tolerant and sensitive soybean varieties, and study the function of this gene in transgenic soybean composite plants and *Arabidopsis*. The sequence variation in *GmNHX6* and its association with alkaline salt tolerance was also analyzed, to gain a better understanding of its role in alkaline salt tolerance. The possible molecular mechanism of *GmNHX6* in response to alkaline salt stress was investigated by the content of Na^+, K^+, the Na^+/K^+ ratio, and promoter luciferase (LUC) assay. This study aims to provide new insights into the role of *GmNHX6* in soybean tolerance to alkaline salt stress and the possible underlying mechanisms.

Materials and methods

Soybean accessions and sodic tolerance rating

A total of 60 soybean accessions (Supplementary Table 1), were obtained from the National Center for Soybean Improvement (Nanjing, China). Twelve seeds of each soybean accession were germinated in plastic pots (φ10 × 8 cm) filled with sterile nutrient soil, and irrigated with 1/2 Hoagland nutrient solution form bottom, at 28°C (day)/24°C (night) with 14 h (light)/10 h (dark) photoperiod.

Fifteen pots were placed in a plastic tray (55 cm × 36 cm × 8 cm) containing 2 L fresh 1/2 Hoagland nutrient solution (pH ≈ 6.5), and the solution was changed every 2 days. After emergence, the seedings were thinned to four plants in each pot. When the second trifoliolate leaves appeared (12-day-old soybean seedlings), plants were treated with 1/2 Hoagland solution containing 0 or 90 mM $NaHCO_3$. Sixteen days after treatment, the alkaline salt tolerance was determined by Sodic Tolerance Rating (STR) according to the previously published method (Tuyen et al., 2010).

RNA extraction and gene expression analysis

RNA-seq data from different soybean tissues were downloaded from Soybase.[1] To experimentally determine the expression of *GmNHX6*, tissues were collected from soybean plants grown in vermiculite in greenhouse with controlled temperature (day/night, 28°C/24°C) and light cycle (day/night, 14/10 h). *Arabidopsis thaliana* plants were grown under controlled temperatures of day/night as 22/22°C and light cycle of day/night for 16/8 h. Fresh tissues were ground and extracted for total RNA using an RNAprep Pure Plant Kit (Tiangen Biotech, China). The gene specific primers were designed using the Primer Premier 5 software[2] and listed in Supplementary Table 2. The cDNA was synthesized by the PrimeScript™ 1st Strand cDNA Synthesis Kit (TaKaRa, Japan). The qRT-PCR was conducted using SYBR Premix ExTaq™ II Mix (TaKaRa, Japan) on a Roche 480 Real-time detection system (Roche Diagnostics, Switzerland) according to the manufacturers' instructions. Each experiment was performed in triplicates. Transcript levels in soybean plants were calculated in relative to the reference gene *GmUKN1* (Guan et al., 2014), using the $2^{-\triangle\triangle CT}$ methods (Livak and Schmittgen, 2001). The expression of *GmUKN1* is stable across all samples in this study (Supplementary Table 3). The qRT-PCR of *GmNHX6* gene in transgenic *A. thaliana* was conducted using the OE-2 line (with lowest expression level of *GmNHX6*) as the control (relative expression = 1) and *AtACTIN7* (Miao et al., 2020) as the reference gene. Data was collected from three biological replicates. The amplification efficiencies (E) of primer pairs were estimated (Supplementary Table 2) by qRT-PCR using 1:10, 1:20, 1:40, 1:80, and 1:160 dilutions of cDNA templates, according to the equation: $E = [10^{-1/slope}] - 1$ (Pfaffl, 2001).

Sequence alignment, phylogenetic analysis, and sequence analysis

The full sequences of GmNHX6 and other NHXs proteins obtained from Phytozome,[3] were used for multiple sequence alignments by ClustalW2.[4] The unrooted phylogenetic tree was then constructed using MEGA 6.0 (Tamura et al., 2013) based on the Maximum Likelihood (ML) algorithm with 1,000 bootstraps. The putative promoter sequence of 2000-bp upstream *GmNHX6* was analyzed using the PlantCARE.[5] Protein transmembrane topology and signal peptides were predicted from amino acid sequences using Protter database.[6]

Subcellular localization of GmNHX6

For subcellular localization of GmNHX6 protein, the open reading frame (ORF) of *GmNHX6* was amplified by PCR using specific primers (Supplementary Table 2). The PCR product was then introduced into the pAN580 vector, which contains a green florescence protein (GFP) reporter gene, under the drive of CaMV 35S promoter for N-terminal GFP fusion (Zhang et al., 2016), using the ClonExpress II One Step Cloning Kit (Vazyme, China) according to the manufacturer's protocol. To further determine the specific subcellular localization of GmNHX6, the fluorescent marker protein, Man1-mCherry, which is characteristic for the *cis*-Golgi (Tse et al., 2004), was co-expressed with GmNHX6. The transient expression of above proteins in *Arabidopsis* mesophyll protoplasts was performed following the published methods (Wu et al., 2009). Confocal imaging analysis was performed using a laser scanning microscope (Zeiss LSM780 META, Jena, Germany).

Obtaining and phenotyping of transgenic soybean composite plants

The coding sequence of *GmNHX6* from soybean accession M8206 was cloned into $pBinGFP_4$ vector under the control of CaMV 35S promoter. The construct containing 35S:*GmNHX6* and the empty vector $pBinGFP_4$, were separately transformed into *Agrobacterium rhizogenes* strain K599 (Kereszt et al., 2007), then used to infect soybean hypocotyls of an alkaline salt sensitive soybean variety Tianlong 1, to obtain transgenic soybean composite plants, according to the previously described method (Kereszt et al., 2007). After 20 days of plant growth, the green fluorescence signals of GFP in positive hairy roots were identified at 488 nm wavelength using a stereoscopic fluorescence microscope (Mshot, China), and the non-transgenic roots were cut-off. The transgenic soybean composite plants were treated with 1/2 Hoagland solution containing 0 or 90 mM $NaHCO_3$ for 7 days, then the Soil and Plant Analysis Development (SPAD) value for chlorophyll content (Pérez-Patricio et al., 2018) and leaf relative water content (LRWC) (Zhang and Zhang, 2021) were measured. The SPAD values of the second, third and fourth trifoliolate leaves were recorded 10 times each by the

1 http://soybase.org
2 http://www.premierbiosoft.com/primerdesign/
3 https://phytozome.jgi.doe.gov/pz/portal.html
4 https://www.ebi.ac.uk/Tools/msa/clustalw2/
5 http://bioinformatics.psb.ugent.be/webtools/plantcare/html/
6 https://wlab.ethz.ch/protter/start/

Chlorophyll Meter SPAD-502 Plus (Konica Minolta, Tokyo, Japan), and the average value was calculated as the SPAD value of the plant. To measure LRWC, the fresh weight (FW) of the second, third and fourth trifoliolate leaves was recorded, and then the leaves were incubated in sterile water at room temperature for 12 h, and the turgid fresh weight (TFW) was measured. The fully turgid leaves were then dried at 80°C for 72 h and the dry weight (DW) was recorded. LRWC was measured according to the following equation: LRWC = (FW – DW)/(TFW – DW) × 100%. Three biological replications were performed and five independent transgenic plants ($n = 3 \times 5 = 15$) were measured for each repeat.

Functional analysis of *GmNHX6* in *Arabidopsis thaliana*

The coding sequence of *GmNHX6* from soybean accession M8206 was cloned into pCAMBIA3301 vector under the control of CaMV 35S promoter. The pCAMBIA3301-35S:*GmNHX6* vector was introduced into the *A. tumefaciens* strain EHA105 for *A. thaliana* transformation using Col-0 wild-type (WT) plants by the floral dipping method (Clough and Bent, 1998). Transgenic plants were screened based on 20 mg L^{-1} glufosinate (PhytoTech, United States) resistance (Ren et al., 2009). The homozygous T_3 generation plants were subjected to *GmNHX6* gene expression analyses and phenotypic evaluation. The *Arabidopsis nhx6* T-DNA insertion mutant (stock, *SALK-100042C*, *nhx6*, Col-0 background) used in this study was obtained from the *Arabidopsis* Biological Resource Center.[7]

To measure the seed germination rate of *A. thaliana*, 36 stratified seeds for each of three independent *GmNHX6*-overexpressed lines (OE-1, OE-9, and OE-18), as well WT and *nhx6* mutant, were plated on 1/2 MS agar medium, or the media supplemented with 150 mM NaCl (Zheng et al., 2021) or 8 mM $NaHCO_3$ (Zhu et al., 2011), and then placed in a growth chamber under the long-day growth conditions (16 h light/8 h dark cycle at 22°C ± 2°C). The rates of cotyledon greening were measured as seed germination rates after 7 days. To measure the root length and fresh weight of *Arabidopsis* seedlings, the sterilized seeds of *Arabidopsis* were vertically grown on 1/2 MS agar medium for 7 days, and then the seedlings were transplanted to fresh 1/2 MS agar medium, or media supplemented with 150 mM NaCl (Zheng et al., 2021) or 8 mM $NaHCO_3$ (Zhu et al., 2011) for another 10 days. The root length was analyzed using Image J software.[8]

Determination of ion content

Seven-day-old *Arabidopsis* seedlings were transplanted to fresh 1/2 MS agar medium, or media supplemented with 150 mM NaCl or 8 mM $NaHCO_3$ for another 10 days. Then the seedlings were used for Na^+ and K^+ content determination. To determine the ion content, plants were rinsed thoroughly in deionized water, separated into roots and shoots, inactivated at 105°C for 2 h, and oven dried at 80°C until they attained a constant mass. The 0.1 g dry material was mineralized using 2 ml 10% HNO_3. Then, the microwave digestion system ETHOS T (Milestone, Italy)was used for ion extraction, and the ion contents in the supernatants were analyzed using an Optima 8,000 ICP-OES DV Spectrometer (PerkinElmer, United States) according to the manufacturer's instructions (Munns et al., 2012). Three biological replicates were performed.

Analysis of sequence variation and regional association study

Nucleotide sequences were downloaded from phytozome v10 (see Footnote 3), and the 2000-bp upstream of the start codon ATG was considered as the promoter region. The full-length coding sequences were amplified using soybean cDNA as template, and the promoter regions were amplified using soybean genomic DNA as template. The amplicons were sequenced at GenScript (Nanjing, China), and sequence variations among soybean accessions were analyzed. Regional association study was performed using TASSEL 5.0 software (Bradbury et al., 2007). The threshold for a significant association was determined using the previously published method (Yang et al., 2014; Wang et al., 2020), which is $p < 3.3 \times 10^{-2}$ ($1/n$, where n is the number of markers) in this study.

Promoter-LUC assays in tobacco leaves

The promoter activities were analyzed by promoter-LUC transient expression assays in tobacco (*Nicotiana benthamiana*) leaves, according to the procedures described previously (Jin et al., 2021). For vector constructs, the full-length 2-kb promoter regions of M8206 and ZY were amplified, respectively, and the amplicons were inserted into pGreenII0800-LUC to obtain different promoter-LUC vectors. Above constructs with LUC as a reporter, Pro(ZY)-LUC, Pro(M)-LUC, were transformed into *A. tumefaciens* strain GV3101 that carries pSoup-19, and introduced into tobacco leaves by infiltration. Then, the tobacco seedlings were transferred to 0 or 100 mM $NaHCO_3$ for 16 h. LUC activity was observed with an *in vivo* plant imaging system (Berthold LB 985, Germany).

Statistical analyses

For qRT–PCR analyses, at least four individual plants were pooled per tissue sample, and at least three qRT–PCR reactions (technical replicates) with at least three biological replicates were performed. For phenotypic evaluation, at least 12 individual plants

7 http://www.arabidopsis.org
8 https://imagej.nih.gov/ij/

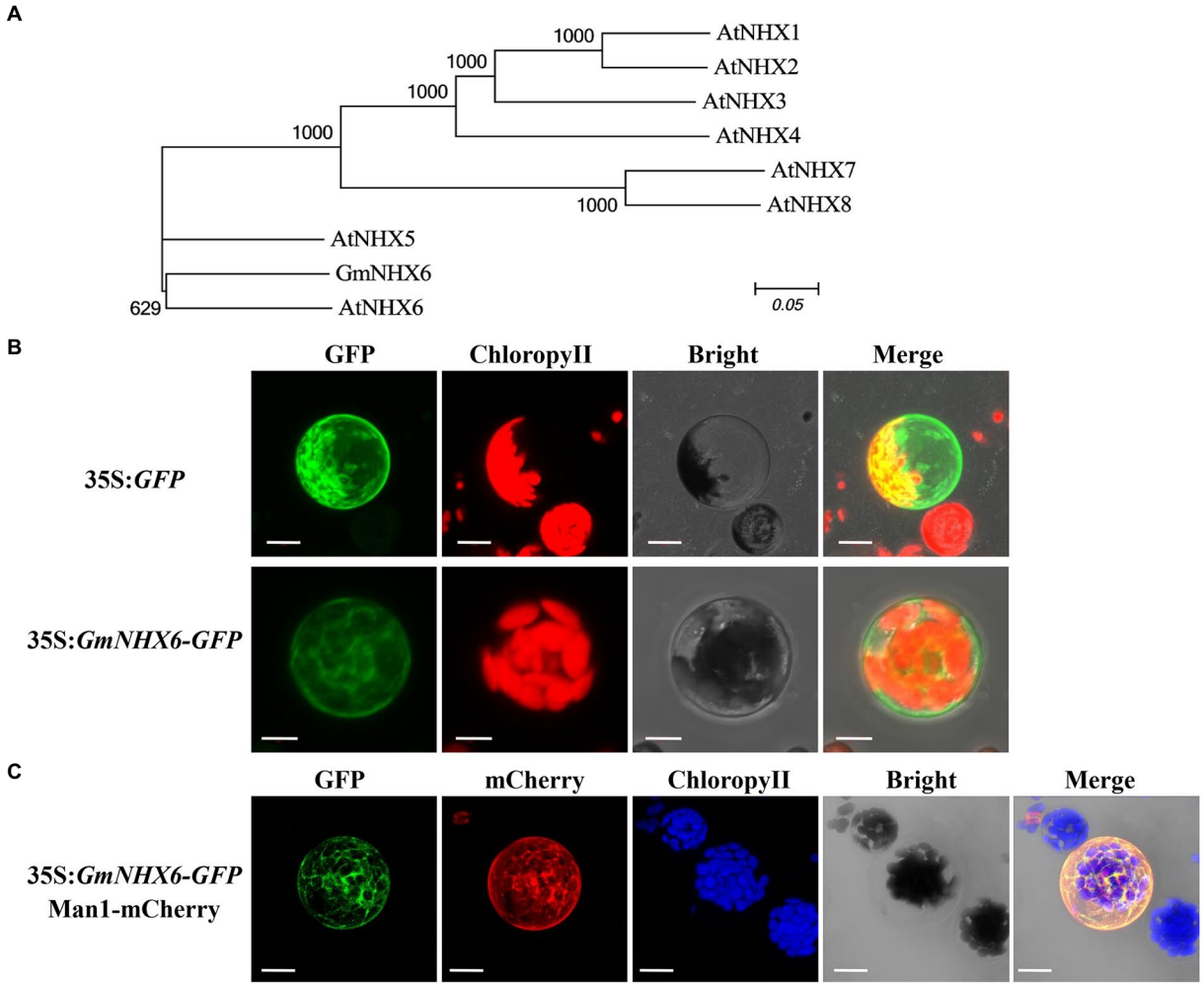


FIGURE 1
Phylogenetic analysis and subcellular localization of GmNHX6. **(A)** Phylogenetic analysis of GmNHX6 (XP_028181010.1) from soybean *Glycine max* and NHX proteins from *Arabidopsis thaliana* including AtNHX1 (NP_198067.1), AtNHX2 (NP_187154.1), AtNHX3 (NP_200358.1), AtNHX4 (NP_187288.2), AtNHX5 (NP_175839.2), AtNHX6 (NP_178079.2), AtNHX7 (NP_178307.2), AtNHX8 (NP_172918.2). The sequences of NHX proteins were downloaded from the Phytozome database. Unrooted phylogenetic tree was constructed using MEGA6.0 based on the Maximum Likelihood algorithm with 1,000 bootstraps. **(B)** Subcellular localization of GmNHX6. GFP protein and GmNHX6-GFP fusion protein expressed under the control of CaMV 35S promoter in *Arabidopsis* protoplasts were observed under a confocal microscope. Bar, 10μm. **(C)** Images of *Arabidopsis* mesophyll protoplasts co-expressing GmNHX6-GFP fusion protein and a mCherry Golgi apparatus marker (Man1-mCherry). Bars=10μm.

were analyzed per accession, and the exact numbers of individuals (*n*) are presented in all figure legends. Differences between groups or genotypes were analyzed using two-sided Wilcoxon test, Duncan's multiple range test or Student's *t*-test by R software and SAS 9.2 (SAS Institute Inc., Cary, NC, United States).

Results

GmNHX6 encodes a Golgi-localized sodium/hydrogen exchanger

GmNHX6 gene encodes a sodium/hydrogen exchanger consisting of 534 amino acids. Phylogenetic analyses of GmNHX6 with NHX proteins in *Arabidopsis* showed that GmNHX6 is closely related with AtNHX6 (Figure 1A). Prediction of GmNHX6 transmembrane topology displays 12 transmembrane domains (Supplementary Figure 1), sharing high similarities with AtNHX6. GmNHX6 contains four conserved acidic residues (Al-Harrasi et al., 2020) in transmembrane domains (Supplementary Figure 2), which were previously proven to be essential for K^+ transport (Wang et al., 2015). To determine the subcellular localization of GmNHX6, the *GmNHX6* gene was expressed in fusion with the *GFP* reporter gene, in the pAN580-GFP expression vector. The constructed vector and empty control vector were transferred to the protoplast of *Arabidopsis*, and the location of GFP or its fusion protein was observed. We found that, GFP itself was distributed evenly in the cytoplasm and the nucleus, whereas the

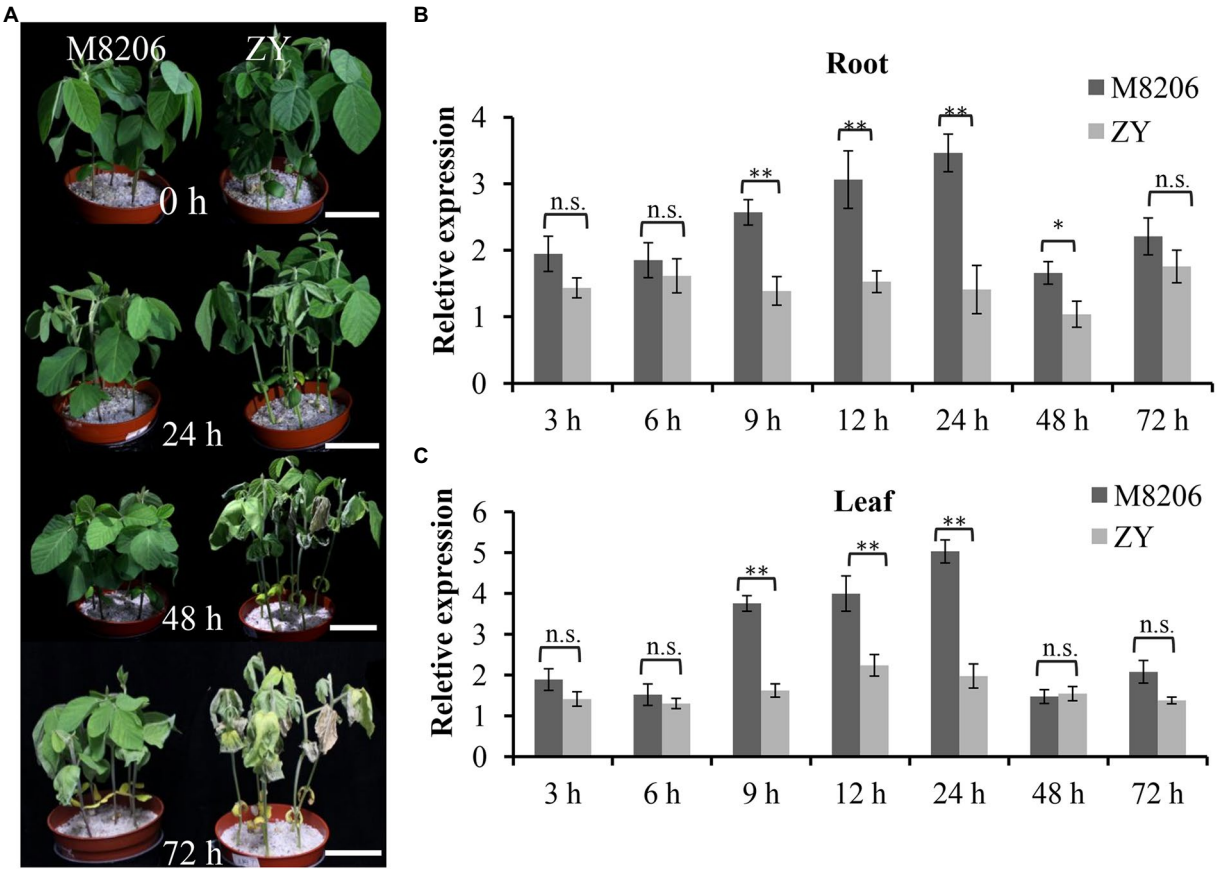


FIGURE 2
Phenotype and relative expression of *GmNHX6* in two soybean varieties in response to alkali stress. **(A)** Phenotypes of two soybean varieties at 0, 24, 48, and 72h after 90mM $NaHCO_3$ treatment. The 12-day-old seedlings were subjected to treatment. Bar=5cm. **(B,C)** Relative expression of *GmNHX6* in response to alkali stress in roots and leaves, respectively. Soybean seedlings were treated with 0 or 90mM $NaHCO_3$ for 3, 6, 9, 12, 24, 48, and 72h. Roots and leaves receiving 0mM $NaHCO_3$ treatment at each time point were used as controls. Data represents the mean±standard deviation of three biological replications with three repeats within each replication (n=3×3=9). Differences between two soybean varieties were evaluated using the two-tailed Student's *t*-tests (n.s., not significant; $^*p<0.05$; $^{**}p<0.01$).

GmNHX6–GFP fusion protein was mainly localized to the cytoplasm and to punctate compartments in the cytosol (Figure 1B). To determine the nature of these punctate compartments, we co-expressed GmNHX6–GFP and the fluorescent marker characteristic for the Golgi, Man1-mCherry (Tse et al., 2004). The co-expression results showed that GmNHX6-GFP fusion protein was mainly located in the Golgi (Figure 1C).

The expression of *GmNHX6* is more upregulated in alkaline salt tolerant soybean in response to $NaHCO_3$ stress

The phenotypic difference between two soybean varieties of M8206 (alkaline salt tolerant) and ZY (alkaline salt sensitive) was obvious under 90 mM $NaHCO_3$ stress (Figure 2A): ZY already showed wilted leaves at 24 h, but the leaves of M8206 were still normal. The leaves of ZY became yellow and dry out at 48 h and more severe at 72 h, while the leaves of M8206 did not show obvious wilting at 48 h until 72 h. The expression patterns of *GmNHX6* in *G. max* (see Footnote 1) showed that the transcript abundance of *GmNHX6* was higher in the young_leaf and root than the other organs (Supplementary Figure 3). The relative expression levels of *GmNHX6* in response to alkaline salt stress (90 mM $NaHCO_3$) in two soybean varieties were investigated using qRT-PCR. *GmNHX6* gene expression was induced by 90 mM $NaHCO_3$ treatment in the roots and leaves of both soybean varieties (Figures 2B,C). The relative expression level of *GmNHX6* gene in the roots of M8206 (alkaline salt tolerant) showed significantly greater increase at 9, 12, 24, and 48 h after 90 mM $NaHCO_3$ treatment than the alkaline salt sensitive soybean variety ZY (Figure 2B), and the relative expression level of *GmNHX6* gene in the leaves of M8206 showed significantly greater increase at 9, 12, and 24 h in response to $NaHCO_3$ treatment than ZY (Figure 2C). There was no significant difference in *GmNHX6* gene expression between two soybean varieties under normal condition (Supplementary Figure 4).

