frontiers | Research Topics

Genetics and epigenetics: Plausible role in development of climate resilient crops

Edited by
Vijay Gahlaut, Abhishek Bohra, Dragan Perovic and Rajeev K. Varshney

Published in
Frontiers in Genetics



FRONTIERS EBOOK COPYRIGHT STATEMENT

The copyright in the text of individual articles in this ebook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this ebook is the property of Frontiers.

Each article within this ebook, and the ebook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this ebook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or ebook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-83252-076-5
DOI 10.3389/978-2-83252-076-5

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

Genetics and epigenetics: Plausible role in development of climate resilient crops

Topic editors

Vijay Gahlaut — Chandigarh University, India
Abhishek Bohra — Murdoch University, Australia
Dragan Perovic — Julius Kühn-Institut, Germany
Rajeev K. Varshney — Murdoch University, Australia

Citation

Gahlaut, V., Bohra, A., Perovic, D., Varshney, R. K., eds. (2023). *Genetics and epigenetics: Plausible role in development of climate resilient crops*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83252-076-5

Table of contents

06 Editorial: Genetics and epigenetics: Plausible role in development of climate resilient crops
Abhishek Bohra, Vijay Gahlaut, Dragan Perovic and Rajeev K. Varshney

09 Genome-Wide Identification and Expression Analysis of AP2/ERF Transcription Factor Related to Drought Stress in Cultivated Peanut (*Arachis hypogaea* L.)
Mengjie Cui, Muhammad Salman Haider, Pengpei Chai, Junjia Guo, Pei Du, Hongyan Li, Wenzhao Dong, Bingyan Huang, Zheng Zheng, Lei Shi, Xinyou Zhang and Suoyi Han

26 Stress Conditions Modulate the Chromatin Interactions Network in *Arabidopsis*
Vikash Kumar Yadav, Swadha Singh, Amrita Yadav, Neha Agarwal, Babita Singh, Siddhi Kashinath Jalmi, Vrijesh Kumar Yadav, Vipin Kumar Tiwari, Verandra Kumar, Raghvendra Singh and Samir Vishwanath Sawant

39 Genomic Selection: A Tool for Accelerating the Efficiency of Molecular Breeding for Development of Climate-Resilient Crops
Neeraj Budhlakoti, Amar Kant Kushwaha, Anil Rai, K K Chaturvedi, Anuj Kumar, Anjan Kumar Pradhan, Uttam Kumar, Rajeev Ranjan Kumar, Philomin Juliana, D C Mishra and Sundeep Kumar

56 Transcriptome Analysis of Bread Wheat Genotype KRL3-4 Provides a New Insight Into Regulatory Mechanisms Associated With Sodicity (High pH) Tolerance
Geeta Prasad, Shikha Mittal, Arvind Kumar, Divya Chauhan, Tanmaya Kumar Sahu, Sundeep Kumar, Rakesh Singh, Mahesh C. Yadav and Amit Kumar Singh

76 An Epigenetic Alphabet of Crop Adaptation to Climate Change
Francesco Guarino, Angela Cicatelli, Stefano Castiglione, Dolores R. Agius, Gul Ebru Orhun, Sotirios Fragkostefanakis, Julie Leclercq, Judit Dobránszki, Eirini Kaiserli, Michal Lieberman-Lazarovich, Merike Sõmera, Cecilia Sarmiento, Cristina Vettori, Donatella Paffetti, Anna M. G. Poma, Panagiotis N. Moschou, Mateo Gašparović, Sanaz Yousefi, Chiara Vergata, Margot M. J. Berger, Philippe Gallusci, Dragana Miladinović and Federico Martinelli

94 RNA-Seq-Based Transcriptomics Study to Investigate the Genes Governing Nitrogen Use Efficiency in Indian Wheat Cultivars
Sarabjit Kaur, M. Shamshad, Suruchi Jindal, Amandeep Kaur, Satinder Singh, Achla sharma and Satinder Kaur

109 Uncovering the Epigenetic Marks Involved in Mediating Salt Stress Tolerance in Plants
Garima Singroha, Satish Kumar, Om Prakash Gupta, Gyanandra Pratap Singh and Pradeep Sharma

119 Moving Beyond DNA Sequence to Improve Plant Stress Responses
Faisal Saeed, Usman Khalid Chaudhry, Allah Bakhsh, Ali Raza, Yasir Saeed, Abhishek Bohra and Rajeev K. Varshney

134 Indian Wheat Genomics Initiative for Harnessing the Potential of Wheat Germplasm Resources for Breeding Disease-Resistant, Nutrient-Dense, and Climate-Resilient Cultivars
Sundeep Kumar, Sherry R. Jacob, Reyazul Rouf Mir, V. K. Vikas, Pawan Kulwal, Tilak Chandra, Satinder Kaur, Uttam Kumar, Suneel Kumar, Shailendra Sharma, Ravinder Singh, Sai Prasad, Anju Mahendru Singh, Amit Kumar Singh, Jyoti Kumari, M. S. Saharan, Subhash Chander Bhardwaj, Manoj Prasad, Sanjay Kalia and Kuldeep Singh