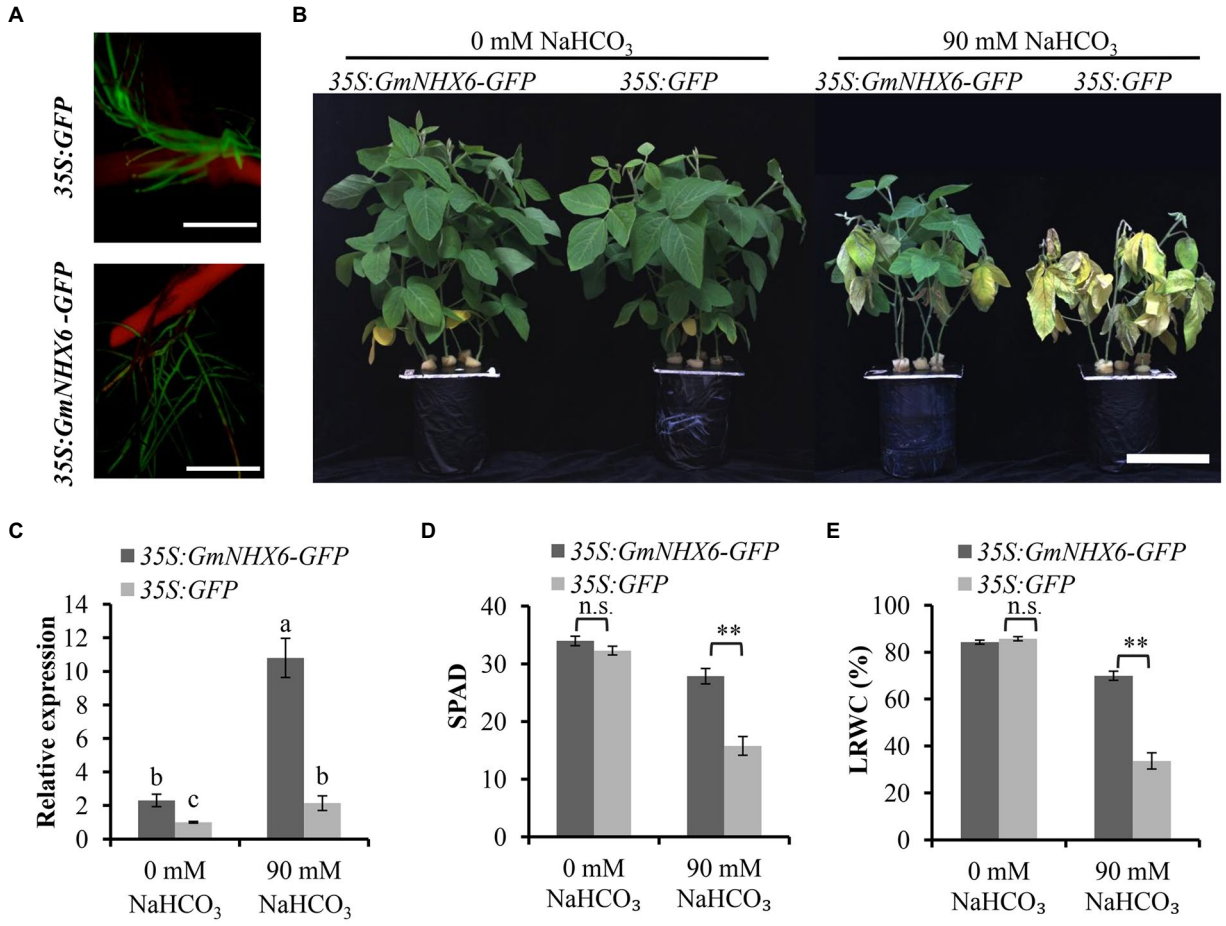


FIGURE 3
Alkaline salt tolerance analyses of transgenic soybean composite plants. **(A)** Identification of positive transgenic soybean lines by green fluorescence using a stereoscopic fluorescence microscope. Bar=1cm. **(B)** Representative phenotype of transgenic soybean composite plants (25-day-old) under 0mM or 90mM $NaHCO_3$ for 7 days. Bar=12cm. **(C)** Relative expression of *GmNHX6* gene in the roots of soybean composite plants at 24h after 0 or 90mM $NaHCO_3$ treatment by qRT-PCR. For relative expression calculation, the sample from soybean composite plants with empty vector (35S:*GFP*) under 0mM $NaHCO_3$ was used as control and *GmUKN1* was the reference gene. Data represents the mean±standard deviation of three biological replications with three repeats within each replication (n=3×3=9). Bars with the same letter in lowercase above bars indicate no significant difference according to Duncan's multiple range test at 0.05 level. **(D,E)** The Soil and Plant Analysis Development (SPAD) value for chlorophyll content, and leaf relative water content (LRWC) of transgenic soybean composite plants under 0mM or 90mM $NaHCO_3$ for 7 days, respectively. Data represents the mean±standard deviation of three biological replications and each repeat contained five independent transgenic plants for each genotype (n=3×5=15). Soybean variety of "TianLong1" was used. Differences were evaluated using the two-tailed Student's t-tests (n.s., not significant; $^{**}p$<0.01).

These results suggested that *GmNHX6* might play an important role in soybean response to alkaline salt stress.

Overexpression of *GmNHX6* improved soybean tolerance to $NaHCO_3$ stress

Since a greater increase in the expression level of *GmNHX6* was observed in alkaline salt tolerant soybean variety than the sensitive variety after $NaHCO_3$ treatment, we next tested whether overexpression of *GmNHX6* could improve soybean tolerance to alkaline salt stress or not. The coding region of *GmNHX6* gene was expressed in fusion with *GFP* (while the empty vector 35S:*GFP* was used as control), and transformed into soybean hypocotyls to obtain transgenic composite plants, in the genetic background of an alkaline salt sensitive soybean variety Tianlong 1. The positive transgenic composite soybean plants were identified through GFP fluorescence signal in roots (Figure 3A). Under normal condition (0mM $NaHCO_3$), all soybean composite plants grew well, with no obvious difference (Figure 3B; Supplementary Figures 5A,E,I). When the transgenic soybean composite plants were treated with 90mM $NaHCO_3$ for 7 days, the 35S:*GFP*-transformed soybean composite plants showed obviously inhibited growth, leaf wilting and chlorosis, while the 35S:*GmNHX6-GFP*-transformed composite plants had much less damage (Figure 3B; Supplementary Figures 5A,E,I). The expression level of *GmNHX6* gene was higher in 35S:*GmNHX6-GFP*-transformed soybean plants than the 35S:*GFP*-transformed plants either under 0 or 90mM $NaHCO_3$, and *GmNHX6* gene expression

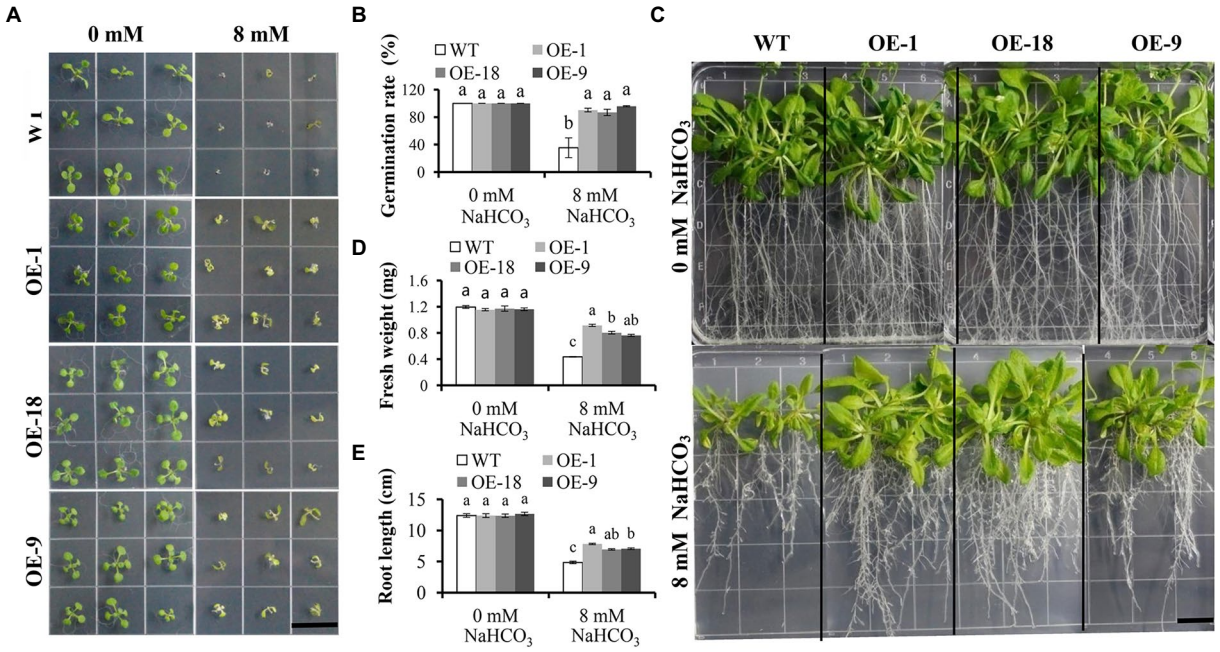


FIGURE 4
Effect of *GmNHX6* overexpression on *Arabidopsis* tolerance to $NaHCO_3$ treatment. **(A)** Seed germination assay of different *Arabidopsis* lines on 1/2 MS medium and 1/2 MS supplied with 8mM $NaHCO_3$. Photographs were taken 7 days after $NaHCO_3$ treatment. WT: wild type; OE-1, OE-18, OE-9: *Arabidopsis* lines overexpressing *GmNHX6*. Bar=1.5cm. **(B)** Seed germination rates of different *Arabidopsis* lines as shown in A. Data represents the mean±standard deviation of three biological replications and each repeat contained 36 seeds per line for each treatment (n=36×3=108). **(C)** Phenotypes of *Arabidopsis* lines subjected to 0 and 8mM $NaHCO_3$ treatment, respectively. Seeds were grown vertically on 1/2 MS agar plates supplemented with 0 or 8mM $NaHCO_3$ for 10days after seven days of normal growth. Bar=1.5cm. **(D,E)** Fresh weight and root length of *Arabidopsis* lines as shown in **(C)**. Data represents the mean±standard deviation of three biological replications and each repeat contained four plants per line for each treatment (n=12). Bars with the same letter in lowercase above bars indicate no significant differences at the 0.05 level between lines under same treatment according to Duncan's multiple range tests.

was significantly upregulated at 24 h post 90 mM $NaHCO_3$ treatment in 35S:*GmNHX6-GFP*-transformed soybean hairy roots (Figure 3C; Supplementary Figures 5B,F,J). The average SPAD value for chlorophyll content (Figure 3D; Supplementary Figures 5C,G,K) and the leaf relative water content (LRWC; Figure 3E; Supplementary Figures 5D,H,L) of 35S:*GmNHX6-GFP*-transformed soybean composite plants were significantly higher than those of the control plants transformed by the 35S:*GFP* under 90 mM $NaHCO_3$ treatment. No significant differences were found in SPAD and LRWC between 35S:*GmNHX6-GFP*-transformed plants and 35S:*GFP*-transformed plants under normal condition (0 mM $NaHCO_3$). These results demonstrated that overexpression of *GmNHX6* reduced the damage of $NaHCO_3$ treatment on soybean plants.

Ectopic expression of *GmNHX6* in *Arabidopsis* enhanced the alkaline salt tolerance by maintaining low Na^+/K^+ ratios

We further investigate the role of *GmNHX6* in response to alkaline salt stress by transgenic *Arabidopsis* plants. *GmNHX6* was overexpressed (OE) using the CaMV 35S promoter in *Arabidopsis*. Seven homozygous transgenic *Arabidopsis* lines containing 35S:*GmNHX6* were obtained at T_3 generation. Then three lines with higher *GmNHX6* expression levels (OE-1, OE-18, and OE-9) were selected for further analyses (Supplementary Figure 6). The germination rates of *Arabidopsis* lines were compared in the absence or presence of $NaHCO_3$, respectively. The results showed that there was no difference in germination rates between the wild type (WT) and *GmNHX6* OE lines under normal condition without $NaHCO_3$, but the germination rates of *GmNHX6* OE lines were significantly higher than that of WT *Arabidopsis* under $NaHCO_3$ stress (Figures 4A,B). To further test the effect of $NaHCO_3$ stress on *Arabidopsis*, we measured the fresh weight and root length of *Arabidopsis*. Under normal condition, no significant difference was observed between *GmNHX6* OE lines and WT (Figures 4C–E). After 10 days of 8 mM $NaHCO_3$ treatment, *GmNHX6* OE lines showed better growth than WT (Figure 4C): the average fresh weight and root length of *GmNHX6* OE lines were significantly larger than those of WT *Arabidopsis* (Figures 4D,E). These results suggest that overexpression of *GmNHX6* enhanced the alkaline salt tolerance in *Arabidopsis*.

We also compared the content of Na^+, K^+, and the Na^+/K^+ ratio in the roots and leaves of *GmNHX6* OE lines with WT *Arabidopsis*

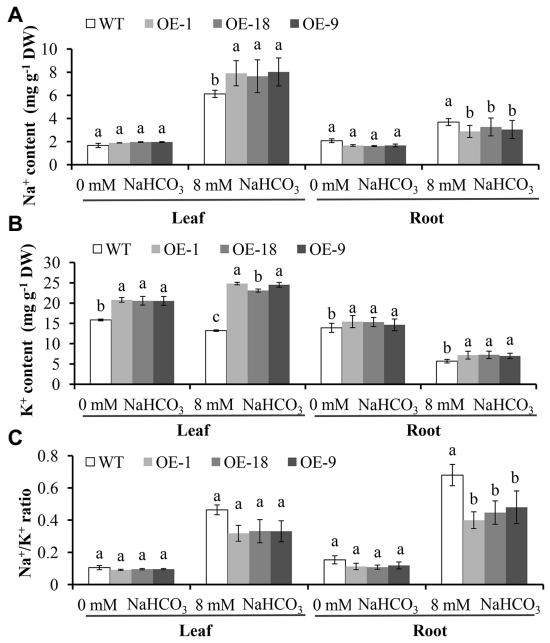


FIGURE 5
Na$^+$ content **(A)**, K$^+$ content **(B)**, and Na$^+$/K$^+$ ratio **(C)** in the leaves and roots of *Arabidopsis* lines. *Arabidopsis* lines receiving 0mM or 8mM $NaHCO_3$ treatment for 10days, respectively. Data represents the mean±standard deviation of three biological replications and each repeat contained 10 plants per line for each treatment (n=30). Bars with the same letter in lowercase above bars indicate no significant differences at the 0.05 level between lines under same treatment according to Duncan's multiple range tests.

under $NaHCO_3$ treatment and normal conditions (Figure 5). For Na$^+$, under normal conditions, its content in leaf or root was similar between all lines. After $NaHCO_3$ treatment, the leaf Na$^+$ content in the *GmNHX6* OE lines was significantly higher than that in the WT *Arabidopsis*, but the Na$^+$ content in root showed the opposite pattern that OE lines had lower Na$^+$ content than WT (Figure 5A). For K$^+$, its content in both leaf and root of the *GmNHX6* OE lines was significantly higher than that in the WT before and after $NaHCO_3$ treatment (Figure 5B). For Na$^+$/K$^+$ ratios, there was no significant difference between *GmNHX6* OE lines and WT under normal conditions in both leaves and roots. However, the Na$^+$/K$^+$ ratio of *GmNHX6* OE lines was lower than that in the WT (Figure 5C). These results suggested that overexpression of *GmNHX6* could improve the alkaline salt tolerance by keeping high K$^+$ content and maintaining low Na$^+$/K$^+$ ratio in plants.

Natural variation in the promoter region of soybean *NHX6* is associated with alkaline salt tolerance

We sequenced the coding and promoter regions of soybean *NHX6* from 30 wild soybean (15 alkaline salt tolerant and 15 alkaline salt sensitive soybean accessions) and 30 cultivated soybean (15 alkaline salt tolerant and 15 alkaline salt sensitive), respectively (Supplementary Table 1; Supplementary Figure 7). The results showed that only the *NHX6* promoter region had sequence variation among 60 soybean accessions. A total of 30 SNPs were found in the 2-kb promoter region of soybean *NHX6* (Supplementary Figure 8). Then, a regional association study was performed using the 30 SNPs in the 60 soybean accessions with extreme sodic tolerance rating (STR). Only one SNP (SNP_{-560}, 560 bp upstream of the start codon) showed significant association with STR by mixed model in EMMAX (Supplementary Figure 9; Supplementary Table 4). Among the 60 soybean accessions, the majority (25 out of 30, 83%) of alkaline salt tolerant accessions had a base "C" at SNP_{-560} upstream of *GmNHX6*, while 70% (21 out of 30) alkaline salt sensitive accessions had a base "T" at SNP_{-560} (Supplementary Figure 8). There was a significant difference in alkaline salt tolerance between the two groups of SNP_{-560}-C and SNP_{-560}-T in 60 soybean accessions, as well as in wild soybean accessions and cultivated soybean accessions (Figure 6A; Supplementary Figure 10). In order to further compare the two types of promoters containing SNP_{-560}-C and SNP_{-560}-T, the representative promoters were cloned from two soybean varieties, M8206 and ZY, and designated as Pro-M8206 (containing SNP_{-560}-C) and Pro-ZY (containing SNP_{-560}-T), respectively. The promoter-LUC transient expression assays in tobacco leaves revealed that the activity of Pro-M was significantly stronger than that of Pro-ZY under alkali stress but not control condition (Figures 6B,C), suggesting that Pro-M is more responsive to alkali stress than Pro-ZY, therefore, leading to higher alkali induced expression level of *GmNHX6* in the alkaline salt tolerant soybean accessions than sensitive accessions, thus resulting in enhanced alkaline salt tolerance.

Discussion

Soil salinization and alkalization can reduce soil osmotic potential, and cause ion imbalance, disrupt physiological processes, inhibit growth and development of plants, leading to a serious decline in its yield and quality, and even the death of plants (Zhu, 2001; Yang et al., 2011). Soil salinization and alkalization frequently co-occur, but in the saline-alkaline land in mainland China, the soil salinization caused by alkali salts such as $NaHCO_3$ and Na_2CO_3 is more serious than that caused by neutral salts such as NaCl and Na_2SO_4 (Yang et al., 2008, 2011). In general, the stress factors of neutral salts are mainly the ion stress of Na$^+$ and osmotic stress of low water potential caused by high salt concentration, but for the alkaline salts, there is an added factor of high pH (Shi et al., 1998). The alkaline soil causes damage to plants not only through salt stress but also through high pH (Shi and Wang, 2005). Therefore, the deleterious effect of high pH stress or salinity alone is significantly less than that of the combined stress of high pH and salinity (Li et al., 2010). There have been lots of progress on plant tolerance to salt stress, while only few reports on alkali salt

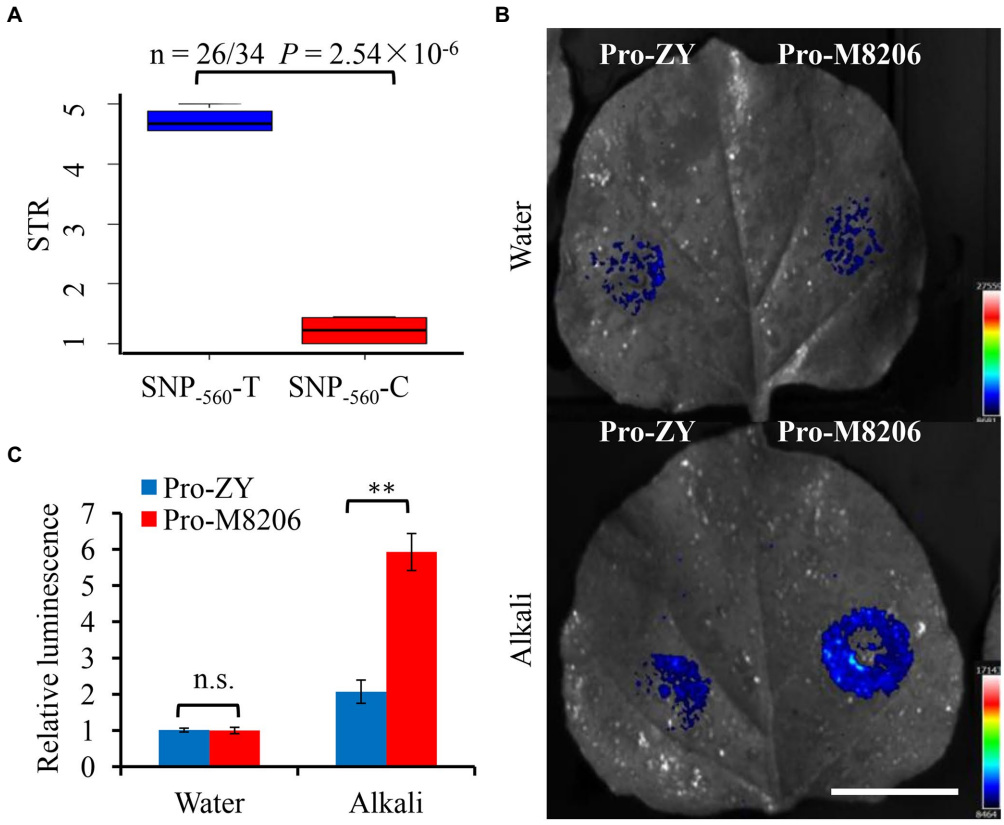


FIGURE 6
Natural variation and activity of soybean *NHX6* promoter. **(A)** Boxplot of STR for two alleles of SNP_{-560} in *GmNHX6* promoter among 60 soybean accessions. *n* denotes number of accessions. Statistical significance was detected by two-sided Wilcoxon test. The center bold line represents the median; box edges indicate the upper and lower quantiles; whiskers show the 1.5 × interquartile range. STR: sodic (alkaline salt) tolerance rating. **(B,C)** Promoter activities of two types of *GmNHX6* promoters by luciferase (LUC) transient expression assays in tobacco leaves after alkaline salt (100mM $NaHCO_3$) treatment for 16h. Pro-M8206 contains SNP_{-560}-C and Pro-ZY contains SNP_{-560}-T. The *LUC* reporter gene was driven by each type of promoter. The photos were taken using an *in vivo* plant imaging system. Bar=3cm. Data represents the mean±standard deviation of four biological replications with two repeats within each replication (n=4×2=8). Differences were evaluated using the two-tailed Student's *t*-tests (**p<0.01; n.s., not significant).

tolerance. Here, we investigated the role of soybean *NHX6* gene in plant tolerance to alkaline salt ($NaHCO_3$), which would broaden our knowledge on alkali salt tolerance.

In plants, sodium hydrogen exchangers allow cation/H^+ such as Na^+/H^+ to transmembrane transport, to regulate pH and maintain ion homeostasis to resist abiotic stress (Rodríguez-Rosales et al., 2009; Bassil et al., 2012). However, there is no report about its role in soybean tolerance to alkali stress. In this study, a sodium/hydrogen exchanger gene from the soybean, *GmNHX6*, was found to be involved in the alkaline salt tolerance of soybean. The expression of *GmNHX6* was induced by $NaHCO_3$ treatment and more upregulated in alkali salt tolerant soybean (Figure 2), suggesting it is involved in plant responses to alkaline stress. Transgenic soybean composite plants and *Arabidopsis* plants have been previously used to study gene functions in saline-alkaline tolerance (Al-Harrasi et al., 2020; Ketehouli et al., 2021). *GmPKS4* overexpressing soybean composite plants and transgenic *Arabidopsis* plants had increased proline content as well as high antioxidant enzyme activities under salt and salt-alkali stress treatments, compared to the empty-vector-transformed or wild-type ones (Ketehouli et al., 2021). Then we employed transgenic soybean composite plants and *Arabidopsis* to investigate the role of *GmNHX6* in alkaline salt tolerance (Figures 3–5; Supplementary Figure 5). We found that overexpression of *GmNHX6* enhanced the alkaline salt tolerance of soybean and *Arabidopsis*.

Since *NHX* genes have been shown to play important roles in salt tolerance in a variety of plant species (Kobayashi et al., 2012; Al-Harrasi et al., 2020; Long et al., 2020; Sun et al., 2021), we further analyzed the function of *GmNHX6* in salt tolerance. The germination rates of *GmNHX6* OE lines were significantly higher than that of WT *Arabidopsis* under NaCl stress (Figures 7A,B). At seedling stage, overexpression of *GmNHX6* in *Arabidopsis* increased the average fresh weight and root length under salt stress (Figures 7C–E). We also compared the contents of Na^+ and K^+ in the roots and leaves of *Arabidopsis* (Figures 7F–H). Under normal conditions, there is no significant difference in the content of Na^+ content between WT and OE lines. However, under NaCl treatment, the Na^+ content in the leaves of the *GmNHX6* OE lines was significantly higher than that in

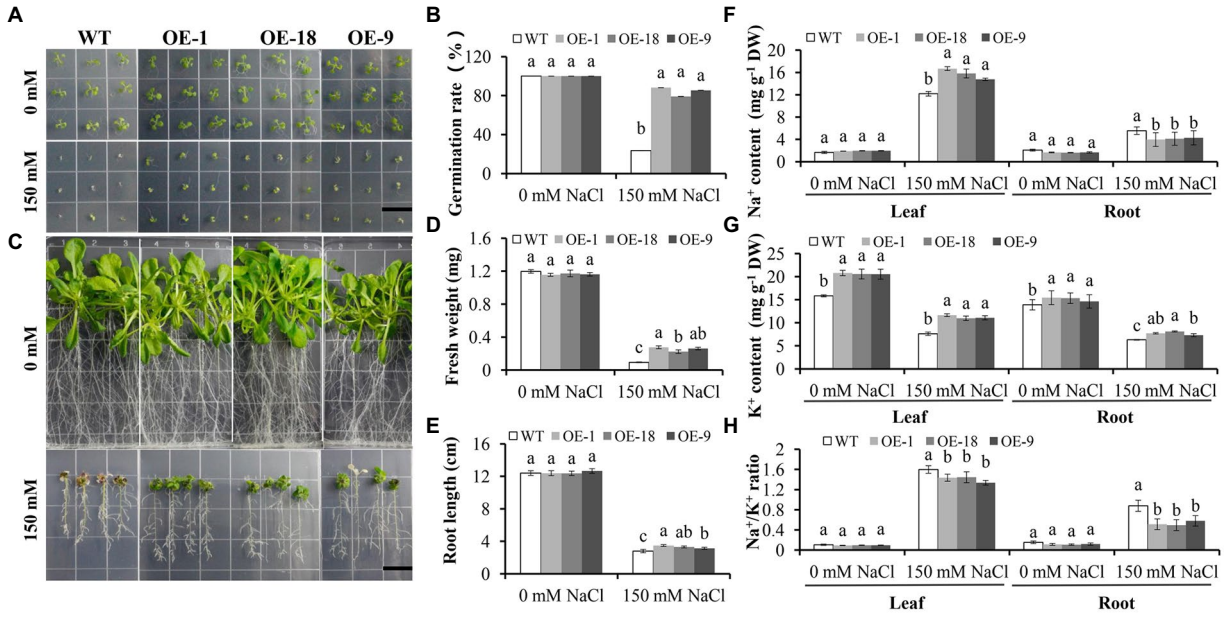


FIGURE 7
The effect of *GmNHX6* overexpression on the salt tolerance of *Arabidopsis*. **(A)** Seed germination assay of different *Arabidopsis* lines on 1/2 MS medium supplied with 0 or 150mM NaCl. Photographs were taken 7 days after treatment. WT: wild type; OE-1, OE-18, OE-9: *Arabidopsis* lines overexpressing *GmNHX6*. Bar=1.5cm. **(B)** Seed germination rates of different *Arabidopsis* lines as shown in A. Data represents the mean±standard deviation of three biological replications and each repeat contained 36 seeds per line for each treatment (n=36×3=108). **(C)** *Arabidopsis* seeds were grown vertically on 1/2 MS agar plates supplemented with 0 or 150mM NaCl treatment for 10days after 7 days of normal growth, respectively. Bar=1.5cm. **(D,E)** The fresh weight and root length of *Arabidopsis* lines shown in **(C)**. Data represents the mean±standard deviation of three biological replications and each repeat contained four plants per line for each treatment (n=12). **(F)** Na^+ content, **(G)** K^+ content, **(H)** Na^+/K^+ ratio in the leaves and roots of *Arabidopsis* lines shown in **(C)**. For **(F–H)**, data represents the mean±standard deviation of three biological replications and each replication has four repeats (n=12), and 10 plants were pooled together for each sample. Bars with the same letter in lowercases above bars indicate no significant differences at the 0.05 level between lines under each condition according to Duncan's multiple range test.

the WT, but in the root Na^+ content was just the opposite (Figure 7F). The K^+ content of *GmNHX6* OE lines was significantly higher than that in the WT *Arabidopsis* in both leaves and roots (Figure 7G). The Na^+/K^+ ratios of the *GmNHX6* OE lines were significantly lower than that in the WT *Arabidopsis* under NaCl treatment (Figure 7H). These results suggest that overexpression of *GmNHX6* enhanced salt tolerance by maintaining Na^+ and K^+ homeostasis.

Different from salt stress, alkali stress caused by $NaHCO_3$ and Na_2CO_3 leads to higher pH (Fang et al., 2021). According to the previous studies, the NHXs are important regulators of cellular pH and ion homeostasis (Rodríguez-Rosales et al., 2009; Bassil et al., 2012). NHX-type antiporters utilize the H^+ electrochemical gradient to facilitate the exchange of H^+ for cations such as Na^+ or K^+, thereby maintaining both pH and ion homeostasis (Sze and Chanroj, 2018). Studies also showed that plants can secrete large amounts of organic acids under alkali stress, which can play a buffer role and maintain intracellular pH stability (Fang et al., 2021). However, further experiments are needed to investigate the role of *GmNHX6* in adjusting pH to alleviate alkaline salt stress.

K^+ is an essential macronutrient for plant growth and the most abundant inorganic cation in plant cells. During plant growth and development, K^+ is involved in the activation of more than 60 enzyme systems, photosynthesis, carbohydrate metabolism and protein synthesis, what's more, K^+ is a key modulator for cell homeostasis (Zhang et al., 2019a; Kumar et al., 2020). Na^+ toxicity is one of the main harmful factors of saline-alkali stress (Apse and Blumwald, 2007). In root cells, Na^+ is compartmentalized into vacuoles, radial transported to the stele cells, and loaded into the xylem, thus establishing the homeostatic control of Na^+ in the cytosol (Apse and Blumwald, 2007). Plants can maintain high K^+ content by increasing K^+ absorption and reducing K^+ loss (Zhu, 2003). Under saline-alkali stress, plant cells absorb a large amount of Na^+, resulting in ion toxicity, which not only inhibits photosynthesis, but also inhibits the absorption of essential element such as K^+ by plants, when Na^+/K^+ is too high, resulting in the change of ion homeostasis and metabolic disorder (Zhu, 2003). Therefore, maintaining a relatively high K^+ content and low Na^+/K^+ ratio in plants under salt-alkali stress is a manifestation of salt-alkali tolerance. The NHXs play important roles in maintaining Na^+ and K^+ homeostasis have been well documented (Wang et al., 2015; Sun et al., 2021). Previous studies have shown that NHXs can compartmentalize Na^+ into vacuoles under salt stress, thus maintaining the intracellular ion balance (Feng et al., 2021). Other studies found that vacuolar NHXs can simultaneously catalyze the exchange of Na^+/H^+ and K^+/H^+ to maintain intracellular ion balance (Sze and Chanroj, 2018). Under salt stress, NHX transporter, such as NHX1, transport sodium ions to and sequester them in the vacuole

by vesicle transport, and release K^+ from vacuolar into the cytoplasm (Barragán et al., 2012). In *Arabidopsis, AtNHX5* and *AtNHX6* are critical to K^+ homeostasis in *Arabidopsis* (Wang et al., 2015). *GmNHX5* positively regulates salt tolerance by maintaining higher K^+/Na^+ ratio in soybean (Sun et al., 2021). In this study, we found that overexpression of *GmNHX6* enhanced *Arabidopsis* tolerance to both alkaline salt ($NaHCO_3$) and salt (NaCl) stress, by maintaining higher K^+ content and low Na^+/K^+ ratios.

We also obtained a homozygous mutant *nhx6* (stock, *SALK-100042C*), a T-DNA insertion *Arabidopsis* mutant of *AtNHX6* (Supplementary Figure 11). However, *nhx6* mutant did not show alkaline salt tolerance or salt tolerance compared with the WT *Arabidopsis* (Supplementary Figure 12), which is consistent with the previous study by Bassil et al. (2011a) that the phenotype of *nhx5* or *nhx6* single-knockout line was not different from the WT, while the double mutant *nhx5 nhx6* has reduced growth, smaller and fewer cells, and increased sensitivity to salinity. Through the EnsemblPlants database,[9] we identified a total of 66 genes in the soybean *NHX* gene family, which might be functionally redundant in soybean tolerance to salt-alkali stress.

Soybean germplasm provides a wide range of saline-alkaline tolerance. To explore the natural variation in soybean *NHX6* gene, we analyzed the sequence polymorphism of the 2-kb promoter region (Supplementary Figure 13) of *NHX6* in 60 soybean accessions, including *G. max* and *G. soja*. Only one SNP (SNP_{-560}) out of a total of 30 SNPs showed significant association with alkaline salt tolerance (STR) by mixed model in EMMAX (Supplementary Table 4). We analyzed the *cis*-elements of the 2-kb promoter region (Supplementary Figure 13) of *GmNHX6* in two soybean accessions, M8206 and ZY. We found that in Pro-M8206 (containing SNP_{-560}-C), the sequence forms four cis-acting elements around SNP_{-560}, including Skn-1_motif (GTCAT, −562 to −558 bp), TCATTT element (−561 to −556 bp), TTGTCA motif (−564 to −559 bp), and TGTCAT motif (−563 to −559 bp). Skn-1_motif is responsive to abiotic stress and could be used in plant genetic engineering research on abiotic stress tolerance (De Silva et al., 2017; Zhang et al., 2018). TTGTCA element is involved in light induction (Mazouni et al., 2003). Previous studies showed that TCATTT element and TGTCAT element are the binding sites of downstream key functional genes. TCATTT-containing element acted as an enhancer (Perrotti et al., 1996), and was essential for inducible expression of the *IL-5* gene (Miao et al., 2006). Tunicamycin (Tm)-activated sXBP1 bound to the TGTCAT element and suppressed *XRCC2* expression to prevent tumor proliferation *in vivo* (Zhao et al., 2021). Interesting, promoter-LUC transient expression assays in tobacco leaves revealed that the promoter of *GmNHX6* from the alkaline salt tolerant soybean variety, Pro-M8206, had a significantly stronger activity than Pro-ZY (from alkaline salt sensitive variety), under alkali stress (Figures 6B,C), suggesting that Pro-M8206 is more responsive to alkali stress than Pro-ZY. Therefore, it is likely that the SNP_{-560} might affect the promoter activity under alkali stress by interrupting the relevant *cis*-acting elements mentioned above. How these *cis*-acting elements in the promoter of *GmNHX6* regulate gene expression in response to alkali stress needs further study.

9 http://plants.ensembl.org/biomart/martview/

Conclusion

In summary, *GmNHX6* encodes a sodium/hydrogen exchanger. The expression of *GmNHX6* was induced by $NaHCO_3$ stress, and greater increase in its transcript abundance was observed in alkaline salt tolerant than in alkaline salt sensitive soybean variety. Overexpression of *GmNHX6* enhanced the alkaline-salt tolerance of soybean composite plants and *Arabidopsis*. The *GmNHX6* overexpressing *Arabidopsis* lines had higher K^+ content and lower Na^+/K^+ ratio than the wild-type plants under $NaHCO_3$ stress. Furthermore, a single nucleotide polymorphism in the promoter region of *NHX6* is associated with the alkali tolerance in soybean. These findings would help to further understand the role of *NHX6* and its regulatory mechanism in soybean tolerance to alkaline-salt stress.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary material.

Author contributions

TJ and YL conceived and designed the experiments, interpreted the results, and wrote and revised the manuscript. TJ, JA, HX, JC, LP, RZ, and NW performed the experiments. TJ analyzed the data and generated the pictures. JG and YL contributed reagents and materials. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Key Research and Development Program of China (2021YFF1001204) and the Core Technology Development for Breeding Program of Jiangsu Province (JBGS-2021-014).

Acknowledgments

We appreciate Xinyuan Huang and Daolong Dou at Nanjing Agricultural University to provide us the vectors of pGreenII0800-LUC and pBinGFP4.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.938635/full#supplementary-material

References

Al-Harrasi, I., Jana, G. A., Patankar, H. V., Al-Yahyai, R., Rajappa, S., Kumar, P. P., et al. (2020). A novel tonoplast Na^+/H^+ antiporter gene from date palm (*PdNHX6*) confers enhanced salt tolerance response in *Arabidopsis*. *Plant Cell Rep.* 39, 1079–1093. doi: 10.1007/s00299-020-02549-5

An, R., Chen, Q. J., Chai, M. F., Lu, P. L., Su, Z., Qin, Z. X., et al. (2007). *AtNHX8*, a member of the monovalent cation: proton antiporter-1 family in *Arabidopsis thaliana*, encodes a putative Li^+/H^+ antiporter. *Plant J.* 49, 718–728. doi: 10.1111/j.1365-313X.2006.02990.x

Apse, M. P., and Blumwald, E. (2007). Na^+ transport in plants. *FEBS Lett.* 581, 2247–2254. doi: 10.1016/j.febslet.2007.04.014

Barragán, V., Leidi, E. O., Andrés, Z., Rubio, L., De Luca, A., Fernández, J. A., et al. (2012). Ion exchangers NHX1 and NHX2 mediate active potassium uptake into vacuoles to regulate cell turgor and stomatal function in *Arabidopsis*. *Plant Cell* 24, 1127–1142. doi: 10.1105/tpc.111.095273

Bassil, E., Coku, A., and Blumwald, E. (2012). Cellular ion homeostasis: emerging roles of intracellular NHX Na^+/H^+ antiporters in plant growth and development. *J. Exp. Bot.* 63, 5727–5740. doi: 10.1093/jxb/ers250

Bassil, E., Ohto, M. A., Esumi, T., Tajima, H., Zhu, Z., Cagnac, O., et al. (2011a). The *Arabidopsis* intracellular Na^+/H^+ antiporters NHX5 and NHX6 are endosome associated and necessary for plant growth and development. *Plant Cell* 23, 224–239. doi: 10.1105/tpc.110.079426

Bassil, E., Tajima, H., Liang, Y. C., Ohto, M. A., Ushijima, K., Nakano, R., et al. (2011b). The *Arabidopsis* Na^+/H^+ antiporters NHX1 and NHX2 control vacuolar pH and K^+ homeostasis to regulate growth, flower development, and reproduction. *Plant Cell* 23, 3482–3497. doi: 10.1105/tpc.111.089581

Bassil, E., Zhang, S., Gong, H., Tajima, H., and Blumwald, E. (2019). Cation specificity of vacuolar NHX-type cation/H^+ antiporters. *Plant Physiol.* 179, 616–629. doi: 10.1104/pp.18.01103

Bouras, H., Choukr-Allah, R., Amouaouch, Y., Bouaziz, A., Devkota, K. P., El Mouttaqi, A., et al. (2022). How does quinoa (*Chenopodium quinoa* Willd.) respond to phosphorus fertilization and irrigation water salinity? *Plant Theory* 11:216. doi: 10.3390/plants11020216

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308

Cao, B., Xia, Z., and Liu, C. (2020). New insights into the structure-function relationship of the endosomal-type Na^+, K^+/H^+ antiporter NHX6 from mulberry (*Morus notabilis*). *Int. J. Mol. Sci.* 21:428. doi: 10.3390/ijms21020428

Chen, C., Yu, Y., Ding, X., Liu, B., Duanmu, H., Zhu, D., et al. (2018). Genome-wide analysis and expression profiling of PP2C clade D under saline and alkali stresses in wild soybean and *Arabidopsis*. *Protoplasma* 255, 643–654. doi: 10.1007/s00709-017-1172-2

Clough, S. J., and Bent, A. F. (1998). Floral dip: a simplified method for *agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16, 735–743. doi: 10.1046/j.1365-313x.1998.00343.x

De Silva, W. S. I., Perera, M. M. N., Perera, K., Wickramasuriya, A. M., and Jayasekera, G. A. U. (2017). *In silico* analysis of *osr40c1* promoter sequence isolated from Indica variety Pokkali. *Rice Sci.* 24, 228–234. doi: 10.1016/j.rsci.2016.11.002

Dragwidge, J. M., Scholl, S., and Schumacher, K. (2019). NHX-type Na^+, K^+/H^+ antiporters are required for TGN/EE trafficking and endosomal ion homeostasis in *Arabidopsis thaliana*. *J. Cell Sci.* 132:jcs226472. doi: 10.1242/jcs.226472

Fan, L., Zhao, L., Hu, W., Li, W., Novák, O., Strnad, M., et al. (2018). Na^+, K^+ $/H^+$ antiporters regulate the pH of endoplasmic reticulum and auxin-mediated development. *Plant Cell Environ.* 41, 850–864. doi: 10.1111/pce.13153

Fang, S., Hou, X., and Liang, X. (2021). Response mechanisms of plants under saline-alkali stress. *Front. Plant Sci.* 12:667458. doi: 10.3389/fpls.2021.667458

Feng, J., Ma, W., Ma, Z., Ren, Z., and Zhou, Y. (2021). *$GhNHX_3D$*, a vacuolar-localized Na^+/H^+ antiporter, positively regulates salt response in upland cotton. *Int. J. Mol. Sci.* 22:4047. doi: 10.3390/ijms22084047

Guan, R., Qu, Y., Guo, Y., Yu, L., Liu, Y., Jiang, J., et al. (2014). Salinity tolerance in soybean is modulated by natural variation in *GmSALT3*. *Plant J.* 80, 937–950. doi: 10.1111/tpj.12695

He, Y., Dong, Y., Yang, X., Guo, D., Qian, X., Yan, F., et al. (2020). Functional activation of a novel R2R3-MYB protein gene, *GmMYB68*, confers salt-alkali resistance in soybean (*Glycine max* L.). *Genome* 63, 13–26. doi: 10.1139/gen-2018-0132

He, C., Gao, H., Wang, H., Guo, Y., He, M., Peng, Y., et al. (2021). GSK3-mediated stress signaling inhibits legume-rhizobium symbiosis by phosphorylating GmNSP1 in soybean. *Mol. Plant* 14, 488–502. doi: 10.1016/j.molp.2020.12.015

Jin, T., Sun, Y., Shan, Z., He, J., Wang, N., Gai, J., et al. (2021). Natural variation in the promoter of *GsERD15B* affects salt tolerance in soybean. *Plant Biotechnol. J.* 19, 1155–1169. doi: 10.1111/pbi.13536

Kereszt, A., Li, D., Indrasumunar, A., Nguyen, C. D., Nontachaiyapoom, S., Kinkema, M., et al. (2007). *Agrobacterium rhizogenes*-mediated transformation of soybean to study root biology. *Nat. Protoc.* 2, 948–952. doi: 10.1038/nprot.2007.141

Ketehouli, T., Zhou, Y. G., Dai, S. Y., Carther, K. F. I., Sun, D. Q., Li, Y., et al. (2021). A soybean calcineurin B-like protein-interacting protein kinase, GmPKS4, regulates plant responses to salt and alkali stresses. *J. Plant Physiol.* 256:153331. doi: 10.1016/j.jplph.2020.153331

Kobayashi, S., Abe, N., Yoshida, K. T., Liu, S., and Takano, T. (2012). Molecular cloning and characterization of plasma membrane- and vacuolar-type Na^+/H^+ antiporters of an alkaline-salt-tolerant monocot, Puccinellia tenuiflora. *J. Plant Res.* 125, 587–594. doi: 10.1007/s10265-012-0475-9

Krishnamurthy, P., Vishal, B., Khoo, K., Rajappa, S., Loh, C. S., and Kumar, P. P. (2019). Expression of *AoNHX1* increases salt tolerance of rice and *Arabidopsis*, and bHLH transcription factors regulate *AtNHX1* and *AtNHX6* in *Arabidopsis*. *Plant Cell Rep.* 38, 1299–1315. doi: 10.1007/s00299-019-02450-w

Kumar, P., Kumar, T., Singh, S., Tuteja, N., Prasad, R., and Singh, J. (2020). Potassium: a key modulator for cell homeostasis. *J. Biotechnol.* 324, 198–210. doi: 10.1016/j.jbiotec.2020.10.018

Li, R., Shi, F., and Fukuda, K. (2010). Interactive effects of various salt and alkali stresses on growth, organic solutes, and cation accumulation in a halophyte *Spartina alterniflora* (Poaceae). *Environ. Exp. Bot.* 68, 66–74. doi: 10.1016/j.envexpbot.2009.10.004

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Long, L., Zhao, J. R., Guo, D. D., Ma, X. N., Xu, F. C., Yang, W. W., et al. (2020). Identification of NHXs in *Gossypium* species and the positive role of *GhNHX1* in salt tolerance. *BMC Plant Biol.* 20:147. doi: 10.1186/s12870-020-02345-z

Lv, S., Wang, L., Zhang, X., Li, X., Fan, L., Xu, Y., et al. (2020). *Arabidopsis* NHX5 and NHX6 regulate PIN6-mediated auxin homeostasis and growth. *J. Plant Physiol.* 255:153305. doi: 10.1016/j.jplph.2020.153305

Mazouni, K., Domain, F., Chauvat, F., and Cassier-Chauvat, C. (2003). Expression and regulation of the crucial plant-like ferredoxin of cyanobacteria. *Mol. Microbiol.* 49, 1019–1029. doi: 10.1046/j.1365-2958.2003.03609.x

Miao, Y., Yan, P. K., Kim, H., Hwang, I., and Jiang, L. (2006). Localization of green fluorescent protein fusions with the seven *Arabidopsis* vacuolar sorting receptors to prevacuolar compartments in tobacco BY-2 cells. *Plant Physiol.* 142, 945–962. doi: 10.1104/pp.106.083618

Miao, L., Yang, S., Zhang, K., He, J., Wu, C., Ren, Y., et al. (2020). Natural variation and selection in *GmSWEET39* affect soybean seed oil content. *New Phytol.* 225, 1651–1666. doi: 10.1111/nph.16250

Moghaieb, R. E., Sharaf, A. N., Soliman, M. H., El-Arabi, N. I., and Momtaz, O. A. (2014). An efficient and reproducible protocol for the production of salt tolerant transgenic wheat plants expressing the *Arabidopsis AtNHX1* gene. *GM Crops Food* 5, 132–138. doi: 10.4161/gmcr.28941

Munns, R., James, R. A., Xu, B., Athman, A., Conn, S. J., Jordans, C., et al. (2012). Wheat grain yield on saline soils is improved by an ancestral Na^+transporter gene. *Nat. Biotechnol.* 30, 360–364. doi: 10.1038/nbt.2120

Oh, D. H., Lee, S. Y., Bressan, R. A., Yun, D. J., and Bohnert, H. J. (2010). Intracellular consequences of SOS1 deficiency during salt stress. *J. Exp. Bot.* 61, 1205–1213. doi: 10.1093/jxb/erp391

Pérez-Patricio, M., Camas-Anzueto, J. L., Sanchez-Alegría, A., Aguilar-González, A., and Gutiérrez-Miceli, F. (2018). Optical method for estimating the chlorophyll contents in plant leaves. *Sensors* 18:650. doi: 10.3390/s18020650

Perrotti, D., Bellón, T., Trotta, R., Martinez, R., and Calabretta, B. (1996). A cell proliferation-dependent multiprotein complex NC-3A positively regulates the CD34 promoter via a TCATTT-containing element. *Blood* 88, 3336–3348. doi: 10.1182/blood.V88.9.3336.bloodjournal8893336

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 29:e45. doi: 10.1093/nar/29.9.e45

Qiu, Q. S. (2016a). Plant endosomal NHX antiporters: activity and function. *Plant Signal. Behav.* 11:e1147643. doi: 10.1080/15592324.2016.1147643

Qiu, Q. S. (2016b). AtNHX5 and AtNHX6: roles in protein transport. *Plant Signal. Behav.* 11:e1184810. doi: 10.1080/15592324.2016.1184810

Quintero, F. J., Blatt, M. R., and Pardo, J. M. (2000). Functional conservation between yeast and plant endosomal Na^+/H^+ antiporters. *FEBS Lett.* 471, 224–228. doi: 10.1016/s0014-5793(00)01412-5

Reguera, M., Bassil, E., and Blumwald, E. (2014). Intracellular NHX-type cation/H^+ antiporters in plants. *Mol. Plant* 7, 261–263. doi: 10.1093/mp/sst091

Reguera, M., Bassil, E., Tajima, H., Wimmer, M., Chanoca, A., Otegui, M. S., et al. (2015). pH regulation by NHX-type antiporters is required for receptor-mediated protein trafficking to the vacuole in *Arabidopsis*. *Plant Cell* 27, 1200–1217. doi: 10.1105/tpc.114.135699

Ren, Y., Lv, J., Wang, H., Li, L., Peng, Y., and Qu, L. J. (2009). A comparative proteomics approach to detect unintended effects in transgenic *Arabidopsis*. *J. Genet. Genomics* 36, 629–639. doi: 10.1016/s1673-8527(08)60155-1

Rodríguez-Rosales, M. P., Gálvez, F. J., Huertas, R., Aranda, M. N., Baghour, M., Cagnac, O., et al. (2009). Plant NHX cation/proton antiporters. *Plant Signal. Behav.* 4, 265–276. doi: 10.4161/psb.4.4.7919

Sellamuthu, G., Jegadeeson, V., Sajeevan, R. S., Rajakani, R., Parthasarathy, P., Raju, K., et al. (2020). Distinct evolutionary origins of intron retention splicing events in *NHX1* antiporter transcripts relate to sequence specific distinctions in *Oryza* species. *Front. Plant Sci.* 11:267. doi: 10.3389/fpls.2020.00267

Shi, D. C., Sheng, Y. M., and Zhao, K. F. (1998). Stress effects of mixed salts with various salinities on the seedlings of *Aneurolepidium chinense*. *Acta Bot. Sin.* 40, 1136–1142.