153 Genome-Wide Identification and Characterization of *Receptor-Like Protein Kinase 1 (RPK1)* Gene Family in *Triticum aestivum* Under Drought Stress
Amna Abdul Rahim, Muhammad Uzair, Nazia Rehman, Obaid Ur Rehman, Nageen Zahra and Muhammad Ramzan Khan

172 Identification and Characterization of DNA Demethylase Genes and Their Association With Thermal Stress in Wheat (*Triticum aestivum* L.)
Vijay Gahlaut, Harsha Samtani, Tinku Gautam and Paramjit Khurana

186 Epigenomics as Potential Tools for Enhancing Magnitude of Breeding Approaches for Developing Climate Resilient Chickpea
B. S. Chandana, Rohit Kumar Mahto, Rajesh Kumar Singh, Rebecca Ford, Niloofar Vaghefi, Santosh Kumar Gupta, Hemant Kumar Yadav, Murli Manohar and Rajendra Kumar

211 Cold adaptation strategies in plants—An emerging role of epigenetics and antifreeze proteins to engineer cold resilient plants
Satyakam, Gaurav Zinta, Rajesh Kumar Singh and Rajiv Kumar

229 Development of *de-novo* transcriptome assembly and SSRs in allohexaploid Brassica with functional annotations and identification of heat-shock proteins for thermotolerance
Kaushal Pratap Singh, Preetesh Kumari and Devendra Kumar Yadava

247 Dissection of a grain yield QTL from wild emmer wheat reveals sub-intervals associated with culm length and kernel number
Mathieu Deblieck, Gergely Szilagyi, Fatiukha Andrii, Yehoshua Saranga, Madita Lauterberg, Kerstin Neumann, Tamar Krugman, Dragan Perovic, Klaus Pillen and Frank Ordon

267 Genome-wide association studies reveal putative QTLs for physiological traits under contrasting phosphorous conditions in wheat (*Triticum aestivum* L.)
Palaparthi Dharmateja, Rajbir Yadav, Manjeet Kumar, Prashanth Babu, Neelu Jain, Pranab Kumar Mandal, Rakesh Pandey, Manoj Shrivastava, Kiran B. Gaikwad, Naresh Kumar Bainsla, Vipin Tomar, S. Sugumar, Nasreen Saifi and Rumesh Ranjan

frontiers | Frontiers in Genetics

TYPE Editorial
PUBLISHED 10 March 2023
DOI 10.3389/fgene.2023.1165843

Check for updates

OPEN ACCESS

EDITED AND REVIEWED BY
Andrew H. Paterson,
University of Georgia, United States

*CORRESPONDENCE
Abhishek Bohra,
abhishek.bohra@murdoch.edu.au
Rajeev K. Varshney,
rajeev.varshney@murdoch.edu.au

SPECIALTY SECTION
This article was submitted to Plant Genomics, a section of the journal Frontiers in Genetics

RECEIVED 14 February 2023
ACCEPTED 06 March 2023
PUBLISHED 10 March 2023

CITATION
Bohra A, Gahlaut V, Perovic D and Varshney RK (2023), Editorial: Genetics and epigenetics: Plausible role in development of climate resilient crops. *Front. Genet.* 14:1165843. doi: 10.3389/fgene.2023.1165843

COPYRIGHT
© 2023 Bohra, Gahlaut, Perovic and Varshney. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Editorial: Genetics and epigenetics: Plausible role in development of climate resilient crops

Abhishek Bohra[1]*, Vijay Gahlaut[2], Dragan Perovic[3] and Rajeev K. Varshney[1]*

[1]State Agricultural Biotechnology Centre, Centre for Crop and Food Innovation, Food Futures Institute, Murdoch University, Murdoch, WA, Australia, [2]University Centre for Research and Development, Chandigarh University, Mohali, Punjab, India, [3]Federal Research Centre for Cultivated Plants, Julius Kuehn Institute, Institute for Resistance Research and Stress Tolerance, Quedlinburg, Germany

KEYWORDS

candidate genes, climate-resilient crops, genetics, epigenetics, stress response

Editorial on the Research Topic
Genetics and epigenetics: Plausible role in development of climate resilient crops

Rising weather extremes and evolving pest and pathogen dynamics associated with climate change exert profound negative impacts on global crop production. Plants are sessile organisms, and intrinsic mechanisms enable them to respond to a variety of challenges posed by stressful conditions. Technological advances, especially in development of new high-throughput sequencing technologies, in recent years have contributed greatly to improving our ability to understand the genetic and epigenetic changes occurring in plants, particularly in response to stresses. This Research Topic on "*Genetics and epigenetics: Plausible role in development of climate resilient crops*" presents 16 articles from leading experts in this field. We summarize key highlights of these articles in this editorial.