Shi, D., and Wang, D. (2005). Effects of various salt-alkaline mixed stresses on *Aneurolepidium chinense* (Trin.) Kitag. *Plant and Soil* 271, 15–26. doi: 10.1007/s11104-004-1307-z

Sun, T., Ma, N., Wang, C., Fan, H., Wang, M., Zhang, J., et al. (2021). A golgi-localized sodium/hydrogen exchanger positively regulates salt tolerance by maintaining higher K^+/Na^+ ratio in soybean. *Front. Plant Sci.* 12:638340. doi: 10.3389/fpls.2021.638340

Sze, H., and Chanroj, S. (2018). Plant endomembrane dynamics: studies of K^+/H^+ antiporters provide insights on the effects of pH and ion homeostasis. *Plant Physiol.* 177, 875–895. doi: 10.1104/pp.18.00142

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197

Tse, Y. C., Mo, B., Hillmer, S., Zhao, M., Lo, S. W., Robinson, D. G., et al. (2004). Identification of multivesicular bodies as prevacuolar compartments in *Nicotiana tabacum* BY-2 cells. *Plant Cell* 16, 672–693. doi: 10.1105/tpc.019703

Tuyen, D. D., Lal, S. K., and Xu, D. H. (2010). Identification of a major QTL allele from wild soybean (*Glycine soja* Sieb. & Zucc.) for increasing alkaline salt tolerance in soybean. *Theor. Appl. Genet.* 121, 229–236. doi: 10.1007/s00122-010-1304-y

Walker, D. J., Leigh, R. A., and Miller, A. J. (1996). Potassium homeostasis in vacuolate plant cells. *Proc. Natl. Acad. Sci. U. S. A.* 93, 10510–10514. doi: 10.1073/pnas.93.19.10510

Wang, L., Wu, X., Liu, Y., and Qiu, Q. S. (2015). AtNHX5 and AtNHX6 control cellular K^+ and pH homeostasis in *Arabidopsis*: three conserved acidic residues are essential for K^+ transport. *PLoS One* 10:e0144716. doi: 10.1371/journal.pone.0144716

Wang, L., Yang, Y., Zhang, S., Che, Z., Yuan, W., and Yu, D. (2020). GWAS reveals two novel loci for photosynthesis-related traits in soybean. *Mol. Genet. Genomics* 295, 705–716. doi: 10.1007/s00438-020-01661-1

Wu, X., Ebine, K., Ueda, T., and Qiu, Q. S. (2016). AtNHX5 and AtNHX6 are required for the subcellular localization of the SNARE complex that mediates the trafficking of seed storage proteins in *Arabidopsis*. *PLoS One* 11:e0151658. doi: 10.1371/journal.pone.0151658

Wu, F. H., Shen, S. C., Lee, L. Y., Lee, S. H., Chan, M. T., and Lin, C. S. (2009). Tape-*Arabidopsis* Sandwich - a simpler *Arabidopsis* protoplast isolation method. *Plant Methods* 5:16. doi: 10.1186/1746-4811-5-16

Yang, C. W., Wang, P., Li, C. Y., Shi, D. C., and Wang, D. L. (2008). Comparison of effects of salt and alkali stresses on the growth and photosynthesis of wheat. *Photosynthetica* 46, 107–114. doi: 10.1007/s11099-008-0018-8

Yang, J., Zaitlen, N. A., Goddard, M. E., Visscher, P. M., and Price, A. L. (2014). Advantages and pitfalls in the application of mixed-model association methods. *Nat. Genet.* 46, 100–106. doi: 10.1038/ng.2876

Yang, J. Y., Zheng, W., Tian, Y., Wu, Y., and Zhou, D. W. (2011). Effects of various mixed salt-alkaline stresses on growth, photosynthesis, and photosynthetic pigment concentrations of *Medicago ruthenica* seedlings. *Photosynthetica* 49, 275–284. doi: 10.1007/s11099-011-0037-8

Zhang, Z., Cheng, Z. J., Gan, L., Zhang, H., Wu, F. Q., Lin, Q. B., et al. (2016). *OsHSD1*, a hydroxysteroid dehydrogenase, is involved in cuticle formation and lipid homeostasis in rice. *Plant Sci.* 249, 35–45. doi: 10.1016/j.plantsci.2016.05.005

Zhang, Y., Deng, G., Fan, W., Yuan, L., and Wang, H. (2019b). *NHX1* and *eIF4A1*-stacked transgenic sweetpotato shows enhanced tolerance to drought stress. *Plant Cell Rep.* 38, 1427–1438. doi: 10.1007/s00299-019-02454-6

Zhang, X., Li, Z., Li, X., Xu, Y., Xie, H., and Qiu, Q. S. (2020b). CBL3 and CIPK18 are required for the function of NHX5 and NHX6 in mediating Li^+ homeostasis in *Arabidopsis*. *J. Plant Physiol.* 255:153295. doi: 10.1016/j.jplph.2020.153295

Zhang, M., Liang, X., Wang, L., Cao, Y., Song, W., and Shi, J. (2019a). A HAK family Na^+ transporter confers natural variation of salt tolerance in maize. *Nat. Plants* 5, 1297–1308. doi: 10.1038/s41477-019-0565-y

Zhang, S., Tajima, H., Nambara, E., and Blumwald, E. (2020a). Auxin homeostasis and distribution of the auxin efflux carrier PIN2 require vacuolar NHX-type cation/H^+ antiporter activity. *Plant Theory* 9:1311. doi: 10.3390/plants9101311

Zhang, F., and Zhang, J. (2021). Bulk analysis by resequencing and RNA-seq identifies candidate genes for maintaining leaf water content under water deficit in maize. *Physiol. Plant.* 173, 1935–1945. doi: 10.1111/ppl.13537

Zhang, H., Zheng, J., Su, H., Xia, K., Jian, S., and Zhang, M. (2018). Molecular cloning and functional characterization of the Dehydrin (*IpDHN*) gene from *Ipomoea pes-caprae*. *Front. Plant Sci.* 9:1454. doi: 10.3389/fpls.2018.01454

Zhao, Z., He, K., Zhang, Y., Hua, X., Feng, M., Zhao, Z., et al. (2021). XRCC2 repairs mitochondrial DNA damage and fuels malignant behavior in hepatocellular carcinoma. *Cancer Lett.* 512, 1–14. doi: 10.1016/j.canlet.2021.04.026

Zheng, J., Lin, R., Pu, L., Wang, Z., Mei, Q., and Zhang, M. (2021). Ectopic expression of *CrPIP2;3*, a plasma membrane intrinsic protein gene from the halophyte *Canavalia rosea*, enhances drought and salt-alkali stress tolerance in *Arabidopsis*. *Int. J. Mol. Sci.* 22:565. doi: 10.3390/ijms22020565

Zhu, J. K. (2001). Plant salt tolerance. *Trends Plant Sci.* 6, 66–71. doi: 10.1016/S1360-1385(00)01838-0

Zhu, J. K. (2003). Regulation of ion homeostasis under salt stress. *Curr. Opin. Plant Biol.* 6, 441–445. doi: 10.1016/s1369-5266(03)00085-2

Zhu, D., Bai, X., Chen, C., Chen, Q., Cai, H., Li, Y., et al. (2011). *GsTIFY10*, a novel positive regulator of plant tolerance to bicarbonate stress and a repressor of jasmonate signaling. *Plant Mol. Biol.* 77, 285–297. doi: 10.1007/s11103-011-9810-0

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 20 September 2022
DOI 10.3389/fpls.2022.942217

OPEN ACCESS

EDITED BY
Mukesh Jain,
Jawaharlal Nehru University, India

REVIEWED BY
Xuan Yao,
Huazhong Agricultural University, China
Miao Liu,
Guizhou University, China

*CORRESPONDENCE
Jian Zhang
yjnkyy@ntu.edu.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Plant Abiotic Stress, a section of the journal Frontiers in Plant Science

RECEIVED 12 May 2022
ACCEPTED 15 August 2022
PUBLISHED 20 September 2022

CITATION
Yu C, Ke Y, Qin J, Huang Y, Zhao Y, Liu Y, Wei H, Liu G, Lian B, Chen Y, Zhong F and Zhang J (2022) Genome-wide identification of calcineurin B-like protein-interacting protein kinase gene family reveals members participating in abiotic stress in the ornamental woody plant *Lagerstroemia indica*. *Front. Plant Sci.* 13:942217. doi: 10.3389/fpls.2022.942217

Genome-wide identification of calcineurin B-like protein-interacting protein kinase gene family reveals members participating in abiotic stress in the ornamental woody plant *Lagerstroemia indica*

Chunmei Yu[1,2†], Yongchao Ke[1†], Jin Qin[1], Yunpeng Huang[1], Yanchun Zhao[1], Yu Liu[1], Hui Wei[1,2], Guoyuan Liu[1,2], Bolin Lian[1,2], Yanhong Chen[1,2], Fei Zhong[1,2] and Jian Zhang[1,2*]

[1]School of Life Sciences, Nantong University, Nantong, China, [2]Key Laboratory of Landscape Plant Genetics and Breeding, Nantong University, Nantong, China

Calcineurin B-like protein-interacting protein kinases (CIPKs) play important roles in plant responses to stress. However, their function in the ornamental woody plant *Lagerstroemia indica* is remains unclear. In this study, the *LiCIPK* gene family was analyzed at the whole genome level. A total of 37 *LiCIPKs,* distributed across 17 chromosomes, were identified. Conserved motif analysis indicated that all LiCIPKs possess a protein kinase motif (S_TKc) and C-terminal regulatory motif (NAF), while seven LiCIPKs lack a protein phosphatase interaction (PPI) motif. 3D structure analysis further revealed that the N-terminal and C-terminal 3D-structure of 27 members are situated near to each other, while 4 members have a looser structure, and 6 members lack intact structures. The intra- and interspecies collinearity analysis, synonymous substitution rate (K_s) peaks of duplicated *LiCIPKs,* revealed that ~80% of *LiCIPKs* were retained by the two whole genome duplication (WGD) events that occurred approximately 56.12–61.16 million year ago (MYA) and 16.24–26.34 MYA ago. The promoter of each *LiCIPK* contains a number of auxin, abscisic acid, gibberellic acid, salicylic acid, and drought, anaerobic, defense, stress, and wound responsive *cis*-elements. Of the 21 members that were successfully amplified by qPCR, 18 *LiCIPKs* exhibited different expression patterns under NaCl, mannitol, PEG8000, and ABA treatments. Given that *LiCIPK30,* the *AtSOS2* ortholog, responded to all four types of stress it was selected for functional verification. *LiCIPK30* complements the *atsos2* phenotype *in vivo*. 35S:LiCIPK-overexpressing lines exhibit increased leaf area increment, chlorophyll a and b content, reactive oxygen species scavenging enzyme activity, and expression of *ABF3* and *RD22*, while the degree of membrane lipid oxidation decreases under NaCl treatment compared to WT. The evolutionary history, and potential mechanism by which *LiCIPK30* may

regulate plant tolerance to salt stress were also discussed. In summary, we identified *LiCIPK* members involved in abiotic stress and found that *LiCIPK30* transgenic Arabidopsis exhibits more salt and osmotic stress tolerance than WT. This research provides a theoretical foundation for further investigation into the function of *LiCIPKs,* and for mining gene resources to facilitate the cultivation and breeding of new *L. indica* varieties in coastal saline-alkali soil.

KEYWORDS

CIPKs, *Lagerstroemia indica*, gene family, abiotic stress, overexpression, salt tolerance, Arabidopsis

Introduction

Stability of coastal ecosystems is vital for minimizing the destruction of sea winds and tides. However, high salt concentrations of coastal soil, as well as a deficiency or imbalance in inorganic nutrients, such as nitrogen (N), phosphorus (P), and potassium (K), can impact plant growth. Hence, plants have evolved various mechanisms to adapt to changes in their environment. Indeed, plants have evolved calcium signaling to regulate development, plant–microbe interactions, and environmental signal (e.g., abiotic stress) perception (Ghosh et al., 2021). More specifically, calcineurin B-like proteins (CBLs) and CBL-interacting protein kinases (CIPKs, a Ser/Thr protein kinase), are plant-specific calcium sensors with numerous functions. For instance, the salt over sensitive (SOS) pathway, which has four main components (SOS1, SOS2, SOS3, and SOS3-like calcium binding protein 8/SCaBP8), is a basic calcium signaling pathway that has been elucidated in higher plants under salt stress (Halfter et al., 2000; Liu et al., 2000; Guo et al., 2001; Gong et al., 2002; Qiu et al., 2002; Quan et al., 2007). In this pathway, AtSOS3/CBL10 (or SCaBP8) functions as a Ca^{2+} sensor, binding to AtCIPK24/SOS2 to form an active complex, it then phosphorylates SOS1—an Na^+/H^+ exchanger located on the cell membrane—to regulate Na^+ exclusion by the cell. Additionally, other CIPKs, such as maize ZmCBL1/4-ZmCIPK42, regulate salt stress tolerance at the seedling stage (Chen et al., 2021). For example, GmPKS4, a soybean CIPK, regulates soybean responses to salt and alkali stresses (Ketehouli et al., 2021). Besides salt stress, a CaCIPK3 drought response cassette was identified in pepper (*Capsicum annuum* L.) (Ma et al., 2021). AtCIPK23, through combining different CBLs, regulates the uptake and homeostasis of different ions, such as nitrate nitrogen (NO_3^-), ammonium (NH_4^+), K^+, Mg^{2+}, and Fe^{2+} (Cheong et al., 2007; Ragel et al., 2015; Tian et al., 2016; Straub et al., 2017; Morales de Los Ríos et al., 2021; Ródenas and Vert, 2021). OsCIPK9 is a multifaceted kinase that responds to salinity, osmotic stress, and K^+ deficiency in rice (Ketehouli et al., 2021). The CIPKs are also involved in root architecture formation for various plants, such as ZmCIPK15 in maize, *zmcipk15*-knockout mutant exhibited a steeper root growth angle, higher nitrogen absorption, and greater shoot biomass compared to the WT. Meanwhile, overexpression of chrysanthemum *CmCIPK23* in Arabidopsis significantly decreases lateral root number and length, primary root length, and nitrogen uptake (Liu et al., 2022). The *zmcipk42* mutant also has fewer branched tassels and reduced salt stress tolerance at the seedling stage (Chen et al., 2021). Additionally, *OsCIPK31* participates in the development of panicle apical spikelets (Peng et al., 2018). Generally, orthologous genes have similar functions among different species, such as the SOS2 orthologs among different plant species. However, certain examples have been reported, for example, the overexpression of *VaCIPK02* (Amur grape), the *AtCIPK6* ortholog, enhances salt sensitivity in Arabidopsis (Xu et al., 2020) while overexpression of chickpea (*Cicer arietinum*) *CaCIPK6* enhances salt tolerance (Tripathi et al., 2009). The sequence difference between orthologous proteins may lead to the recruitment of different partners or downstream targets. Therefore, functions of orthologs among different plants require further experimental verification.

Due to three common whole genome duplication (WGD) events of dicotyledonous plants (ζ, zeta seed plant-wide WGD; ε, epsilon angiosperm-wide WGD event; and γ, gamma triplicated of dicotyledon-wide WGD) (Tang et al., 2008; Clark and Donoghue, 2018; Li and Barker, 2020), the retention of CIPKs during the evolution of a plant and their role in plant adaption represents an interesting research direction. In fact, genome-wide analysis of the *CIPK* gene family in several plants, including Arabidopsis, rice, grape, *Prunus mume*, tea (*Camellia sinensis* var. Sinensis), and turnip (*Brassica rapa* var. *rapa*), has identified members participating in abiotic stress (Kolukisaoglu et al., 2004; Yu et al., 2007; Xi et al., 2017; Li et al., 2019). For example, in a horticultural/ornamental plant, *P. mume*,

Abbreviations: CIPKs, CBL-interacting protein kinases; ROS, reactive oxygen species; CBLs, calcineurin B-like proteins; SOS, salt over sensitive; WGD, whole genome duplication; WT, Arabidopsis, Col-0; T-DNA SALK_056101C, *atsos2* mutant; COM, LiCIPK30/*atsos2*; OE, LiCIPK30/WT; qPCR, quantitative polymerase chain reaction; pI, isoelectric point; MW, molecular weight; NJ, neighbor-joining; CAT, catalase; POD, peroxidase; MAD, malondialdehyde; RBOH, respiratory burst oxidase homolog; *GOlS2*, galactinol synthase 2; PPI, protein phosphatase interaction.

a total of 16 CIPK genes were identified. Twelve *PmCIPK* genes are up-regulated during cold stress treatment, implying that *PmCIPKs* may be involved in distribution of different *P. mume* varieties (Li et al., 2019). In turnip, 51 *BrrCIPK* genes have diverse expression patterns during development and different stimulation with several CIPKs found to have more than one CLB partner (Yin et al., 2017). These studies indicate that genome-wide analysis can identify stress-related *CIPKs* efficiently.

Lagerstroemia indica (crape myrtle) is an important ornamental shrub (tree) characterized by its long flowering period, different flower color, and bark trunks. It is also a traditional medicinal plant with several effective secondary metabolites (Yang et al., 2011; Labib et al., 2013). During the past decade, research has primarily focused on genes determining ornamental traits, such as leaf and flower color (Wang et al., 2017; Qiao et al., 2019; Li et al., 2020a; Yu et al., 2021), as well as the dwarf and weeping architecture (Ye et al., 2015; Li et al., 2020b). However, during its long life, *L. indica* faces several challenges, including the threat of diseases and soil stress (nutrient and water shortages, high pH, or high salt in some coastal areas). Studies have been undertaken to determine the pathogenicity of leaf spot and powdery mildew (Shi and Mmbaga, 2006; Babu et al., 2014; Kim, 2021), to identify genes resistant to powdery mildew (Wang et al., 2015). However, there are currently few reports on the response of *L. indica* to salt stress, the existence of salt-tolerant varieties, and genetic resources that could be used to improve salt stress tolerance during the breeding program. To our knowledge, the functions of *LiCIPKs* have not been reported yet; thus, rendering an incomplete understanding of the mechanism by which crape myrtle responds to external stress signals.

As CIPKs are the homeostat of several ions in plants, we sought to determine the roles of LiCIPKs in crape myrtle adaptation to abiotic stress. More specifically, we assessed *LiCIPK* gene family characteristics, evolutionary history, and expression patterns under several abiotic stress conditions, and defined their associated functions. Our objectives are to: (1) better understand the *L. indica* genome; (2) identify LiCIPK members that participate in abiotic stress; and (3) verify LiCIPK function. Our results provide a theoretical foundation for further functional analysis of CIPKs in plant adaptation to stress.

Materials and methods

Plant material, growth conditions and chemical reagents

Plant materials in this study include the *L. indica* var. Black Diamond 'Blush V2' and four genotypes of Arabidopsis, Col-0 (WT), *atsos2* mutant (T-DNA SALK_056101C), LiCIPK30/*atsos2* (COM) and LiCIPK30/WT (OE). All Arabidopsis lines were germinated on half-strength Murashige & Skoog media supplied with 2% sucrose (1/2 MS) for 10–14 days, then the seedlings were planted on mixed soil (50% Pindstrup Substrate and 50% vermiculite) and grown in an artificial climate chamber (16 h day/8 h night, 22°C day/18°C night).

The chemical reagents used in this study were purchased from SINOPHARM (Beijing, China) or Ameresco (Framingham, MA, United States). The RNA extraction reagent (MiniBEST Plant RNA Extraction Kit), first strand cDNA synthesis kit (PrimeScript^{TM} RT reagent Kit with gDNA Eraser), and TB GREEN reagents for quantitative polymerase chain reaction (qPCR) were purchased from TaKaRa (Beijing, China). The plant genomic DNA extraction kit was purchased from TIANGEN (DP305, Beijing, China), and plasmid DNA extraction kit from Beyotime (Shanghai, China). Primers were synthesized by General Bio (China, Anhui, Chuzhou). All oligo primers used in this study are listed in **Supplementary Table 1**.

Identification of *L. indica* calcineurin B-like protein-interacting protein kinase genes

The HMM files of the protein kinase (PF00069) and NAF motif (PF03822) were downloaded from the pfam protein database[1]. The candidate LiCIPKs in the *L. indica* genome (Accession number CNP0003018, unpublished data from our lab) were obtained by HMM search (*E*-value < 1e-5, Identity ≥ 50%) using the TBtools software (V1.098689) (Chen et al., 2020). Candidates were also aligned using BLAST against the AtCIPKs (Kolukisaoglu et al., 2004) and filtered according to the methods described by Zhang et al. (2021). The distribution of *LiCIPKs* on chromosomes was depicted by three files (chromosome length, Gene ID, and GFF3) using TBtools (Chen et al., 2020). The theoretical isoelectric point (pI) and molecular weight (MW) of LiCIPKs were predicted using ExPASy[2]. Subcellular location was estimated using Wolf PSORT[3] and SignalP[4].

Gene structure and conserved motifs analysis

The intron–exon structures of *LiCDPKs* were obtained using the genome sequences of *L. indica*, GFF3 files, and CDS sequences of all *LiCDPKs*. The conserved LiCDPK sequences

1 https://pfam.xfam.org
2 http://web.expasy.org/protparam/
3 https://wolfpsort.hgc.jp/
4 http://www.cbs.dtu.dk/services/SignalP/

were analyzed using MEME program with the parameters: motif 10 and width between 6–100 amino acid residues[5]. Graphic of genes structure and conserved motifs were drawn using TBtools software (Chen et al., 2020).

Three-dimensional (3D) structure prediction of LiCIPKs

The 3D structures of LiCIPKs were predicted on the SWISS-MODEL website through homologous modeling[6]. The pdb files were opened using chimera soft[7] and the structure of LiCIPKs was compared to that of AtCIPK24/SOS2 reported previously (Sánchez-Barrena et al., 2007; Chaves-Sanjuan et al., 2014).

Calcineurin B-like protein-interacting protein kinase phylogenic tree construction

Arabidopsis AtCIPKs (*Arabidopsis thaliana, At*), rice OsCIPKs (*Oryza sativa, Os*), and grape VvCIPKs (*Vitis vinifera, Vv*) were extracted from their genomes according to previous reports (Kolukisaoglu et al., 2004; Yu et al., 2007; Kanwar et al., 2014; Xi et al., 2017). The amino acid sequences of CIPKs from four plants were used to construct a neighbor-joining (NJ) phylogenetic tree by MEGA X soft using default parameters (Kumar et al., 2017). The phylogenic tree was designed by web-based soft Evolgenius[8] (Zhang et al., 2012; He et al., 2016; Subramanian et al., 2019).

The synteny of calcineurin B-like protein-interacting protein kinase loci among three species

Synteny of CIPK gene loci between Arabidopsis, *L. indica,* and grape (*V. vinifera*) was analyzed using the One Step MCScanX in TBtools. Cognate loci intra *L. indica* was analyzed using the advanced Circos in TBtools.

Divergence time calculation of duplicated genes

After obtaining duplicated gene pairs, the synonymous substitution rate (K_s) and non-synonymous substitution rate (K_a) of gene pairs were calculated using the "Simple K_a/K_s calculator" tool of TBtools. Based on the *Lythraceae* specific rate (λ) of 1.14×10^{-8} substitutions per site per year (Feng et al., 2021), the divergence time (million years ago, MYA) of duplicated gene pairs was calculated according to the formula $T = K_s/2\lambda$. Two common rates, 1.5×10^{-8} or 6.1×10^{-9}, were also used as references (Lynch and Conery, 2000; Blanc and Wolfe, 2004).

Cis-elements analysis of *LiCIPK* promoters

Cis elements in the promoter (−2,000 bp upstream ATG) were predicted on the PlantCARE website[9]. A combined diagram of LiCIPK phylogenic tree and *cis*-element distribution was drawn using the TBtools software (Chen et al., 2020).

Stress treatment of *L. indica*

L. indica var. Black Diamond 'Blush V2' was used in this study. Under normal conditions, the semi-hardwood healthy branches were cut into 10–12 cm long pieces with at least four axillary buds to form cuttings, which were sterilized using 0.0625% carbendazim for 15–20 min, and subsequently planted in vermiculite soil in pots. All cuttings were grown in a greenhouse at 25 ∼ 35°C, under 16 h day/8 h night conditions (2020–2021). When new adventitious roots had grown to a length of ∼2–3 cm (at least 45 days after planting), they were treated with irrigation water containing 200 mmol L^{-1} NaCl, 15% PEG8000 and 10×10^{-4} mol L^{-1} abscisic acid (ABA), or 200 mmol L^{-1} mannitol to induce salt, drought, or osmotic stress (three biological repeats each), respectively (Liu X. et al., 2020). After treatment for 0, 1, 2, 3, 4, 5, and 6 days, adventitious roots were collected for RNA extraction, performed on the same day.

RNA extraction and cDNA synthesis

RNA from the roots of all stress-treated materials was extracted using the MiniBEST Plant RNA Extraction Kit (TaKara). The RNA was converted to first strand cDNA using the PrimeScript^{TM} RT reagent Kit with the gDNA Eraser kit. All procedures were performed according to the manufacturer's instructions.

Real-time quantitative PCR

Primers used for qRT-PCR are listed in **Supplementary Table 1**. The program was performed using ABI7500 according

5 http://meme-suite.org/tools/meme
6 https://swissmodel.expasy.org
7 http://www.cgl.ucsf.edu/chimera/download.html
8 https://evolgenius.info/
9 http://bioinformatics.psb.ugent.be/webtools/plantcare/html/

to the manufacturer's instructions. PCR mixes were made following the protocols of the TB GREEN kit. The expression levels were calculated using $2^{-\Delta\Delta Ct}$ compared to the internal control and CK sample (Yu et al., 2021).

Cloning of *LiCIPK30* and vector construction

Leaf RNA of *L. indica* and first-strand cDNA were obtained using the methods described above. *LiCIPK30* was amplified by 2 × Pfu MasterMix (CWBIO, CW0686, Beijing, China) using the primers listed in **Supplementary Table 1**. The PCR products were cloned into a pGEMT-T easy vector (Promega, Shanghai, China), according to the manufacturer's protocols. The pWM101-*35S:LiCIPK30* was constructed using an in-fusion strategy (ClonExpress II One Step Cloning Kit, C112-01, Vazyme, Nanjing, China).

Transformation of Arabidopsis

The pWM101-*35S:LiCIPK30* construct was first transformed into Arabidopsis WT (Col-0) and at*sos2* mutants through agrobacterium-mediated (GV3101) floral dip method reported previously (Clough and Bent, 1998). The positive LiCIPK30/*atsos2* (COM) and LiCIPK30/WT (OE) T1 plants were screened using half strength (1/2) MS with 20 mg L^{-1} hygromycin and genomic PCR with *LiCIPK30*-specific primers (**Supplementary Table 1**).

Salt and mannitol stress treatment of Arabidopsis

The WT, *atsos2,* T3 Arabidopsis lines of COM and OE were treated at two different developmental stages. At germination, the seeds were planted in 1/2 MS medium with 0, 75, and 125 mmol L^{-1} NaCl or 0, 100, and 200 mmol L^{-1} mannitol. Root length of different lines was observed 10 days after planting. At the rosette stage (6–8 full-size leaves), the plants were treated by irrigating with 0, 100, and 200 mmol L^{-1} NaCl solution. The leaves were harvested for further analysis after 0, 3, and 7 days of treatment. All treatments were applied to at least three replicates.

Detection physiological parameters

Leaves of WT and two independent LiCIPK30 lines were homogenized using liquid nitrogen. The activity of catalase (CAT, A007-1-1), and peroxidase (POD, A084-3-1), as well as the content of malondialdehyde (MAD, A003-3-1), and chlorophyll a and b (A147-1-1) were detected using the respective kits according to the manufacturer's instructions (NJJCBIO, Nanjing, China). All treatments were applied to at least three replicates.

Data analysis

Graphs of the data were constructed using Origin 2018 (Originlab, MA, United States). Differences between treatments or genotypes were analyzed on SPSS23.0 (IBM SPSS, NY, United States) using *t*-test or two-way analysis of variance (ANOVA).

Results

Identification and characterization of LiCIPKs

LiCIPKs were identified through HMM and BLAST search using the amino acid sequence of 26 AtCIPKs. After filtering the amino acid sequences based on length limits, a total of 37 full length proteins with a NAF domain and high sequence similarity to AtCIPKs were identified as LiCIPK family members. The name, chromosome location, peptide length, MW, and subcellular locations are listed in **Table 1**. The theoretical pI varied from 6.11 (LiCIPK3) to 9.33 (LiCIPK21), indicating different residues on the protein surface which may recruit different partners *in vivo*. LiCIPKs were predicted to localize to the cytoplasm (7 members), nuclear compartment (5 members), chloroplast (20 members), endoplasmic reticulum (1 members), and cytoskeleton (1 member). This variable distribution of LiCIPK family members implies that they may be involved in multiple biological processes (**Table 1**).

From the exon–intron patterns of *LiCIPKs,* we found 21 *LiCIPKs* to be intron-less (two introns or less) accounting for 56.76% of the total CIPKs. All others were intron-rich (containing 11–14 introns) genes. The exon-intron patterns were similar to that of the *VvCIPKs* of grape and *CsCIPK* of tea (Yin et al., 2017; Liu et al., 2019) (**Figure 1A**), implying that the gene structure of CIPKs diverged before the evolution of these species.

The MEME motif analysis showed that most LiCIPks had ten conserved motifs, including motifs 1 and 2 in the N-terminal kinase domain, which are in all LiCIPKs (**Figure 1B** and **Supplementary Figure 1**), and the C-terminal-regulating NAF/FISL domain in motif 8, which was identified in 35 LiCIPKs. This result differs from that of the multiple-alignment of the N-end and C-end active domains, which indicates that all CIPKs possess NAF/FISL sequences (**Figures 1B,C** and **Supplementary Figure 1**). Motif 7 is the protein phosphatase interaction (PPI) domain, which exists in 30 LiCIPKs (**Figure 1**

TABLE 1 Characteristic of LiCIPKs.

Subgroup	Gene name	Gene ID	Chr	AAlength	PI	MW(KDa)	Predicted localization
Group Aa	*LiCIPK18*	evm.model.Chr8.403.2	8	452	8.03	51.59	Chl
	LiCIPK23	evm.model.Chr11.440	11	439	6.46	50.38	Chl
	LiCIPK31	evm.model.Chr15.863	15	439	6.86	50.43	Chl
	LiCIPK35	evm.model.Chr22.156.2	22	464	8.99	53.11	Chl
	LiCIPK36	evm.model.Chr23.882	23	439	6.56	50.17	Chl
Group Ab	*LiCIPK9*	evm.model.Chr4.1714	4	449	8.74	50.42	Chl
	LiCIPK25	evm.model.Chr12.477	12	465	8.94	51.5	Chl
Group Ac	*LiCIPK26*	evm.model.Chr12.934	12	448	6.69	50.86	Cyt
	LiCIPK16	evm.model.Chr7.12	7	446	9.07	50.31	Mit
	LiCIPK30	evm.model.Chr15.86.2	15	492	8.94	55.46	Chl
Group Ad	*LiCIPK8*	evm.model.Chr4.815	4	454	7.55	50.85	Chl
	LiCIPK11	evm.model.Chr5.959	5	443	6.22	50.09	Cyt
Group Ae	*LiCIPK2*	evm.model.Chr2.353	2	436	6.87	49.04	Nuc
	LiCIPK3	evm.model.Chr3.750.1	3	467	6.11	52.81	Chl
	LiCIPK17	evm.model.Chr7.317	7	436	6.65	48.9	NuC
	LiCIPK29	evm.model.Chr15.77	15	466	7.06	51.56	Nuc
Group B	*LiCIPK1*	evm.model.Chr2.66	2	436	8.93	47.8	Chl
	LiCIPK5	evm.model.Chr3.1533	3	432	9.1	47.58	Chl
	LiCIPK27	evm.model.Chr13.774	13	435	9	48.14	Chl
	LiCIPK34	evm.model.Chr21.370	21	433	9.2	47.84	Chl
Group Ca	*LiCIPK6*	evm.model.Chr4.21	4	448	9.13	50.53	Chl
	LiCIPK14	evm.model.Chr5.1704	5	448	8.86	50.42	Cyt
Group Cb	*LiCIPK4*	evm.model.Chr3.994	3	442	9.04	49.53	Cyt
	LiCIPK24	evm.model.Chr11.1405	11	443	8.8	49.74	PM
	LiCIPK28	evm.model.Chr15.32	15	450	9.06	50.54	Chl
Group Cc	*LiCIPK15*	evm.model.Chr6.1537	6	461	8.8	52.26	Cyt
	LiCIPK33	evm.model.Chr17.458	17	448	9.04	50.96	Cyt
Group Cd	*LiCIPK37*	evm.model.Chr24.188	24	462	9.02	52.01	Chl
	LiCIPK21	evm.model.Chr9.1257	9	335	9.33	38.29	Chl
	LiCIPK7	evm.model.Chr4.131	4	466	9.2	52.61	Nuc
	LiCIPK13	evm.model.Chr5.1568	5	462	9.01	52.1	Chl
Group D	*LiCIPK12*	evm.model.Chr5.1567	5	441	8.39	49.46	Cyt
	LiCIPK20	evm.model.Chr9.1255	9	451	9.13	50.9	Cyt
	LiCIPK22	evm.model.Chr11.4	11	440	9.09	49.56	Nuc
	LiCIPK32	evm.model.Chr15.1211	15	439	8.89	49.31	CytS
Group E	*LiCIPK10*	evm.model.Chr5.724	5	476	8.6	53.47	ER
	LiCIPK19	evm.model.Chr9.246	9	476	8.46	53.62	Chl

Chl, chloroplast; ER, endoplasmic reticulum; Nuc, nucleus; Mit, mitochondria; Cyt, cytoplasm; CytS, cytoskeleton; PM, plasma membrane. MW, molecular weight. pI, isoelectric point. Groups of the *LiCIPKs* were divided by results of phylogenic analysis (Figure 3).

and Supplementary Figure 1). To elucidate whether less conserved sequences of motifs 7 and 8 affect the 3D structure, we used a homologous-based model from the expasy website (see Text Footnote 6). The results showed that the 37 LiCIPKs could be divided into eight classes according to their three-dimensional structure (Figure 2 and Supplementary Figure 2). Classes A to E are compact, with the N- and C-ends adjacent to each other (Figure 2A and Supplementary Figures 2A–E). The major differences between classes A to E are the α-helix numbers and arrangements ahead of NAF/FISL domain compared to that of LiCIPK24/SOS2 (Figure 2A and Supplementary Figures 2A–E). The 3D structure of class F is looser, with the N-end far away from the C-end (Supplementary Figure 2F and Figure 2B), similar to the active structure model of AtCIPK24 (Chaves-Sanjuan et al., 2014). The 3D structure of classes G and H only contain the N-terminal sequence, however, the 3D structure of class H is the homodimer of N-terminal sequences (Figures 2C,D and Supplementary Figures 2G,H). Of the CIPKs lacking motif 7 (e.g., LiCIPK1, -5, -19, -27, and -37), although the sequences are less conserved (Figure 1), they fold

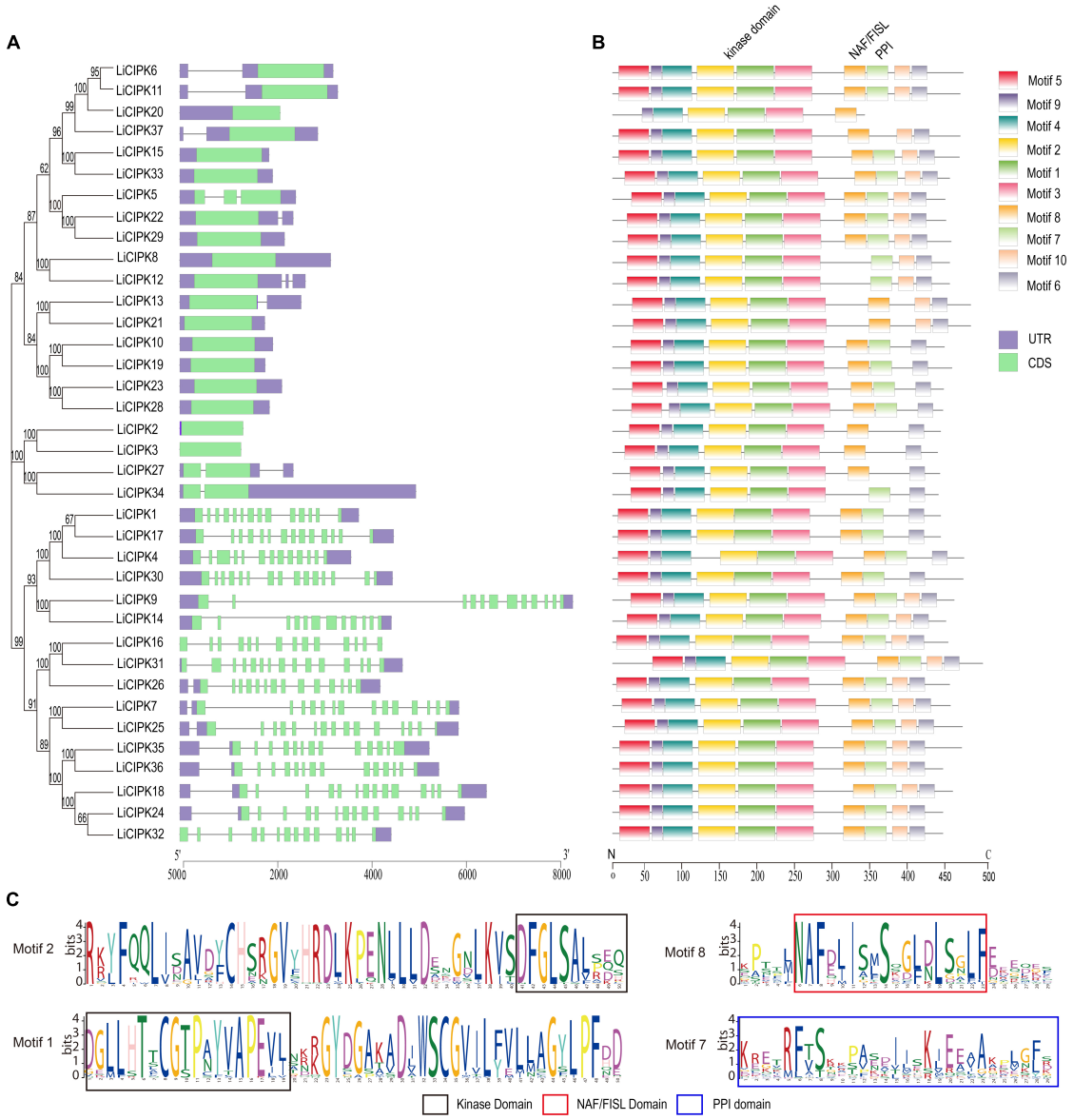


FIGURE 1
Phylogenetic tree, conserved motifs, and gene structure of *LiCIPKs*. **(A)** Exon–intron patterns. **(B)** Conserved motifs arrangement of LiCIPKs. **(C)** The functional domains. Motifs and exons are indicated as rectangles, other sequences as lines.

into a PPI α-helix (Supplementary Figures 2A,C,E,F). Only LiCIPK10 lacked the C-terminal structure (except LiCIPK21 with C-terminal sequence deletion; Figure 1). In summary, we found that the 3D structure analysis provided additional details on the active domain compared to the conserved sequences analysis.

Phylogenic analysis of LiCIPKs

To reveal the phylogenic relationship of the LiCIPKs, an NJ phylogenetic tree was constructed using the full-length amino acid sequences of CIPKs from *L. indica* and three other species (Arabidopsis, rice, and grape). A total of 117 CIPKs were divided into five groups (A–E) (Figure 3 and Supplementary Table 2). Of these, group A was the largest, containing 44 members and group D was the smallest with only 9 members. We found that groups A and C could be further divided into several subgroups (Figure 2 and Supplementary Table 1). Generally, the evolutionary relationship between *LiCIPKs* and *VvCIPKs* is closer than that between *LiCIPKs* and *AtCIPK* or *OsCIPKs*. Hence, the evolutionary rate of the LiCIPK gene family is faster than predicted.

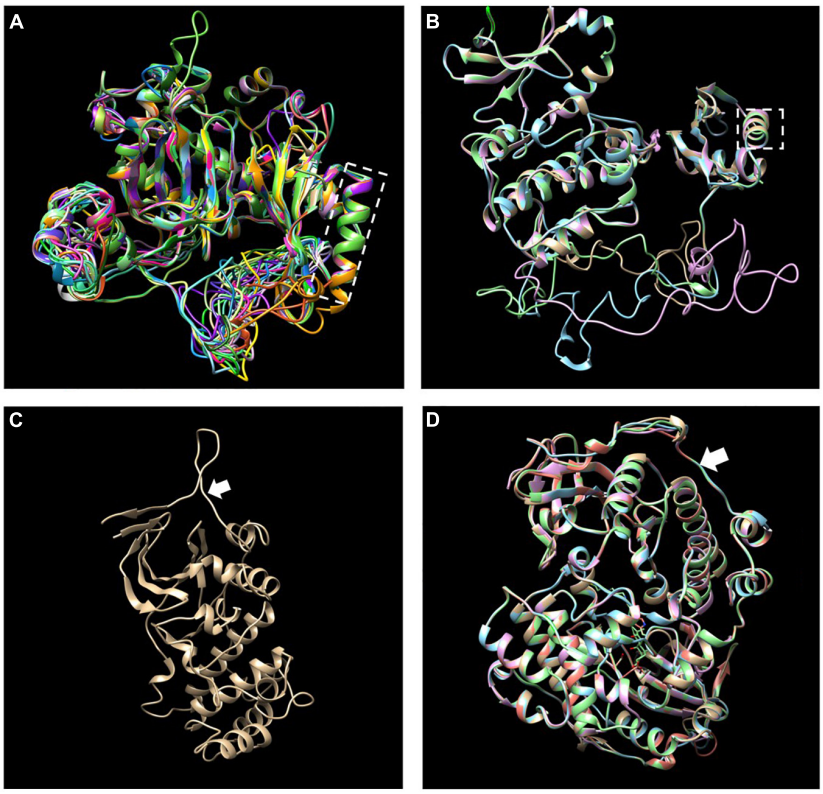


FIGURE 2
Merged three-dimensional (3D) structure of 37 LiCIPKs. **(A)** Merged 3D structure of LiCIPKs belong to class A to E. **(B)** Merged structure of LiCIPKs belong class F. **(C)** 3D structure of class G. **(D)** Merged structure of LiCIPKs class H. The classes A to H are depicted on **Supplementary Figure 2**. The dashed box indicates the α-helix of PPI domain; arrow indicates the junction between the N and C-terminal.

Due to an apparent increase in *LiCIPK* gene numbers, we further analyzed the inter-species collinear relationship of *CIPK* loci (*L. indica* vs. Arabidopsis, *L. indica* vs. grape). Our results showed that *CIPKs* loci were maintained differently during the evolution of the three species (**Figure 4**). As **Figure 4** illustrates, only 10 *VvCIPK* and 15 *AtCIPK* collinear loci were identified in *L. indica* (**Figures 4A,B**). Meanwhile, various *CIPKs* were lost during evolution, including grape *VvCIPK33/30* and Arabidopsis *AtCIPK9* (*At1g01140*), which lack an orthologous gene in *L. indica* (**Figure 4A**). In contrast, *AtCIPK* and *VvCIPK* usually have more than one orthologous loci in *L. indica*. These results indicate that although some of the ancient CIPK have been lost, the remaining *LiCIPKs* have been duplicated during evolution (**Figures 4A,B**). As a result, the total number of *LiCIPK* gene family members is higher than that of grape and Arabidopsis.

To further elucidate the evolution of *LiCIPKs*, we surveyed cognate loci intraspecies, and discovered that 29 (78.38%) of the 37 *LiCIPKs* formed 21 duplicated pairs, which could be divided into four types according to their relationships (**Figures 4C,D** and **Supplementary Figure 3**). The type A duplicated genes contained two members, which was also observed in other plants such as grapes (Yin et al., 2017). Type B contained three members, however, these were not mutually duplicated members. Type C and D involved four (or more) members with complicated relationships. Types B–D have not been previously identified in other plants (**Figure 4D**). Additionally, the three loci harbored *LiCIPKs* very closely on chromosomes (*LiCIPK12* and *LiCIPK13* on chromosome 5; *LiCIPK20* and *LiCIPK21* on chromosome 9; *LiCIPK29* and *LiCIPK30* on chromosome 15), that were not tandem repeat loci (**Supplementary Figure 3**). This result coincides with the phylogenetic analysis, such as that LiCIPK30 has high similarity to *LiCIPK16*, but not to the adjacent *LiCIPK29* (**Figure 3** and **Supplementary Table 2**).