The epigenetic modifications are defined as heritable changes occurring beyond DNA sequences. In this context, Saeed et al. reviewed the recent advances to analyse epigenetic changes when plants are exposed to abiotic and abiotic stress conditions. The article underscores emerging techniques to analyse genome-wide epigenetic modifications such as sodium bisulphite sequencing, methylated DNA immunoprecipitation (MeDIP) and integration of these with evolving sequencing platforms including the next and third generation platforms. The authors also discuss how the improved methods can contribute to separate the contributions of epigenetic modifications to phenotypes from other source of variations, e.g., sequence variations. The article also explores the possibilities to implement modern genome editing tools such as engineered endonucleases to analyse the epigenetic changes for improving plant stress response. In another article, Singroha et al. reviewed epigenetic dynamics and resulting alterations in gene expression reported in plants under salt-stressed scenarios. The article discusses the influence of salt stress on plant epigenetic machinery that involves DNA methylation, histone modifications, histone variants and non-coding RNA molecules including the long non-coding RNAs and microRNAs. For a better understanding of the epigenetic variations; Guarino et al. define the "epigenetic code" and recommend translating this code into "epigenetic

syntax" for developing new crops carrying climate adaptation traits. The article highlights challenges that have cropped up as the field of epigenetics evolves. The authors call for modern sequencing technologies, optimized breeding strategies, standardized workflows for data analysis and integration of multi-omics data, to efficiently exploit epigenetic variation for crop improvement. DNA methylation is the most common epigenetic phenomenon leading to transcriptional silencing of genes and enabled by four different methyltransferases (MTases) in three sequence contexts: CG, CHG and CHH. Equally important to DNA methylation homeostasis are demethylases (dMTases) that remain instrumental to C methylation removal. Gahlaut et al. identified and characterized 12 dMTase in wheat by analysing the genome sequence information. The identified genes belonged to two subfamilies: DEMETER-LIKE (DML) and REPRESSOR OF SILENCING1 (ROS1), and mapped onto nine chromosomes. The study suggested a higher number of dMTase genes in wheat than other plant species, which are reported to vary between 2 and 10 in different plant species. Using these genes, the study demonstrated phylogenetic relationships, gene structure, regulatory function, nuclear localization signals (NLS) and DNA marker development. Analysis of gene expression patterns indicated a role for these genes in tolerance to heat stress in wheat.

Transcription factors (TFs) are known to regulate the cellular processes. AP2/ERF family represents the largest group of TFs among the 60 different TF families discovered so far in plants (Joshi et al., 2016). The availability of whole genome sequence has facilitated genome-wide identification and characterization TFs in plants. Since the role of the AP2/ERF family was demonstrated in flower development in Arabidopsis (Jofuku et al., 1994), growing literature has provided evidence in support of involvement of the AP2/ERF family in plant growth and development and stress responsiveness. Cui et al. used genome sequence information of Tifrunner, a popular groundnut (*Arachis hypogaea*) variety and identified 185 AP2/ERF family genes belonging to five sub-families e.g., AP2 (59), ERF (76), DREB (41), RAV (4), and Soloist (5). The study examined the phylogenetic relations and intron-exon structure and demonstrated that the identified genes are unevenly distributed among 20 chromosomes. Further analysis of the orthologous gene pairs between *A. hypogaea, Medicago truncatula* and *Glycine max* predicted the divergence times between *A. hypogaea* and *M. truncatula* (64.7 Mya), and *A. hypogaea* and *Glycine max* (66.44 Mya). Differential expression of 35 selected AP2/ERF family genes supported their roles in abiotic stress response. A similar survey of wheat genome sequence for receptor-like kinase (RLK) gene family led to the identification and characterization of 15 *TaRPK* genes (Rahim et al.). Gene expression analysis of tolerant (Pakistan 13 and Galaxy) and susceptible (Shafaq) varieties suggested *TaRPK*'s participation in drought tolerance in wheat (*Triticum aestivum*). Also, the cis-regulatory element (CRE) prediction showed abundance of drought-responsive elements for binding in the promoter regions.

Next-generation sequencing (NGS) has revolutionized the field of functional genomics by facilitating detailed inquiries into regulation of gene expression. Singh et al. employed Illumina technology to construct a *de novo* transcriptome assembly of an allohexaploid *Brassica* (AABBSS), comprising 486,066 transcripts. The novel allohexaploid has resulted from somatic hybridization between an amphidiploid *Brassica juncea* (AABB) and diploid *Sinapsis alba* (SS). In wheat, an NGS-based profiling of flag leaf transcriptome of cultivar KRL 3–4 increased understanding of its high level of tolerance against sodicity (Prasad et al.). The analysis revealed a set of 1,980 genes that respond differentially to sodicity stress. Authors provide a list of 18 candidate genes and 39 SNPs potentially associated with the sodicity-responsive genes. Similarly, RNA-Seq analysis of the root and shoot transcriptomes of PBW677 (nitrogen-efficient) and PBW703 (N-inefficient) revealed differential expression of 2,406 genes between the two contrasting genotypes. Nitrogen use efficiency (NUE) is an important breeding target in wheat improvement, owing to the increasing environmental risks associated with greater use of nitrogen fertilisers. The higher nitrogen efficiency of PBW 677 could be explained by expression abundance of the genes belonging to nitrogen metabolism and protein kinases.

By using the high-throughput chromosome conformation capture (Hi-C) technique, Yadav et al. examined the impact on stress conditions and hormone application in the dynamics of chromatin interactions in *Arabidopsis*. The study showed that the changes in chromatin interactions from stress conditions did not alter the expression profiles of the interacting genes. Interacting genes were found to be enriched in the heterochromatic regions and likely to belong to the same epigenetic state.