The high percentage of duplicated *LiCIPKs* in *L. indica* prompted us to investigate the time of the duplication events. To this end, we calculated the *Ks* of *LiCIPK* duplicated pairs, and orthologous pairs between *LiCIPKs* and *AtCIPK*, *LiCIPKs* and *VvCIPKs* gene pairs. The results showed that *LiCIPK* paralogs have two apparent Ks peaks (**Figure 4D** and **Supplementary Figure 4**), indicating that the existing *LiCIPKs* experienced two duplication events (**Table 2**). According to the Myrtales specific molecular clock (1.14×10^{-8}) reported previously

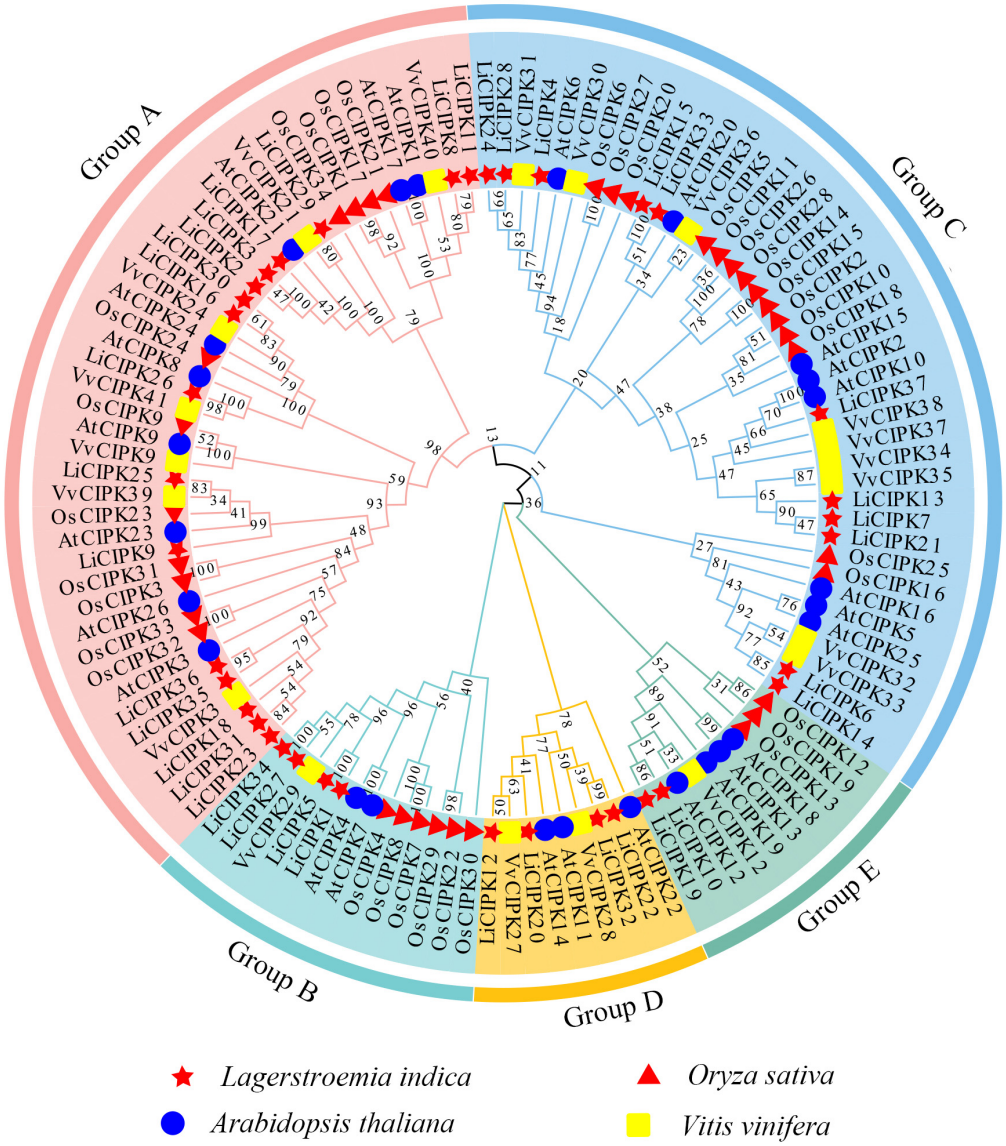


FIGURE 3
NJ-Phylogenic tree of AtCIPKs, OsCIPKs, VvCIPKs, and LiCIPKs. The group A to E is shaded by Jade red, Mint cyan, wathet blue, orange, and aquatic green, respectively. Genes of *L. indica* (*Li*), *Arabidopsis thaliana* (*At*), *Oryza sativa* (*Os*), and *Vitis vinifera* (*Vv*) are marked by red star, blue circle, red triangle, and yellow rectangle, respectively.

(Feng et al., 2021), the two duplication events of *L. indica* were estimated to have occurred around 16.24–26.34 MYA and 56.12–61.16 MYA, respectively (**Table 2**). Interestingly, we found that type A and B duplicated genes were maintained by the recent duplication events, whereas type C and D duplicated genes experienced two duplicated events.

The K_s value between *LiCIPKs–VvCIPKs* was 1.5–2, while that between *LiCIPKs–AtCIPK* was more than 2, implying that the *LiCIPKs* are more highly divergent from *AtCIPKs* than *VvCIPKs* (**Supplementary Figure 4**). This result is also consistent with the phylogenetic analysis (**Figure 3**). The average K_a/K_s ratios of *LiCIPKs–VvCIPKs, LiCIPKs–VvCIPKs,* and *LiCIPKs* pairs are 0.103, 0.106, and 0.142, respectively. Hence, the *CIPKs* genes among the three species were under strong purifying selection (**Table 2**).

Cis-elements in the promoter of *LiCIPKs*

To clarify the regulatory mechanism of *LiCIPK* genes under abiotic stress, the *cis*-elements of the LiCIPKs promoters (−2,000 bp upstream ATG), which respond to plant hormones and abiotic stress were analyzed using PlantCARE software.

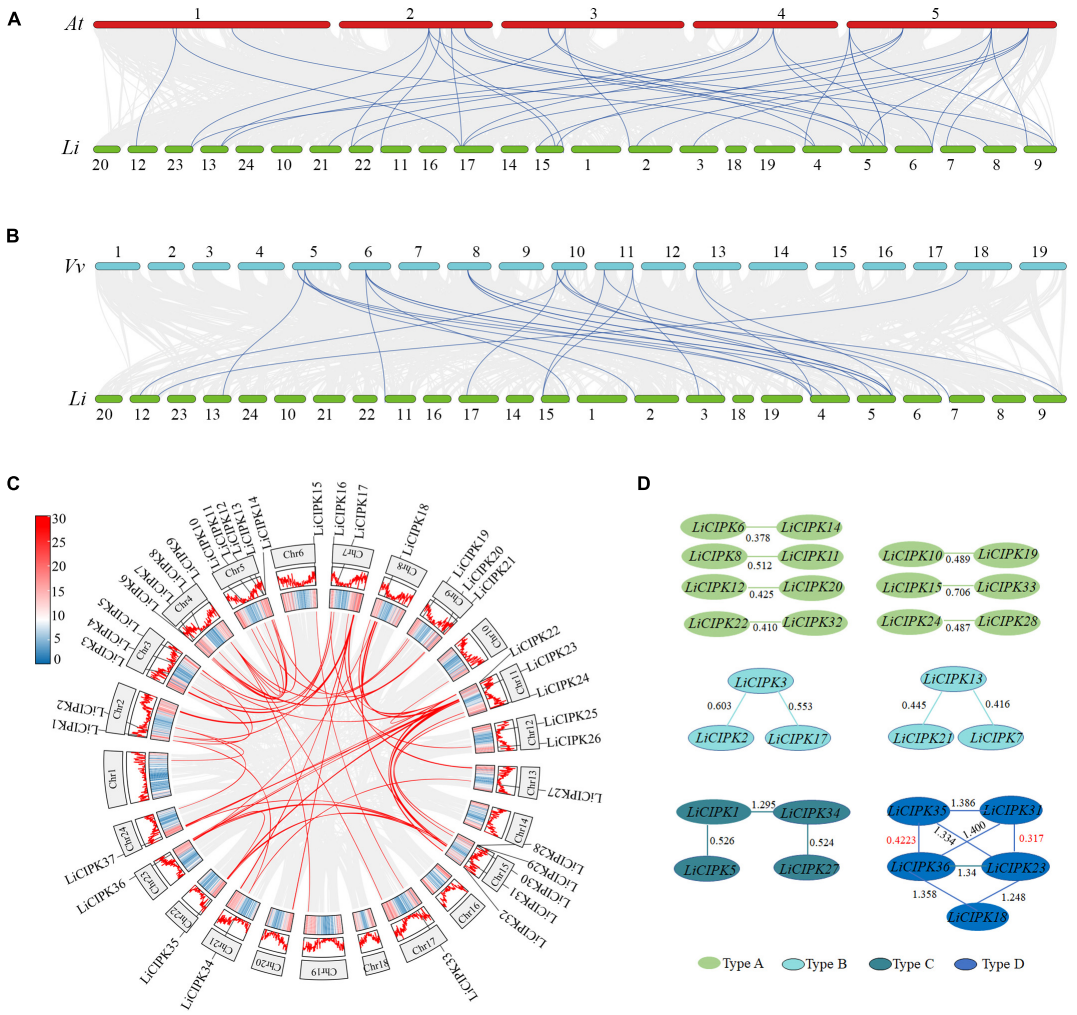


FIGURE 4
Inter- and intraspecies collinearity of *CIPK* loci. **(A)** *AtCIPKs* vs. *LiCIPKs*. **(B)** *VvCIPKs* vs. *LiCIPKs*. The light gray lines indicate the synteny between species, blue lines indicate cognate loci of *CIPKs*. **(C)** *LiCIPKs* vs. *LiCIPKs*. Red lines show synteny loci of *LiCIPKs*. Gene density across chromosomes are indicated by hot map (inner circle) and column map (medium circles), out circle show length of chromosomes. **(D)** Classification of duplicated *LiCIPK*. Light green, cyan, dark green, and blue indicate duplicated *LiCIPKs* containing two, three, four, and five members, respectively.

TABLE 2 Divergence time of CIPKs among three species.

Species – species		*Li – Li*	*Vv – Li*	*At – Li*
K_a		0.0887 ± 0.0492	0.1857 ± 0.0722	0.2585 ± 0.0722
K_s	Mean 1	0.485 ± 0.115	2.013 ± 0.735	2.664 ± 1.151
	Mean 2	1.337 ± 0.0574	–	–
K_a/K_s (mean)		0.142 ± 0.072	0.103 ± 0.045	0.106 ± 0.071
Divergence time (MYA)	$\lambda = 1.5 \times 10^{-8}$ (mean 1)	16.18 ± 3.84	67.11 ± 24.62	88.99 ± 38.38
	$\lambda = 1.14 \times 10^{-8}$ (mean 1)	21.29 ± 5.05	88.23 ± 32.24	116.84 ± 50.48
	$\lambda = 6.1 \times 10^{-9}$ (mean 1)	39.79 ± 9.45	165.03 ± 62.31	218.33 ± 94.38
	$\lambda = 1.5 \times 10^{-8}$ (mean 2)	44.57 ± 1.91	–	–
	$\lambda = 1.14 \times 10^{-8}$ (mean 2)	58.64 ± 2.52	–	–
	$\lambda = 6.1 \times 10^{-9}$ (mean 2)	109.60 ± 4.70	–	–

All data are mean ± SD. K_a, non-synonymous substitutions per synonymous; K_s, synonymous substitutions per synonymous. MYA, million years ago.

The plant hormone and abiotic stress responsive elements were broadly distributed in *LiCIPK* promoters. Of these, the top three elements are ABRE (abscisic acid responsive element), MeJA (MeJA responsive element), and anaerobic inducible element (Figure 5). The overlapping of different elements on promoters is a common phenomenon. For example, defense, ABRE, GA, and auxin responsive elements were arranged in an array on the promoter of *LiCIPK13*, *-15*, *-34*, etc. From the *cis*-elements in the *LiCIPK* promoters, we deduced that *LiCIPKs* may be widely involved in plant hormone signaling and stress response. Furthermore, the number and location of *cis*-elements differed in the promoter of CIPK duplicated pairs, for example, *LiCIPK24* and 28 pairs (Figures 4, 5). These results imply functional differentiation of the duplicated genes (Figure 5).

Different *LiCIPKs* respond to abiotic stress differently

To investigate the regulatory mechanism and potential function of LiCIPKs, we analyzed their expression profiles under

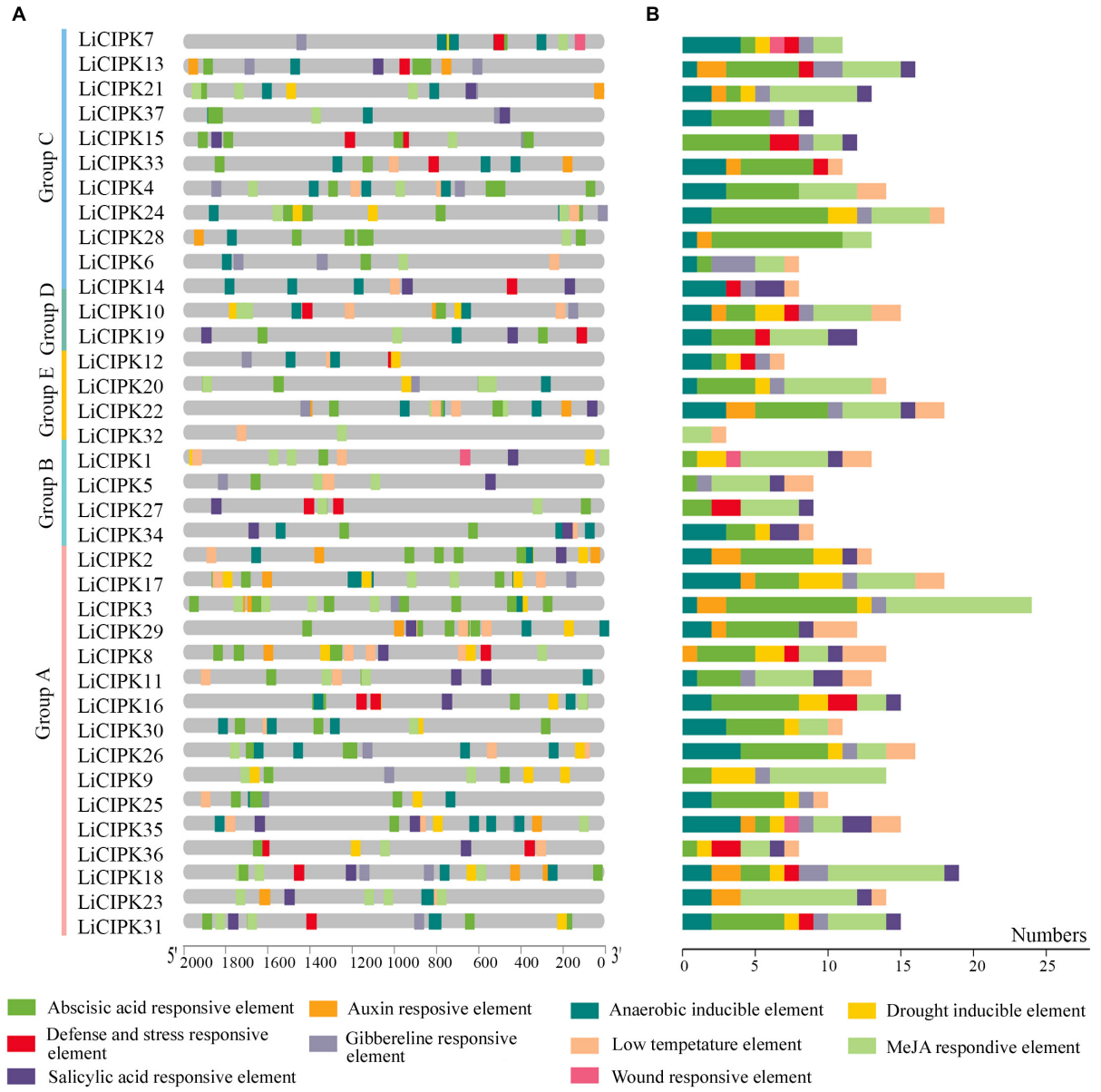


FIGURE 5
Distribution and numbers of abiotic stress and hormone responsive *cis*-elements in promoters of *LiCIPKs*. **(A)** The distribution of different *cis*-elements. **(B)** Numbers of different *cis*-elements. Elements are indicated as rectangles, others sequences as lines.

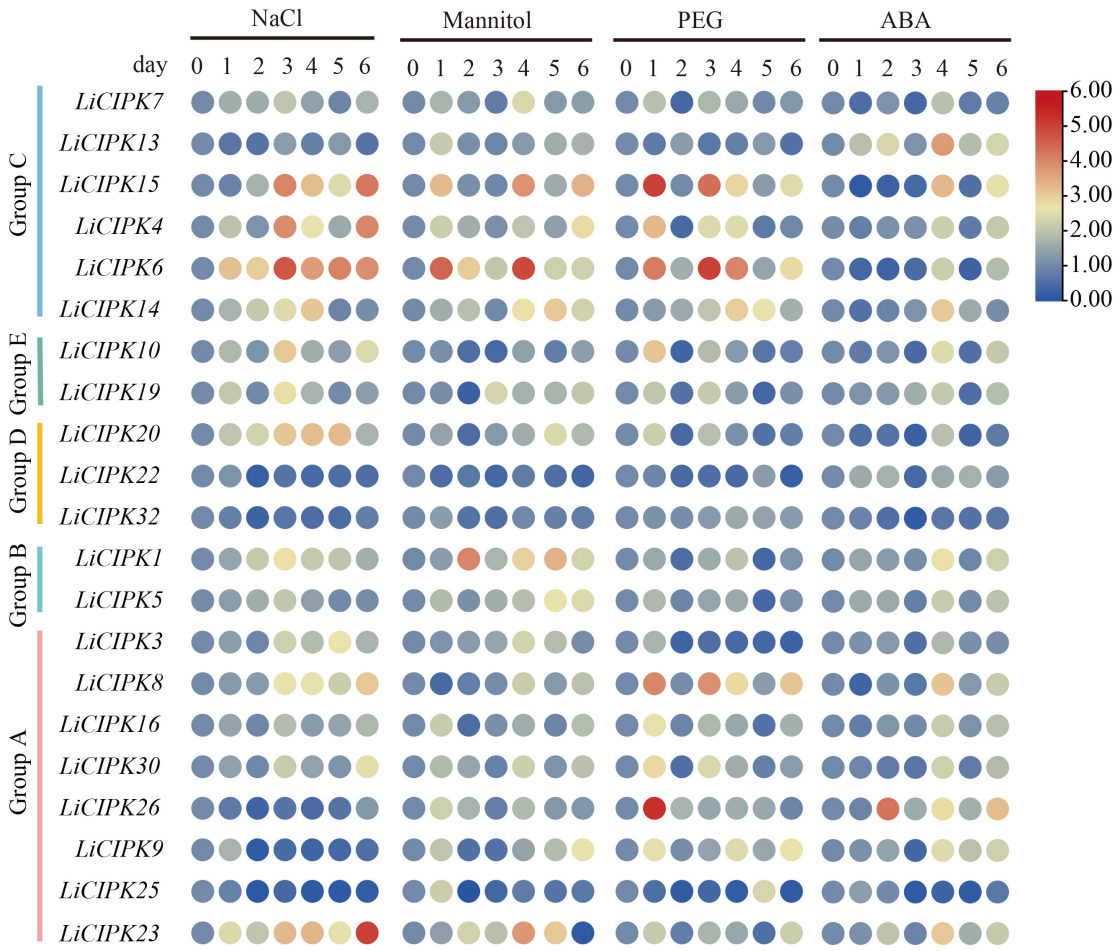


FIGURE 6
Expression profiles of *LiCIPKs* under four kinds of abiotic stress. Cuttings were treated by NaCl solution (200 mmol·L^{-1}), mannitol solutions (200 mmol·L^{-1}), 15% PEG8000 and ABA (10 × 10^{-4} mol·L^{-1}), respectively. Roots were collected after treatment of 0, 1, 2, 3, 4, 5, and 6 days. The mean fold changes of the expression level are depicted by heat map. $n \geq 3$.

NaCl, PEG ABA, and mannitol induced salt, drought, and osmotic stress using qRT-PCR. Of the 30 gene-specific primer pairs (Supplementary Table 1), 21 *LiCIPKs* were successfully amplified (Figure 6). This result indicates that under different abiotic stresses, expression levels of the responding members differ. Under salt stress, *LiCIPK4*, *−6*, and *−15* of group C; *LiCIPK10* and *−19* of group E; and *LiCIPK1* of group B reached their highest expression after 3 days of treatment, whereas *LiCIPK23* and *−30* peaked after 6 days of treatment. Under mannitol (osmotic) stress, *LiCIPK1* of group B and *LiCIPK4*, *−6*, and *−15* of group C exhibited fluctuating patterns during the 6 days of treatment. *LiCIPK15*, *−4*, *−6*, *−8*, and *−26* responded to PEG treatment more rapidly than to the other types of stress treatment. Most LiCIPKs responded to ABA until 4 days of treatment. Moreover, the expression patterns of certain members showed opposite tendencies, such as LiCIPK3 under salt and PEG treatment (Pearson $r = -0.3659$), and *LiCIPK1* under mannitol and PEG treatment (Pearson $r = -0.6001$). The members *LiCIPK4*, *−6*, *−15*, *−14*, *−10*, *−1*, *−5*, *−8*, *−16*, *−30*, −9, and −23 responded to all four stress treatments, indicating that they may be involved in stress signaling interplay. However, other members, including *LiCIPK22*, *−25*, and −32 exhibited relatively no changes under the four treatments (fold changes < 2) and thus, members did not likely participate in the stress response under our experimental conditions.

LiCIPK30 complements *AtSOS2* in Arabidopsis

Phylogenetic analysis showed that LiCIPK30 is an orthologous gene of AtCIPK24 (AtSOS2) (Figure 3) and responds to the four abiotic stresses (Figure 6). However, whether LiCIPK30 is a bona fide SOS2 gene requires verification. To clarify the function of LiCIPK30 *in vivo*,

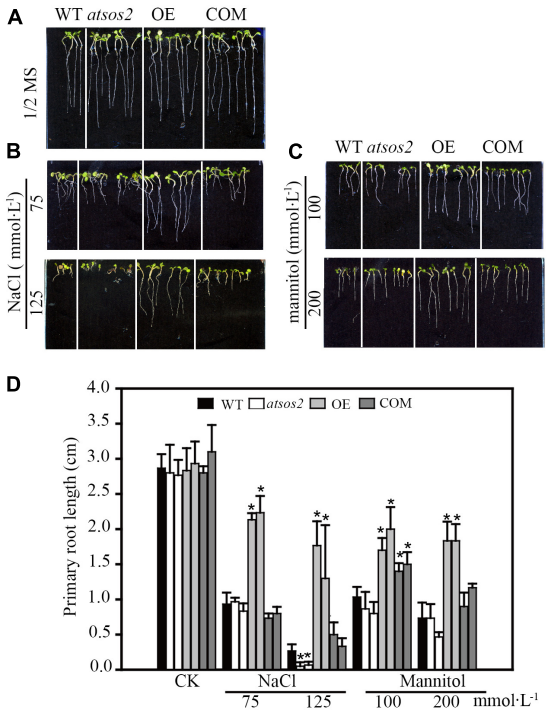


FIGURE 7
Phenotype of WT, *atsos2*, OE, and COM under salt and mannitol stress. **(A)** 1/2 MS. **(B)** 1/2 MS with 75, 125 mmol L^{-1} NaCl. **(C)** 1/2 MS with 100 and 200 mmol L^{-1} mannitol. **(D)** Primary root length. Data were collected from seeds grown on medium 10 days after planting. Mean ± SD, $n \geq 3$, $^{*}p < 0.05$ (Student's *t*-test).

we developed a 35S:LiCIPK30 construct, and transformed it into the *atsos2* mutant and Arabidopsis WT (Col-0). After genotype identification and expression analysis of the transformed lines, the complementary lines (COM) and over-expression lines (OE) were successfully obtained (**Supplementary Figures 5, 6**). The seeds of the four lines (WT, *atsos2*, OE, and COM) were germinated under salt (0, 75, and 125 mmol L^{-1} NaCl) and osmotic (0, 100, and 200 mmol L^{-1} mannitol) stress conditions. After 10 days, the primary root length of all four lines showed no difference under normal conditions (1/2 MS; **Figure 7A**). Under salt stress, OE lines exhibited the highest growth rate, while atsos2 had the lowest, and that of COM and WT lines were between OE and the mutant, however, were all inhibited (**Figure 7**). Under osmotic stress, WT, *atsos2*, OE, and COM lines have similar phenotype as that of salt stress (**Figure 7**). Collectively, LiCIPK30 could salvage the salt- and osmotic-sensitive phenotype of *atsos2*. In fact, over-expression of LiCIPK30 in WT enhanced salt/osmotic tolerance of Arabidopsis during the germination and seedling stages. By combining the results of phylogenic analysis, we refer to LiCIPK30 as LiSOS2 hereafter.

The physiological mechanism of LiSOS2 confers stress tolerance of Arabidopsis

To further uncover the function of LiSOS2, we observed the phenotype of OE lines during the rosette leaf stage under different salt stress conditions (0, 100, and 200 mmol L^{-1} NaCl). All OE lines maintained growth well under stress conditions, while WT growth appeared to be inhibited with lower relative water content and smaller leaf area after 7 days of salt treatment (**Figures 8A–C**). The total chlorophyll content (chlorophyll a and b) appeared to decrease in WT and decreased weakly in the OE lines under 200 mmol L^{-1} NaCl treatment (no significant statistical difference) (**Figures 8D–F**). Moreover, the activity of reactive oxygen species (ROS) scavenging enzymes, POD, and CAT increased in the OE lines and decreased in WT (**Figures 8G,H**). MDA content increased under the harsher salt treatment (increased concentration and prolonged time) in WT and decreased in OE lines (**Figure 8I**). Collectively, these results indicate that the overexpression of LiSOS2 in Arabidopsis confers salt stress tolerance through developmental adaptation (regulating leaf size), decreased damage to the leaf photosynthetic system, membrane lipid peroxidation, and enhanced ROS scavenging ability (physiological adaptation).

LiSOS2 enhances the expression of *AtABF3* and *AtRD22 in vivo* under salt stress

After the plant perceives a salt stress signal, it responds through an interplay of several pathways to decrease the detrimental effect. Hence, we detected expression profiles of the endogenous genes, including those of the SOS pathway (*SOS1–SOS3*), mitogen-activated protein kinase (which functions through ABA pathway), ABA-dependent signaling pathway (*ABF3*, *ABI5*, *RD22*, *RD29A*, and *RD29B*), ABA-independent signaling pathway (*DREB2A*, *RD20*, and *RD29A*), ROS signal (respiratory burst oxidase homolog, RBOH), and ion homeostasis (Na^{+}. K^{+}), as well as small molecular proline biosynthesis-related gene and membrane signal-related gene *GOlS2* (galactinol synthase 2). These genes, excluding *AtHKT1*, were induced under salt stress in both WT and OE lines. It was also evident that the fold changes of these induced genes differed between WT and OE lines. Among the 18 upregulated genes, only *ABF3, AtRD22,* and *GOlS2* expressions in OE were higher than that of WT under higher salt stress (**Figure 9**). According to the induced expression patterns in WT, the upregulated genes could be divided into three classes: Class I, genes that were continuously induced as salt stress was enhanced, namely, *SOS2*,

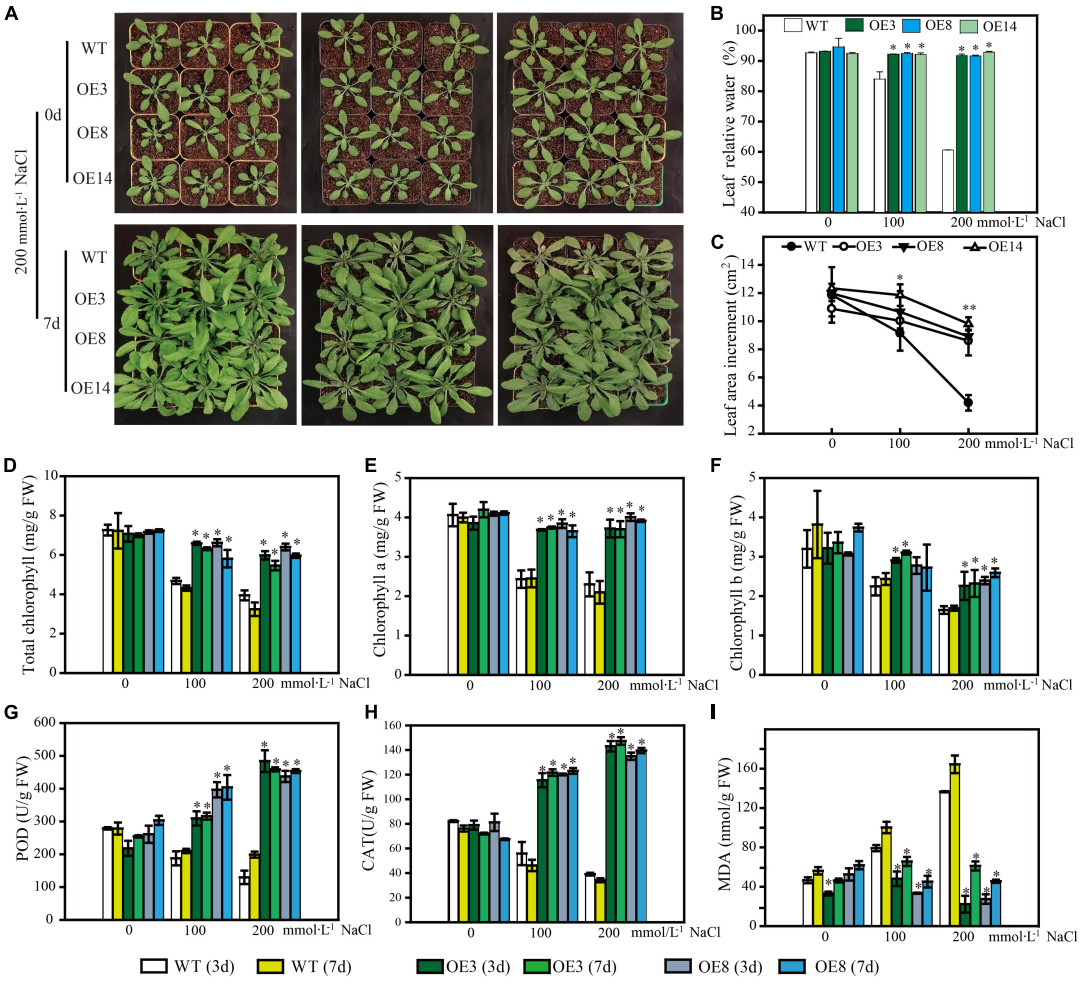


FIGURE 8
Phenotype of LiSOS2 OE lines. **(A)** Phenotype of plants before and after salt stress. **(B)** Relative water contents of leaves. **(C)** Leaves area increment per plant after salt treatment. **(D)** Total content of chlorophyll. **(E)** Chlorophyll a. **(F)** Chlorophyll b. **(G)** Activity of POD. **(H)** Activity of CAT. **(I)** Content of MAD. Mean ± SD, $n \geq 3$, $^{*}p < 0.05$, $^{**}p < 0.01$ (Student's *t*-test, compared to the WT at same conditions).

MPK4, *MPK6*, *ABI5*, *RD29B*, *RBOHD*, *P5CS1*, and *NHX1*. Class II, genes induced under lower salt stress (100 mmol L^{-1} NaCl) but downregulated under 200 mmol L^{-1} NaCl stress, namely, *ABF3*, *DREB2A*, *RD20*, and *RD22*. Class III, invariable genes, namely, *SOS3*, *MYB2*, *RD29A*, *RBOHF*, and *GOlS2* (**Figure 9**). Additionally, the expression of most genes in the OE lines remained relatively invariable under the two stress conditions, indicating that these genes may be regulated under invariable signals. Based on these results, we concluded that under the "protection" of LiSOS2, OE plants did not respond as strongly to harsh stress conditions as WT.

Discussion

In this study, we surveyed the *CIPK* gene family of the ornamental plant *L. indica* through mining recently sequenced reference genome data (manuscript under preparation). Our results indicated that there are 37 *LiCIPKs* in *L. indica* that can be divided into two classes according to their intron/exon patterns, or five groups according to the phylogenic relationship of Arabidopsis, grape, rice, and *L. indica CIPKs* (**Figures 1**, **3**). The intron-rich (43.25%) and intron-less (56.75%) patterns of *LiCIPK* genes were similar to that of *AtCIPKs*, *VvCIPKs,* and *OsCIPKs* (Kolukisaoglu et al., 2004; Kanwar et al., 2014; Xi et al., 2017). Besides these four species, the exon/intron structure of *CIPKs* of *Populus* (Yu et al., 2007), maize (Chen et al., 2011), canola (Zhang et al., 2014), and pepper (*Capsicum annuum* L.) (Ma et al., 2019) were highly similar. These results suggest that the diversity of *CIPK* gene structure predated the split of angiosperms. A previous investigation indicated that intron-less CIPKs first appeared in the basal angiosperm *Amborella trichopoda* and were derived from retrotransposition events that occurred in the ancestor of angiosperm plants (Kleist et al., 2014;

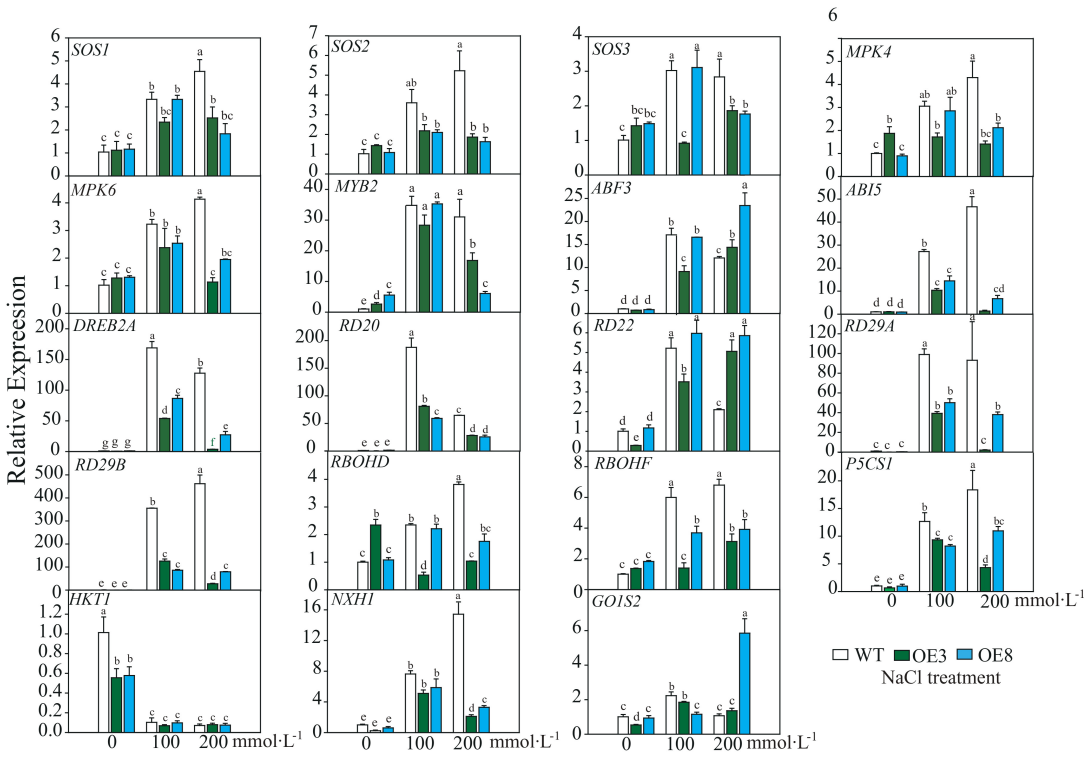


FIGURE 9
Relative expression of salt stress responsive genes of WT and OE lines. *SOS*, Salt Overly Sensitive; MPK, mitogen-activated protein kinas; MYB, V-MYB avian myeloblastosis viral oncogene homolog; ABF, ABA response element binding factor; ABI, abscisic acid insensitive, DREB, dehydration responsive element binding proteins 2; RD, responsive to desiccation; RBOH, respiratory burst oxidase homolog; P5CS, $^{\Delta}$1-pyrroline-5-carboxylate synthetase; HKT, high-affinity potassium transporter; NHX, Na^+/H^+ antiporter; GolS, galactinol synthase. Mean ± SD, $n = 3$, lowercase letters indicate significant difference between samples ($p \leq 0.05$, two ways ANOVA).

Zhang et al., 2020). Hence, the gene structure of *LiCIPKs* was inherited from their ancestor species.

To date the reported CIPK gene family members in plant genomes have been under pure selection (Zhang et al., 2014; Sun et al., 2015; Li et al., 2016; Xi et al., 2017; Yin et al., 2017; Liu et al., 2019), indicating relatively conserved amino acid sequences among plant *CIPKs*. In this study, the 37 *LiCIPKs* could be divided into several groups according to their phylogenetic relationships, conserved motifs, and 3D structures. Regarding kinase function, we considered that information from 3D-structures may be more credible than that from MEME motifs. From the regulatory mode of AtCIPK (Chaves-Sanjuan et al., 2014), we deduced that class A–E LiCIPKs depend on CBL to switch on and PP2C (or other types of phosphatase) to switch off kinase activity, due to their compact 3D-structure (**Figure 2** and **Supplementary Figure 2**). The class F LiCIPKs may have higher basal kinase activity due to their open structure, however, the full activity requires binding to a specific CBL (**Figure 2** and **Supplementary Figure 2**). Few LiCIPKs lacked the C-lobe, which was not anticipated as their NAF/FISL and PPI were intact (**Figures 1**, **2** and **Supplementary Figures 1, 2**). The reason for this may be the divergent sequences at the C-terminal. However, the N-terminal 3D-structures of LiCIPKs highly resemble those of AtCIPK24 (Chaves-Sanjuan et al., 2014). We speculate that the kinase activity of these types of LiCIPKs still depend on CBL, however, is independent of phosphatase. In addition, the serine insertion in the NAF/FISL motif of LiCIPK6 and -14 increases the hydrophilicity of these regions (**Supplementary Figure 1**), which may decrease binding activity between CBL and CIPKs according to the interactions of AtCIPK24/SOS2 and AtCBL4/SOS3 (Sánchez-Barrena et al., 2007). Hence, future work should identify the CBL partner, kinase activity and downstream targets.

The most striking evolutionary characteristic of *LiCIPKs* is the high percentage (∼80%) of duplicated *LiCIPKs* in the *L. indica* genome (**Figures 4C,D** and **Supplementary Figure 3**). This indicates that they are the vestiges of WGD events, not of chromosome segmental duplication events. From the two separate K_s peak distributions (**Supplementary Figure 4**), combined with the results of previous investigations (Myburg et al., 2014; Qin et al., 2017; Luo et al., 2020; Feng et al., 2021), we inferred that the older WGD1 event (56.12–61.16 MYA) occurred commonly in the *Lythraceae* ancestor, while the recent WGD2 event (16.24–26.34 MYA)

may have occurred exclusively in *Lagerstroemia* species (**Table 2**). Our results are supported by the phylogenomic analysis of 20 *Lagerstroemia* chloroplast genomes, which shows that *Lagerstroemia* originated in the late Paleocene (∼60 MYA), and *Lagerstroemia* species were thriving around 11.8–31.6 MYA (Dong et al., 2021). However, currently, only one genome dataset exists for *Lagerstroemia*, whether this WGD event occurred commonly in *Lagerstroemia* species still requires verification.

It is well known that dicotyledon experienced three common WGD events: ζ, ε, and γ (Tang et al., 2008; Clark and Donoghue, 2018; One Thousand Plant Transcriptomes, 2019; Li and Barker, 2020). As for *L. indica,* it experienced five WGD events contained two WGD events we uncovered in this research, hence, it is interesting how the *LiCIPKs* were retained after five WGD events. Based on the phylogenic relationships of CIPKs among Arabidopsis, gape, rice and *L. indica*, duplicated pairs of *LiCIPKs* (**Figures 3**, **4**, **Table 2**, and **Supplementary Table 2**), and the previous report (Zhang et al., 2020), we deduced the evolutionary history of *LiCIPKs* (**Supplementary Figure 7**). Retaining of LiCIPKs after two linkages specific WGDs were summarized in **Supplementary Table 3**. After two WGD events and genome reshuffling, different genes numbers were kept. We deduced that the types A–C may have originated from one ancestral gene, but type D cluster has two segmental duplicated ancestral genes. In type A, only two WGD2 duplicated members were retained; in type B, one of the WGD1 duplicated members was lost; In type C, all four genes were retained; in type D, three WGD1 members and two WGD2 members were retained, however, collinearity of five loci was more intact than that of loci of types A–C, genes involved have multilateral relationships. Apart from the duplication pairs, the eight *LiCIPKs* (∼20%) lacking duplicated pairs were likely the members retained in the genome by ancient WGDs, or their duplicated paralogous were lost (**Supplementary Figure 7**). Interestingly, their orthologous pairs in Arabidopsis experienced similar evolutionary mechanisms, for example, *AtCIPK6, AtCIPK8*, *AtCIPK23*, and *AtCIPK24*/*SOS2* were not expanded (or their paralogs were lost) after linkage specific α, and β WGD events and they are the retained members following ancient WGD event (Zhang et al., 2020). Gene balance hypothesis have been widely accepted to explain genes retaining after WGD (Birchler and Veitia, 2021), future work should be undertook to identify the specific interaction LiCBL-LiCIPK and their function in *L. indica*.

In this study, we surveyed the *cis*-elements of *LiCIPK* promoters and detected their expression through qRT-PCR (**Figures 5**, **6**). We found that expression levels and the numbers of several *cis*-elements were not positively correlated, which may have been due to gene expression depending on the co-operation between *cis*-elements and *trans*-factors (such as TFs, RNA Pol, etc.) and the post-transcription regulation mechanism. As kinases, CIPKs regulate activities of several downstream proteins, for instance, MdCIPK22 regulates MdAREB2 in apples (Ma et al., 2017). However, since the transcriptional monitor of CIPKs received little attention, future studies will aim to identify the factors affecting *LiCIPK* expression under abiotic stress, which will facilitate a comprehensive elucidation of the complete regulatory mechanism of CIPKs. Additionally, the variety we used in this study differs from used for genome sequencing, hence, sequence polymorphisms of the promoter may also account for the difference between the predicted and observed results.

Comparing the expression of *LiCIPK30*/*SOS2* in *L. indica* (**Figure 6**) and *AtCIPK24*/*SOS2* in Arabidopsis (**Figure 9**), we found that *LiCIPK30* is not induced like *AtCIPK24*/*SOS2* under salt stress (**Figures 6**, **9**). This may be a key difference between the two species. The *L. indica* variety (Black Diamond 'Blush V2') we used in this study is salt tolerant, with the ability to grow new buds and roots under 75 mmol·L^{-1} NaCl stress (unpublished lab data). Hence, under the conditions of the current study, we deduced that it may not have sensed the high stress, causing the dissimilarity in the expression patterns of *LiCIPK30*/*SOS2* to that of Arabidopsis (**Figure 9**). Based on the phenotype of OE Arabidopsis, we inferred that upregulation of *LiCIPK30*/*SOS2* expression protects plants from the detrimental impact of salt stress through developmental adaptation and physiological adaptation (**Figures 7**, **8**). The function of *LiCIPK30* in Arabidopsis may also partially occur through the ABA pathway as both *AtRD22* and *AtABF3* are ABA-responsive genes (**Figure 9**) (Yoshida et al., 2010; Liu et al., 2018). *ABF3* overexpression confers tolerance to multiple abiotic stresses in alfalfa (Wang et al., 2016) and drought tolerance in Arabidopsis and rice (Oh et al., 2005; Yoshida et al., 2010). However, currently the mechanism by which *LiCIPK30*/*SOS2* regulates the transcription of *ABF3* is unclear. Nevertheless, yeast-two-hybrid and BiFC assays have revealed that VaCIPK02 of amur grape modulates ABA signaling through interacting with the ABA receptor-PYL9 (Xu et al., 2020). Hence, it is possible that *LiCIPK30*/*SOS2* activates the components upstream of *ABF3*. Elevated expression of *GOlS2* may lead to an increase in raffinose accumulation, which serves as an osmotic compound *in vivo*. The GOlSs have been reported to confer abiotic stress, particular to drought and cold (Li et al., 2020c; Liu Y. et al., 2020; Dai et al., 2022). Collectively, the enhanced Na^+ exclusion, ABA pathway signaling, and ROS scavenging, as well as small osmotic compounds, coordinate to improve the performance of OE plants under salt and osmotic stress.

Conclusion

Our data reveal the characteristics and evolutionary history of *LiCIPKs*, as well as the gene resources involved in abiotic stress. Ectopic expression of *LiCIPK30* in Arabidopsis enhances salt stress tolerance. This work also advances the

current understanding regarding the complex interaction between *L. indica* and its harsh environmental conditions. Further studies are required for an in-depth elucidation of these interactions.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary material**.

Author contributions

CY and JZ contributed to conception and design of the study. YK, YH, YZ, YL, HW, GL, BL, YC, and FZ performed the experiments. CY and YK prepared the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by fund of the Forestry Science and Technology of Jiangsu Province (Su[2021]TG03), fund of Science and Technology of Nantong (MS12020070), grants of key projects of Jiangsu R&D plan (Modern Agriculture) (BE2018326), and the Undergraduate Innovation Training Program of Jiangsu Province (202010304091Y).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.942217/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Multi-alignment of LiCIPKs N and C conserved motifs. The top is the logo of the conserved amino acids. The red boxes show the amino acid residues may affect the interaction between CBL sensor.

SUPPLEMENTARY FIGURE 2
The three-dimensional (3D) structure of 37 LiCIPKs. **(A–E)** Member with 3D Structure reassemble AtCIPK24/SOS2. The "a", "b," and "c" indicates three α-helixes ahead the NAF domain, respectively. **(F)** Members have a looser structure. **(G,H)** Members only with N-terminal 3D structure. Member of class G is monomer, of class H is homo-dimer.

SUPPLEMENTARY FIGURE 3
Chromosome location of *LiCIPKs*. Chromosome length show on left. Blue lines show segmental duplication of LiCIPKs.

SUPPLEMENTARY FIGURE 4
K_s of synteny CIPKs pairs of three species. *Li, L. indica; At, Arabidopsis thaliana*; *Vv*, Vitis vinifera. ND, genes pairs are too divergence to detect.

SUPPLEMENTARY FIGURE 5
Identification of *LiCIPK30* OE Arabidopsis. **(A)** Genotype of different single plant. Lane M: DL2,000. Lanes 1–3, WT, negative control. Lanes 4–22 different individual plant with hygromycin resistance. *LiCIPK30* specific primers were designed for PCR detection the genotype of the individual plant. PCR products were separated by 1% agarose gel. Positive OE plants were recoded (white number). **(B)** qRT-PCR detected the expression of *LiCIPK30*.

SUPPLEMENTARY FIGURE 6
Identification of atsos2 mutant *and LiCIPK30* COM lines. **(A)** Identification of *atsos2* mutant. Lanes labeled by odd numbers are PCR products of SALK_056101 specific forward and reverse primers (LP and RP), lanes labeled by even numbers are PCR products of LBb1.3 (BP) and RP. **(B)** Genotype of COM lines. **(C)** qRT-PCR detected the expression of *LiCIPK30*.

SUPPLEMENTARY FIGURE 7
Phylogenic tree of LiCIPKs. Whole genome duplication (WGD) events of angiosperm-plants are indicated on the branches (ζ, zeta seed plant-wide WGD; ε, epsilon angiosperm-wide WGD event; and γ, gamma triplicated of dicotyledon-wide WGD). The "T" indicates tandem repeat duplication happened in the ancestor of *L. indica*. WGD1 and WGD2 indicate two WGD events in *L. indica*.