An article by Kumar et al. provides an overview of the Indian Wheat Genomics Initiative aimed at harnessing the untapped genetic potential of the wheat germplasm collection held at National genebank of ICAR-National Bureau of Plant Genetic Resources (NBPGR), India. NBPGR, India houses more than 31,000 wheat accessions from 51 species, representing the largest germplasm collection of wheat in Asia. Major challenges that limit the progress of Indian wheat improvement programs include a range of biotic (rust, Karnal bunt, Fusarium head blight, spot blotch, powdery mildew and other pathogens/pests) and abiotic (drought, heat and salinity) stresses. Other important breeding targets of Indian wheat improvement programs include nutrition and quality traits and nutrient use efficiency. Chandana et al. discuss the role of epigenomics for developing improved chickpea that can withstand agricultural conditions exposed to a variety of biotic and abiotic stresses. While citing recent research demonstrating the role of epigenetics stress response in chickpea and other legume crops, the authors advocate for embracing newer technologies to profile epigenomic variations and multi-disciplinary approaches to enable their efficient deployment in chickpea improvement programs by combining multi-omics science and targeted epigenetic manipulation.

Rising temperature extremities negatively impact normal growth and development of plants. In this context, low temperature leading to chilling stress (0°C and 15°C) and freezing stress (<0°C) remains a greater concern for sustaining global crop production (Jha et al., 2017). Satyakam et al. explain genetic, epigenetic, physiological and biochemical basis of cold adaptation in plants, with emphasis on cold acclimation. The authors outline strategies to improve cold tolerance in plant, such as targeted manipulation of anti-freeze proteins.

Identification and manipulation of genomic loci associated with climate adaptation is crucial for developing future crops.

Artificially created populations (biparental/multiparent population) and diverse collections have facilitated understanding the genetic architectures of complex plant traits for (Bohra et al., 2020; Varshney et al., 2021a; Varshney et al., 2021b). A QTL for high grain yield and harvest index on chromosome 2BS was transferred from wild emmer (*Triticum turgidum* subsp. dicoccoides) to durum wheat cultivar Uzan (*Triticum turgidum* subsp. *durum*) following fine mapping and marker-assisted backcrossing approach (Deblieck et al.). The study shows implications of a high-throughput phenotyping platform for monitoring plant response under stress conditions. The resulting set of 2B introgression lines carrying QTL for culm length and kernel number would serve as a good resource for breeding wheat for water limiting conditions. A popular alternative to biparental QTL mapping is genome-wide association studies (GWAS) which was implemented by Dharmateja et al. for identification of genomic loci associated with phosphorus use efficiency. The study identified 45 QTL in P-limiting conditions in 158 wheat genotypes including popular varieties and advanced breeding lines analysed with Axiom BreedWheat 35 K genotyping array. The need for identification of a significant set of genetic markers to inform selection decisions is bypassed by genomic selection (Crossa et al., 2017; Bhat et al., 2021). In this Research Topic, the application of genomic selection (GS) was discussed in improving stress tolerance and quality traits of various crop species including cereals, pulses, oilseeds and horticultural crops (Budhlakoti et al.). The authors compare different genomic prediction models and outline key considerations related to genomic prediction accuracies in plant breeding programs. The cutting-edge genomic tools including whole genome sequence, candidate or causative genes, gene-trait associations or associated DNA markers, and genomic prediction models reported in these studies would be helpful for designing breeding strategies to obtain modern cultivars having new traits enabling cultivation in target production environments. As shown by Kumar et al., the multi-disciplinary initiatives will leverage of modern genomic resources and methods like GWAS and GS, thus paving the way for "Genebank genomics" and use of diverse germplasm collections for sustainable crop improvement.

Author contributions

AB and RKV prepared the first draft of the editorial. VG and DP contributed to different versions of the manuscript. All authors approved the submitted version.

Acknowledgments

AB and RKV are thankful to Food Futures Institute, Murdoch University (Australia), for financial support. We thank authors of the articles published in this Research Topic for their valuable contributions and the referees for their rigorous review. We also thank the Frontiers team for their support.

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.

References

Bhat, J. A., Yu, D., Bohra, A., Ganie, S. A., and Varshney, R. K. (2021). Features and applications of haplotypes in crop breeding. *Commun. Biol.* 4, 1266. doi:10.1038/s42003-021-02782-y

Bohra, A., Chand Jha, U., Godwin, I. D., and Kumar Varshney, R. (2020). Genomic interventions for sustainable agriculture. *Plant Biotechnol. J.* 18, 2388–2405. doi:10.1111/pbi.13472

Crossa, J., Pérez-Rodríguez, P., Cuevas, J., Montesinos-López, O., Jarquín, D., de Los Campos, G., et al. (2017). Genomic selection in plant breeding: Methods, models, and perspectives. *Trends Plant Sci.* 22, 961–975. doi:10.1016/j.tplants.2017.08.011

Jha, U. C., Bohra, A., and Jha, R. (2017). Breeding approaches and genomics technologies to increase crop yield under low-temperature stress. *Plant Cell Rep.* 36, 1–35. doi:10.1007/s00299-016-2073-0

Jofuku, K. D., den Boer, B. G., Van Montagu, M., and Okamuro, J. K. (1994). Control of Arabidopsis flower and seed development by the homeotic gene APETALA2. *Plant Cell* 6, 1211–1225. doi:10.1105/tpc.6.9.1211

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

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

Varshney, R. K., Bohra, A., Yu, J., Graner, A., Zhang, Q., and Sorrells, M. E. (2021a). Designing future crops: Genomics-assisted breeding comes of age. *Trends Plant Sci.* 26, 631–649. doi:10.1016/j.tplants.2021.03.010

frontiers in Genetics

ORIGINAL RESEARCH
published: 13 October 2021
doi: 10.3389/fgene.2021.750761

Genome-Wide Identification and Expression Analysis of AP2/ERF Transcription Factor Related to Drought Stress in Cultivated Peanut (*Arachis hypogaea* L.)

Mengjie Cui[1,2], Muhammad Salman Haider[3], Pengpei Chai[2], Junjia Guo[2], Pei Du[1,2], Hongyan Li[2], Wenzhao Dong[2], Bingyan Huang[2], Zheng Zheng[2], Lei Shi[2], Xinyou Zhang[1,2] and Suoyi Han[1,2]**

[1]College of Agriculture, Nanjing Agricultural University, Nanjing, China, [2]Henan Institute of Crop Molecular Breeding, Henan Academy of Agricultural Science/Key Laboratory of Oil Crops in Huang-Huai-Hai Plains, Ministry of Agriculture/Henan Provincial Key Laboratory for Oil Crops Improvement, Zhengzhou, China, [3]Department of Horticulture, Ghazi University, Dera Ghazi Khan, Pakistan

OPEN ACCESS

Edited by:
Abhishek Bohra,
Indian Institute of Pulses Research (ICAR), India

Reviewed by:
Hui Song,
Qingdao Agricultural University, China
Cheng Qin,
Zunyi Vocational and Technical College, China

***Correspondence:**
Xinyou Zhang
haasxinyou@163.com
Suoyi Han
suoyi_han@126.com

Specialty section:
This article was submitted to Plant Genomics, a section of the journal Frontiers in Genetics

Received: 31 July 2021
Accepted: 17 September 2021
Published: 13 October 2021

Citation:
Cui M, Haider MS, Chai P, Guo J, Du P, Li H, Dong W, Huang B, Zheng Z, Shi L, Zhang X and Han S (2021) Genome-Wide Identification and Expression Analysis of AP2/ERF Transcription Factor Related to Drought Stress in Cultivated Peanut (Arachis hypogaea L.). Front. Genet. 12:750761. doi: 10.3389/fgene.2021.750761

APETALA2/ethylene response element-binding factor (AP2/ERF) transcription factors (TFs) have been found to regulate plant growth and development and response to various abiotic stresses. However, detailed information of *AP2/ERF* genes in peanut against drought has not yet been performed. Herein, 185 AP2/ERF TF members were identified from the cultivated peanut (*A. hypogaea cv.* Tifrunner) genome, clustered into five subfamilies: AP2 (APETALA2), ERF (ethylene-responsive-element-binding), DREB (dehydration-responsive-element-binding), RAV (related to ABI3/VP), and Soloist (few unclassified factors)). Subsequently, the phylogenetic relationship, intron–exon structure, and chromosomal location of *AhAP2/ERF* were further characterized. All of these *AhAP2/ERF* genes were distributed unevenly across the 20 chromosomes, and 14 tandem and 85 segmental duplicated gene pairs were identified which originated from ancient duplication events. Gene evolution analysis showed that *A. hypogaea cv.* Tifrunner were separated 64.07 and 66.44 Mya from *Medicago truncatula* L. and *Glycine max* L., respectively. Promoter analysis discovered many *cis*-acting elements related to light, hormones, tissues, and stress responsiveness process. The protein interaction network predicted the exitance of functional interaction among families or subgroups. Expression profiles showed that genes from *AP2*, *ERF*, and *dehydration-responsive-element-binding* subfamilies were significantly upregulated under drought stress conditions. Our study laid a foundation and provided a panel of candidate AP2/ERF TFs for further functional validation to uplift breeding programs of drought-resistant peanut cultivars.

Keywords: AP2/ERF1, peanut, phylogenetic analysis, evolution analysis, drought stress

INTRODUCTION

Transcription factors (TFs) (or *trans*-acting factors) are the main class of regulatory proteins that can specifically combine with DNA-binding domains and perform a key role by regulating the expression of downstream genes (Singh et al., 2002; Licausi et al., 2013). Nearly 60 different TF families have been found in higher plants, such as AP2/ERF (Gutterson and Reuber, 2004; Xu et al., 2011; Li M.-Y. et al., 2015), ARF (Finet et al., 2010; Rademacher et al., 2011), bHLH (Li et al., 2006), bZIP (Ulm et al., 2004; Liu et al., 2012), C2H2 (Tsutsui et al., 2011), MADS (Trevaskis et al., 2003; Terol et al., 2019), MYB (Dubos et al., 2010; Feller et al., 2011), NAC (Mao et al., 2012; Nakashima et al., 2012), SBP (Kandori et al., 2006; Ferreira et al., 2014), and WRKY (Rushton et al., 2010). Among these TFs, the APETALA2/ethylene-responsive element-binding factor (AP2/ERF) superfamily contains the largest group of TFs in plant, which are reportedly involved in plant growth progress and abiotic stress responsiveness according to relevant reports (Licausi et al., 2013; Feng et al., 2020). The first AP2/ERF TF was found to regulate flower development in *Arabidopsis* (Jofuku et al., 1994). Subsequently, AP2/ERF genes were widely found in leaf, root, seed, fruit, and other tissues (Chuck et al., 2002; Hirota et al., 2007; El-Sharkawy et al., 2009; Pietsch et al., 2009; Kitomi et al., 2011; Soares et al., 2011). Not only in plants, related AP2/ERF superfamily proteins are also found in ciliates and protists that may be associated with the His- and Asn-rich HNH class of homing endonucleases (Magnani et al., 2004; Wuitschick et al., 2004).

AP2/ERF TFs usually contain one or two AP2-conserved domains (60–70 amino acid residues) which combine with the *cis*-acting elements in the promoter regions of targeted genes (Okamuro et al., 1997; Allen et al., 1998; Riechmann and Meyerowitz, 1998). The *AP2/ERF* superfamily genes are mainly divided into AP2 (APETALA2), DREB (dehydration-responsive-element-binding), ERF (ethylene-responsive-element-binding), RAV (related to ABI3/VP), and Soloist (few unclassified factors) subfamilies based on the sequence characteristics and the number of AP2-conserved domains (Nakano et al., 2006; Tamura et al., 2011). In most cases, the AP2 subfamily contains proteins with two AP2 domains involved in regulating plant developmental processes (El et al., 2010). ERF, DREB, and RAV subfamily members contain only one single AP2 domain, while RAV members are often associated with an additional B3 DNA-binding domain (Licausi et al., 2010). Discrepancy of 14th and 19th amino acid sequences is the main differences between ERF and DREB subfamilies; the ERF subfamily consists of alanine (Ala) and aspartate (Asp) whereas the DREB subfamily consists of valine (Val) and glutamic acid (Elu) of 14th and 19th amino acid sequences, respectively (Sakuma et al., 2002). Additionally, other members with special gene structure and AP2-like domain are known as Soloist ones (Li H. et al., 2017).

With more draft genomic information of plants released, AP2/ERF superfamily members have been identified and characterized in eudicots, i.e., *Arabidopsis* (Sakuma et al., 2002; Nakano et al., 2006), grapevine (Licausi et al., 2010), cucumber (Hu and Liu, 2011), Chinese plum (Du et al., 2013), apple (Girardi et al., 2013), sweet orange (Ito et al., 2014), pineapple (Huang et al., 2020), canola (Ghorbani et al., 2020), Chinese cherry (Zhu et al., 2021), and dark jute (Kabir et al., 2021), and in monocots, i.e., rice (Sharoni et al., 2011), common wheat (Zhuang et al., 2011), sugarcane (Li et al., 2020), maize (Liu et al., 2013), barley (Guo et al., 2016), and foxtail millet (Lata et al., 2014). In general, AP2 TFs have been found to regulate various developmental processes, such as the development of floral organs (Irish and Sussex, 1990; Jofuku et al., 1994; Chuck et al., 1998 and, 2008; Maes et al., 2001; Aukerman and Sakai, 2003) and embryo and seed growth (Boutilier et al., 2002; Jofuku et al., 2005; Krizek and Beth, 2009). ERF and DREB subfamily proteins mainly function in the resistance to diverse biological and environmental stresses, such as biotic stresses (microbial pathogens and herbivorous insects) and abiotic stresses (drought, heat, cold, and salinity) (Feng et al., 2020). Additionally, RAV subfamily proteins play a crucial role against biotic and abiotic stress responses (Sohn et al., 2006; Li et al., 2011; Fu et al., 2014) by responding to the signal of plant hormones (ethylene and brassinosteroid) (Alonso et al., 2003; Hu et al., 2004).

Peanut, an important oil and economic crop worldwide, is used to provide oil and proteins for humans (Zhuang et al., 2019). In particular, with the characteristics of underground fruit, its yield is being devastatingly affected by drought stress (Vahdati and Lotfi, 2013). Additionally, drought will also result in the increase of aflatoxin contamination and the frequency of diseases and pests (Boyer, 1982; Nimitr et al., 2003). To date, drought has been the most serious abiotic stress which negatively affects the quality and distribution of peanut (Reddy et al., 2003; Farombi, 2006; Cardwell and Henry, 2008; Sun et al., 2013; Jayaprakash et al., 2019; Soni et al., 2020). Previous studies have demonstrated the critical role of *AP2/ERF* genes in mediating drought stress resistance. For example, AhERF3 and AhERF5 in root were upregulated by PEG treatment (Chen et al., 2012), indicating its association with drought stress. Notably, overexpression of *AhERF019* could enhance tolerance to drought in transgenic *Arabidopsis* (Wan et al., 2014). Thus, there may be more important *AP2/ERF* members that act in enhancing the resistance to drought stress of peanut. In recent years, with the release of genome information with cultivated peanut (*A. hypogaea cv.* Tifrunner), an allotetraploid (AABB 2n = 4x = 40), and related wild-type ones (diploids: *A. duranensis* and *A. ipaensis*) (Bertioli et al., 2016; Zhuang et al., 2019). Several gene families, including monosaccharide transporter MST genes (Wan et al., 2020), GRF (Zhao et al., 2019), bHLH (Chao et al., 2017), and WRKY (Song et al., 2016), have been characterized at a genome scale. However, very limited information of *AP2/ERF* genes is available in cultivated peanut.

In the present study, 185 AP2/ERF superfamily members of *A. hypogaea cv.* Tifrunner were investigated by phylogenetic relationship, sequence structures, chromosomal distributions, duplication events, and promoter region analysis. The expression patterns of *AhAP2/ERF* under drought stress were quantified by using quantitative real-time polymerase chain reaction (qRT-PCR). Our investigation will be beneficial to identify the drought-responsive candidate genes for further

functional characterization to breed drought-resistant peanut cultivars.

MATERIALS AND METHODS

Screening and Identification of the *AP2/ERF* Superfamily Genes

To accurately collect all members of *AhAP2/ERF* genes and avoid nonspecific amplification, multiple-database searches were performed. The *A. hypogaea cv.* Tifrunner genome sequences were downloaded from peanutbase (https://www.peanutbase.org/). The AP2 domain (PF00847) profile was obtained from the pfam database (http://pfam.janelia.org), which was used to match each member of AP2/ERF protein in genomes using HMMER 3.1 software (E-value<$1e^{-5}$) (Finn et al., 2011). To avoid the omission of AP2/ERF members, we also performed searches in the Transcription Factor database (http://planttfdb.cbi.pku.edu.cn/) (Jin et al., 2014). All protein sequences acquired were then verified for the AP2 domain by using the SMART (Simple Modular Architecture Research Tool: http://smart.embl-heidelberg.de/) (Letunic et al., 2012) and Pfam (http://pfam.xfam.org) databases (Finn et al., 2008) and NCBI (https://www.ncbi.nlm.nih.gov/cdd). Proteins lacking compete AP2 domains were identified by manual examination. Physichochemical profiling of *AP2/ERF* genes was performed by using online ExPASy (Gasteiger, 2005; Panu et al., 2012). The subcellular localization analysis of curated *AP2/ERF* superfamily genes was conducted on the Plant-Ploc server (http://www.csbio.sjtu.edu.cn/bioinf/plant/) (Chou and Shen, 2008).

Phylogenetic Analysis of AP2/ERF Proteins

The AP2 domains were extracted based on results of SMART (Simple Modular Architecture Research Tool: http://smart.embl-heidelberg.de/) (Letunic et al., 2012). Multiple-sequence alignment was executed by DNAMAN and CLUSTAL program (Thompson et al., 1994; Burner and Legendre, 2000). To construct phylogenetic trees, MEGA 7.0 software was used with the neighbor-joining model (1,000 replicates) (Tamura et al., 2013). AP2/ERF family gene names in *A. hypogaea cv.* Tifrunner were given according to the ascending order of location on chromosomes.

Gene Structure and Conserved Motif Analysis

Information of the intron–exon structure was obtained from the reference peanut genome (*A. hypogaea cv.* Tifrunner, https://www.peanutbase.org/). The Multiple Expectation Maximization for Motif Elicitation program (MEME, http://meme-suite.org/tools/meme) was used to identify potential conserved motifs shared by 185 *AhAP2/ERF* genes (Bailey et al., 2009). Basic information extraction and preliminary drawing of sequence structure were conducted using TBtools (South China Agricultural University, Guangdong, China) (Chen et al., 2020).

Chromosome Localization, Duplications, and Evolutionary Analysis of *AhAP2/ERF*

MapChart 2.3 software developed by Wageningen University and Research in Wageningen, Netherlands, is used to locate genes on chromosomes (Voorrips, 2002). Tandem and segmental genes, Ka/Ks values, and circos figures for chromosome locations with AP2/ERF duplication links were completed by TBtools software (South China Agricultural University, Guangdong, China) (Chen et al., 2020). Duplication and divergence time were calculated by the following formula as described by Bertioli et al. (2016):

$$T = Ks/2\lambda \left(\lambda = 8.12 \times 10^{-9}\right)$$

Promoter Analysis of *AhAP2/ERF* Genes

Approximately 1,500-bp upstream sequences of the *AhAP2/ERF* genes were used to get a better knowledge of the potential function of promoter. PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) (Lescot et al., 2002) was used to identify the *cis*-regulatory elements exited in the gene promoters related to stress responses and hormone effects, and these results were visualized by TBtools.

Prediction of the Protein Interaction Network

Prediction of the protein interaction network was conducted on the basis of the STRING database (https://string-db.org/. accessed on January 28, 2021) (Szklarczyk et al., 2019). *Arabidopsis thaliana* L., the well-characterized model plant, was the subject organism (combined score≥ 0.4). PPI networks were constructed by Cytoscape software v 3. 8.0 (Shannon et al., 2003).

Plant Materials and Drought Stress Treatment

Seeds of *Arachis hypogaea* L., 'YUANZA 9102', laboratory homozygous material, were sown in 392 cm^3 (7 cm long, 7 cm wide, and 8 cm high) pots which were filled with a mixture of vermiculite and perlite (3:1 v/v). Plants were put in a fully controlled growth room (relative humidity: 70%, 16 h/8 h light/dark; 30°C/28°C day/night; light intensity 17,000 lx). Watering was stopped in one part of pots (drought treatment) when seedlings were 5 weeks old, whereas the watering regime remained unchanged in the control plants (every 4 days). Roots were collected from the control and treatment groups every 4 days from seedlings aged 5–9 weeks. All samples were frozen in liquid nitrogen immediately and then stored at -80°C for RNA isolation.

RNA Isolation and Quantitative RT-PCR Assays

RNAprep Pure Plant Plus Kit (Tiangen Biotech, Co., Beijing, China) was utilized to extract RNA from control and treated peanut samples. The cDNA was prepared by following the

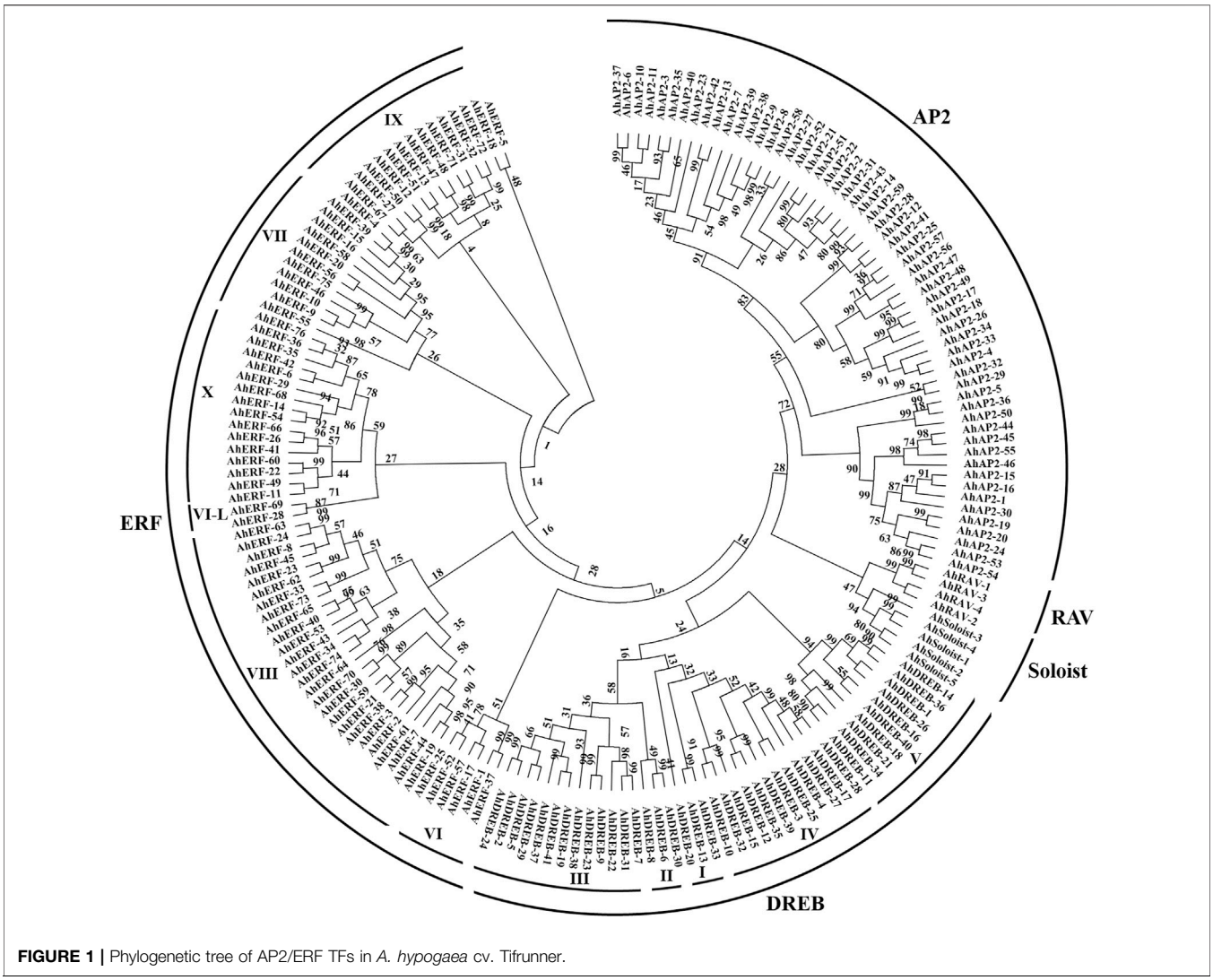


FIGURE 1 | Phylogenetic tree of AP2/ERF TFs in *A. hypogaea* cv. Tifrunner.

user manual of the PrimeScript™ RT Kit with gDNA eraser (perfect real-time, Takara Biomedical Technology, Ltd., Beijing, China). Thirty-five genes from each family were designed by Primer Premier 5.0 and are shown in **Supplementary Table S1**. The alcohol dehydrogenase class III (*AhADH3*, *Arahy. VYWU26.2*) (forward primer: 5′-GAC GCTTGGCGAGATCAACA-3′, reverse primer: 5′-AACCGG ACAACCACCACATG-3′) was selected as the internal reference control (Brand and Hovav, 2010). Subsequently, quantitative RT-PCR was performed using the