References

Babu, B., Newberry, E., Dankers, H., Ritchie, L., Aldrich, J., Knox, G., et al. (2014). First Report of Xanthomonas axonopodis Causing Bacterial Leaf Spot on Crape Myrtle. *Plant Dis.* 98:841. doi: 10.1094/PDIS-10-13-1082-PDN

Birchler, J. A., and Veitia, R. A. (2021). One Hundred Years of Gene Balance: How Stoichiometric Issues Affect Gene Expression, Genome Evolution, and Quantitative Traits. *Cytogenet. Genome Res.* 161, 529–550. doi: 10.1159/000519592

Blanc, G., and Wolfe, K. H. (2004). Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. *Plant Cell* 16, 1667–1678. doi: 10.1105/tpc.021345

Chaves-Sanjuan, A., Sanchez-Barrena, M. J., Gonzalez-Rubio, J. M., Moreno, M., Ragel, P., Jimenez, M., et al. (2014). Structural basis of the regulatory mechanism of the plant CIPK family of protein kinases controlling ion homeostasis and abiotic stress. *Proc. Natl. Acad. Sci. U.S.A.* 111:E4532–E4541. doi: 10.1073/pnas.1407610111

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020). TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol. Plant* 13, 1194–1202. doi: 10.1016/j.molp.2020.06.009

Chen, X., Chen, G., Li, J., Hao, X., Tuerxun, Z., Chang, X., et al. (2021). A maize calcineurin B-like interacting protein kinase ZmCIPK42 confers salt stress tolerance. *Physiol. Plant* 171, 161–172. doi: 10.1111/ppl.13244

Chen, X., Gu, Z., Xin, D., Hao, L., Liu, C., Huang, J., et al. (2011). Identification and characterization of putative CIPK genes in maize. *J. Genet. Genomics* 38, 77–87. doi: 10.1016/j.jcg.2011.01.005

Cheong, Y. H., Pandey, G. K., Grant, J. J., Batistic, O., Li, L., Kim, B. G., et al. (2007). Two calcineurin B-like calcium sensors, interacting with protein kinase CIPK23, regulate leaf transpiration and root potassium uptake in Arabidopsis. *Plant J.* 52, 223–239. doi: 10.1111/j.1365-313X.2007.03236.x

Clark, J. W., and Donoghue, P. C. J. (2018). Whole-Genome Duplication and Plant Macroevolution. *Trends Plant Sci.* 23, 933–945. doi: 10.1016/j.tplants.2018.07.006

Clough, S. J., and Bent, A. F. (1998). Floral dip: A simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16, 735–743.

Dai, H., Zhu, Z., Wang, Z., Zhang, Z., Kong, W., and Miao, M. (2022). Galactinol synthase 1 improves cucumber performance under cold stress by enhancing assimilate translocation. *Hortic. Res.* 9:uhab063. doi: 10.1093/hr/uhab063

Dong, W., Xu, C., Liu, Y., Shi, J., Li, W., and Suo, Z. (2021). Chloroplast phylogenomics and divergence times of *Lagerstroemia* (*Lythraceae*). *BMC Genom.* 22:434. doi: 10.1186/s12864-021-07769-x

Feng, C., Feng, C., Lin, X., Liu, S., Li, Y., and Kang, M. (2021). A chromosome-level genome assembly provides insights into ascorbic acid accumulation and fruit softening in guava (*Psidium guajava*). *Plant Biotechnol. J.* 19, 717–730. doi: 10.1111/pbi.13498

Ghosh, S., Bheri, M., and Pandey, G. K. (2021). Delineating Calcium Signaling Machinery in Plants: Tapping the Potential through Functional Genomics. *Curr. Genom.* 22, 404–439. doi: 10.2174/1389202922666211130143328

Gong, D., Guo, Y., Jagendorf, A. T., and Zhu, J. K. (2002). Biochemical characterization of the Arabidopsis protein kinase SOS2 that functions in salt tolerance. *Plant Physiol.* 130, 256–264. doi: 10.1104/pp.004507

Guo, Y., Halfter, U., Ishitani, M., and Zhu, J. K. (2001). Molecular characterization of functional domains in the protein kinase SOS2 that is required for plant salt tolerance. *Plant Cell* 13, 1383–1400. doi: 10.1105/tpc.13.6.138

Halfter, U., Ishitani, M., and Zhu, J. K. (2000). The Arabidopsis SOS2 protein kinase physically interacts with and is activated by the calcium-binding protein SOS3. *Proc. Natl. Acad. Sci. U.S.A.* 97, 3735–3740. doi: 10.1073/pnas.97.7.3735

He, Z., Zhang, H., Gao, S., Lercher, M. J., Chen, W. H., and Hu, S. (2016). Evolview v2: An online visualization and management tool for customized and annotated phylogenetic trees. *Nucleic Acids Res.* 44:W236–W241. doi: 10.1093/nar/gkw370

Kanwar, P., Sanyal, S. K., Tokas, I., Yadav, A. K., Pandey, A., Kapoor, S., et al. (2014). Comprehensive structural, interaction and expression analysis of CBL and CIPK complement during abiotic stresses and development in rice. *Cell Calcium* 56, 81–95. doi: 10.1016/j.ceca.2014.05.003

Ketehouli, T., Zhou, Y. G., Dai, S. Y., Carther, K. F. I., Sun, D. Q., Li, Y., et al. (2021). A soybean calcineurin B-like protein-interacting protein kinase, GmPKS4, regulates plant responses to salt and alkali stresses. *J. Plant Physiol.* 256:153331. doi: 10.1016/j.jplph.2020.153331

Kim, K. W. (2021). Morphology and surface characteristics of the anamorphic stage of powdery mildew *Erysiphe australiana* on crape myrtle leaves. *Micron* 143:103013. doi: 10.1016/j.micron.2021.103013

Kleist, T. J., Spencley, A. L., and Luan, S. (2014). Comparative phylogenomics of the CBL-CIPK calcium-decoding network in the moss Physcomitrella, Arabidopsis, and other green lineages. *Front. Plant Sci.* 5:187. doi: 10.3389/fpls.2014.00187

Kolukisaoglu, U., Weinl, S., Blazevic, D., Batistic, O., and Kudla, J. (2004). Calcium sensors and their interacting protein kinases: Genomics of the Arabidopsis and rice CBL-CIPK signaling networks. *Plant Physiol.* 134, 43–58. doi: 10.1104/pp.103.033068

Kumar, S., Stecher, G., Suleski, M., and Hedges, S. B. (2017). TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. *Mol. Biol. Evol.* 34, 1812–1819. doi: 10.1093/molbev/msx116

Labib, R. M., Ayoub, N. A., Singab, A. B., Al-Azizi, M. M., and Sleem, A. (2013). Chemical constituents and pharmacological studies of *Lagerstroemia indica*. *Phytopharmacology* 4, 373–389.

Li, J., Jiang, M. M., Ren, L., Liu, Y., and Chen, H. Y. (2016). Identification and characterization of CBL and CIPK gene families in eggplant (*Solanum melongena* L.). *Mol. Genet. Genom.* 291, 1769–1781. doi: 10.1007/s00438-016-1218-8

Li, P., Zheng, T., Li, L., Zhuo, X., Jiang, L., Wang, J., et al. (2019). Identification and comparative analysis of the CIPK gene family and characterization of the cold stress response in the woody plant. *PeerJ* 7:e6847. doi: 10.7717/peerj.6847

Li, S., Wang, S., Wang, P., Gao, L., Yang, R., and Li, Y. (2020a). Label-free comparative proteomic and physiological analysis provides insight into leaf color variation of the golden-yellow leaf mutant of *Lagerstroemia indica*. *J. Proteom.* 228:103942. doi: 10.1016/j.jprot.2020.103942

Li, S., Zheng, T., Zhuo, X., Li, Z., Li, L., Li, P., et al. (2020b). Transcriptome profiles reveal that gibberellin-related genes regulate weeping traits in crape myrtle. *Hortic. Res.* 7:54. doi: 10.1038/s41438-020-0279-3

Li, T., Zhang, Y., Liu, Y., Li, X., Hao, G., Han, Q., et al. (2020c). Raffinose synthase enhances drought tolerance through raffinose synthesis or galactinol hydrolysis in maize and Arabidopsis plants. *J. Biol. Chem.* 295, 8064–8077. doi: 10.1074/jbc.RA120.013948

Li, Z., and Barker, M. S. (2020). Inferring putative ancient whole-genome duplications in the 1000 Plants (1KP) initiative: Access to gene family phylogenies and age distributions. *Gigascience* 9:giaa004. doi: 10.1093/gigascience/giaa004

Liu, H., Wang, Y. X., Li, H., Teng, R. M., Wang, Y., and Zhuang, J. (2019). Genome-Wide Identification and Expression Analysis of Calcineurin B-Like Protein and Calcineurin B-Like Protein-Interacting Protein Kinase Family Genes in Tea Plant. *DNA Cell Biol.* 38, 824–839. doi: 10.1089/dna.2019.4697

Liu, J., Ishitani, M., Halfter, U., Kim, C. S., and Zhu, J. K. (2000). The *Arabidopsis thaliana* SOS2 gene encodes a protein kinase that is required for salt tolerance. *Proc. Natl. Acad. Sci. U.S.A.* 97, 3730–3734.

Liu, S., Lv, Z., Liu, Y., Li, L., and Zhang, L. (2018). Network analysis of ABA-dependent and ABA-independent drought responsive genes in *Arabidopsis thaliana*. *Genet. Mol. Biol.* 41, 624–637. doi: 10.1590/1678-4685-GMB-2017-0229

Liu, X., Zhang, H., Ma, L., Wang, Z., and Wang, K. (2020). Genome-Wide Identification and Expression Profiling Analysis of the Trihelix Gene Family Under Abiotic Stresses in *Medicago truncatula*. *Genes* 11:1189. doi: 10.3390/genes11111389

Liu, Y., Zhang, L., Meng, S., Liu, Y., Zhao, X., Pang, C., et al. (2020). Expression of galactinol synthase from *Ammopiptanthus nanus* in tomato improves tolerance to cold stress. *J. Exp. Bot.* 71, 435–449. doi: 10.1093/jxb/erz450

Liu, B., Fan, H., Sun, C., Yuan, M., Geng, X., Ding, X., et al. (2022). New insights into the role of chrysanthemum calcineurin B-like interacting protein kinase CmCIPK23 in nitrate signaling in *Arabidopsis* roots. *Sci. Rep.* 12, 1018. doi: 10.1038/s41598-021-04758-8

Luo, X., Li, H., Wu, Z., Yao, W., Zhao, P., Cao, D., et al. (2020). The pomegranate (*Punica granatum* L.) draft genome dissects genetic divergence between soft- and hard-seeded cultivars. *Plant Biotechnol. J.* 18, 955–968. doi: 10.1111/pbi.13260

Lynch, M., and Conery, J. S. (2000). The evolutionary fate and consequences of duplicate genes. *Science* 290, 1151–1155. doi: 10.1126/science.290.5494.1151

Ma, Q. J., Sun, M. H., Lu, J., Liu, Y. J., You, C. X., and Hao, Y. J. (2017). An apple CIPK protein kinase targets a novel residue of AREB transcription factor for ABA-dependent phosphorylation. *Plant Cell Environ.* 40, 2207–2219. doi: 10.1111/pce.13013

Ma, X., Gai, W. X., Qiao, Y. M., Ali, M., Wei, A. M., Luo, D. X., et al. (2019). Identification of CBL and CIPK gene families and functional characterization of CaCIPK1 under Phytophthora capsici in pepper (*Capsicum annuum* L.). *BMC Genom.* 20:775. doi: 10.1186/s12864-019-6125-z

Ma, X., Li, Y., Gai, W. X., Li, C., and Gong, Z. H. (2021). The CaCIPK3 gene positively regulates drought tolerance in pepper. *Hortic. Res.* 8:216. doi: 10.1038/s41438-021-00651-7

Morales de Los Ríos, L., Corratgé-Faillie, C., Raddatz, N., Mendoza, I., Lindahl, M., de Angeli, A., et al. (2021). The Arabidopsis protein NPF6.2/NRT1.4 is a plasma membrane nitrate transporter and a target of protein kinase CIPK23. *Plant Physiol. Biochem.* 168, 239–251. doi: 10.1016/j.plaphy.2021.10.016

Myburg, A. A., Grattapaglia, D., Tuskan, G. A., Hellsten, U., Hayes, R. D., Grimwood, J., et al. (2014). The genome of Eucalyptus grandis. *Nature* 510, 356–362. doi: 10.1038/nature13308

Oh, S. J., Song, S. I., Kim, Y. S., Jang, H. J., Kim, S. Y., Kim, M., et al. (2005). Arabidopsis CBF3/DREB1A and ABF3 in transgenic rice increased tolerance to abiotic stress without stunting growth. *Plant Physiol.* 138, 341–351.

One Thousand Plant Transcriptomes, I. (2019). One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* 574, 679–685. doi: 10.1038/s41586-019-1693-2

Peng, Y., Hou, F., Bai, Q., Xu, P., Liao, Y., Zhang, H., et al. (2018). Rice Calcineurin B-Like Protein-Interacting Protein Kinase 31 (OsCIPK31) Is Involved

in the Development of Panicle Apical Spikelets. *Front. Plant Sci.* 9:1661. doi: 10.3389/fpls.2018.01661

Qiao, Z., Liu, S., Zeng, H., Li, Y., Wang, X., Chen, Y., et al. (2019). Exploring the Molecular Mechanism underlying the Stable Purple-Red Leaf Phenotype in *Lagerstroemia indica* cv. Ebony Embers. *Int. J. Mol. Sci.* 20:5636. doi: 10.3390/ijms20225636

Qin, G., Xu, C., Ming, R., Tang, H., Guyot, R., Kramer, E. M., et al. (2017). The pomegranate (*Punica granatum* L.) genome and the genomics of punicalagin biosynthesis. *Plant J.* 91, 1108–1128. doi: 10.1111/tpj.13625

Qiu, Q. S., Guo, Y., Dietrich, M. A., Schumaker, K. S., and Zhu, J. K. (2002). Regulation of SOS1, a plasma membrane Na^+/H^+ exchanger in *Arabidopsis thaliana*, by SOS2 and SOS3. *Proc. Natl. Acad. Sci. U.S.A.* 99, 8436–8441. doi: 10.1073/pnas.122224699

Quan, R., Lin, H., Mendoza, I., Zhang, Y., Cao, W., Yang, Y., et al. (2007). SCABP8/CBL10, a putative calcium sensor, interacts with the protein kinase SOS2 to protect Arabidopsis shoots from salt stress. *Plant Cell* 19, 1415–1431. doi: 10.1105/tpc.106.042291

Ragel, P., Ródenas, R., García-Martín, E., Andrés, Z., Villalta, I., Nieves-Cordones, M., et al. (2015). The CBL-Interacting Protein Kinase CIPK23 Regulates HAK5-Mediated High-Affinity K^+ Uptake in Arabidopsis Roots. *Plant Physiol.* 169, 2863–2873. doi: 10.1104/pp.15.01401

Ródenas, R., and Vert, G. (2021). Regulation of Root Nutrient Transporters by CIPK23: 'One Kinase to Rule Them All'. *Plant Cell Physiol.* 62, 553–563. doi: 10.1093/pcp/pcaa156

Sánchez-Barrena, M. J., Fujii, H., Angulo, I., Martínez-Ripoll, M., Zhu, J. K., and Albert, A. (2007). The structure of the C-terminal domain of the protein kinase AtSOS2 bound to the calcium sensor AtSOS3. *Mol. Cell* 26, 427–435.

Shi, A., and Mmbaga, M. T. (2006). Perpetuation of Powdery Mildew Infection and Identification of *Erysiphe australiana* as the Crape Myrtle Pathogen in Mid-Tennessee. *Plant Dis.* 90, 1098–1101. doi: 10.1094/PD-90-1098

Straub, T., Ludewig, U., and Neuhauser, B. (2017). The Kinase CIPK23 Inhibits Ammonium Transport in *Arabidopsis thaliana*. *Plant Cell* 29, 409–422. doi: 10.1105/tpc.16.00806

Subramanian, B., Gao, S., Lercher, M. J., Hu, S., and Chen, W. H. (2019). Evolview v3: A webserver for visualization, annotation, and management of phylogenetic trees. *Nucl. Acids Res.* 47:W270–W275. doi: 10.1093/nar/gkz357

Sun, T., Wang, Y., Wang, M., Li, T., Zhou, Y., Wang, X., et al. (2015). Identification and comprehensive analyses of the CBL and CIPK gene families in wheat (*Triticum aestivum* L.). *BMC Plant Biol.* 15:269. doi: 10.1186/s12870-015-0657-4

Tang, H., Wang, X., Bowers, J. E., Ming, R., Alam, M., and Paterson, A. H. (2008). Unraveling ancient hexaploidy through multiply-aligned angiosperm gene maps. *Genome Res.* 18, 1944–1954. doi: 10.1101/gr.080978.108

Tian, Q., Zhang, X., Yang, A., Wang, T., and Zhang, W. H. (2016). CIPK23 is involved in iron acquisition of Arabidopsis by affecting ferric chelate reductase activity. *Plant Sci.* 246, 70–79. doi: 10.1016/j.plantsci.2016.01.010

Tripathi, V., Parasuraman, B., Laxmi, A., and Chattopadhyay, D. (2009). CIPK6, a CBL-interacting protein kinase is required for development and salt tolerance in plants. *Plant J.* 58, 778–790. doi: 10.1111/j.1365-313X.2009.03812.x

Wang, S., Wang, P., Gao, L., Yang, R., Li, L., Zhang, E., et al. (2017). Characterization and Complementation of a Chlorophyll-Less Dominant Mutant GL1 in *Lagerstroemia indica*. *DNA Cell Biol.* 36, 354–366. doi: 10.1089/dna.2016.3573

Wang, X., Shi, W., and Rinehart, T. (2015). Transcriptomes That Confer to Plant Defense against Powdery Mildew Disease in *Lagerstroemia indica*. *Int. J. Genet. Genom.* 2015:528395. doi: 10.1155/2015/528395

Wang, Z., Su, G., Li, M., Ke, Q., Kim, S. Y., Li, H., et al. (2016). Overexpressing Arabidopsis ABF3 increases tolerance to multiple abiotic stresses and reduces leaf size in alfalfa. *Plant Physiol. Biochem.* 109, 199–208. doi: 10.1016/j.plaphy.2016.09.020

Xi, Y., Liu, J., Dong, C., and Cheng, Z.-M. M. (2017). The CBL and CIPK Gene Family in Grapevine (*Vitis vinifera*): Genome-Wide Analysis and Expression Profiles in Response to Various Abiotic Stresses. *Front. Plant Sci.* 8:978. doi: 10.3389/fpls.2017.00978

Xu, W., Shen, W., Ma, J., Ya, R., Zheng, Q., Wu, N., et al. (2020). Role of an Amur grape CBL-interacting protein kinase VaCIPK02 in drought tolerance by modulating ABA signaling and ROS production. *Environ. Exp. Bot.* 172:103999. doi: 10.1016/j.envexpbot.2020.103999

Yang, E. J., Lee, J. S., Song, B. B., Yun, C. Y., Kim, D. H., and Kim, I. S. (2011). Anti-inflammatory effects of ethanolic extract from *Lagerstroemia indica* on airway inflammation in mice. *J. Ethnopharmacol.* 136, 422–427. doi: 10.1016/j.jep.2010.05.066

Ye, Y. J., Liu, Y., Cai, M., He, D., Shen, J. S., Ju, Y. Q., et al. (2015). Screening of molecular markers linked to dwarf trait in crape myrtle by bulked segregant analysis. *Genet. Mol. Res.* 14, 4369–4380. doi: 10.4238/2015.April.30.10

Yin, X., Wang, Q., Chen, Q., Xiang, N., Yang, Y., and Yang, Y. (2017). Genome-Wide Identification and Functional Analysis of the Calcineurin B-like Protein and Calcineurin B-like Protein-Interacting Protein Kinase Gene Families in Turnip (*Brassica rapa* var. rapa). *Front. Plant Sci.* 8:1191. doi: 10.3389/fpls.2017.01191

Yoshida, T., Fujita, Y., Sayama, H., Kidokoro, S., Maruyama, K., Mizoi, J., et al. (2010). AREB1, AREB2, and ABF3 are master transcription factors that cooperatively regulate ABRE-dependent ABA signaling involved in drought stress tolerance and require ABA for full activation. *Plant J.* 61, 672–685. doi: 10.1111/j.1365-313X.2009.04092.x

Yu, C., Lian, B., Fang, W., Guo, A., Ke, Y., Jiang, Y., et al. (2021). Transcriptome-based analysis reveals that the biosynthesis of anthocyanins is more active than that of flavonols and proanthocyanins in the colorful flowers of *Lagerstroemia indica*. *Biol. Futur.* 72, 473–488. doi: 10.1007/s42977-021-00094-0

Yu, Y., Xia, X., Yin, W., and Zhang, H. (2007). Comparative genomic analysis of CIPK gene family in Arabidopsis and Populus. *Plant Growth Regul.* 52, 101–110. doi: 10.1007/s10725-007-9165-3

Zhang, H., Gao, S., Lercher, M. J., Hu, S., and Chen, W.-H. (2012). EvolView, an online tool for visualizing, annotating and managing phylogenetic trees. *Nucl. Acids Res.* 40:W569–W572. doi: 10.1093/nar/gks576

Zhang, H., Yang, B., Liu, W.-Z., Li, H., Wang, L., Wang, B., et al. (2014). Identification and characterization of CBL and CIPK gene families in canola (*Brassica napus* L.). *BMC Plant Biol.* 14:8. doi: 10.1186/1471-2229-14-8

Zhang, J., Shi, S. Z., Jiang, Y., Zhong, F., Liu, G., Yu, C., et al. (2021). Genome-wide investigation of the AP2/ERF superfamily and their expression under salt stress in Chinese willow (*Salix matsudana*). *PeerJ* 9:e11076. doi: 10.7717/peerj.11076

Zhang, X., Li, X., Zhao, R., Zhou, Y., and Jiao, Y. (2020). Evolutionary strategies drive a balance of the interacting gene products for the CBL and CIPK gene families. *New Phytol.* 226, 1506–1516. doi: 10.1111/nph.16445

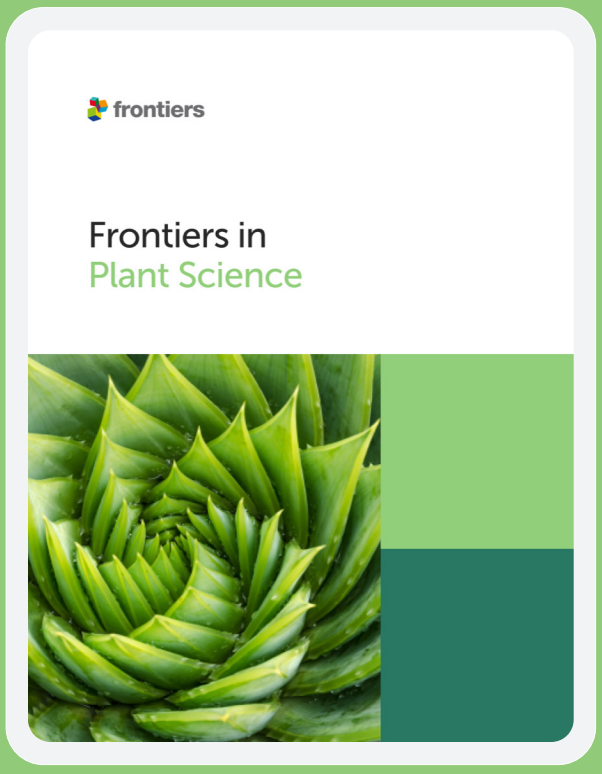
Frontiers in
Plant Science

Cultivates the science of plant biology and its applications

The most cited plant science journal, which advances our understanding of plant biology for sustainable food security, functional ecosystems and human health.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact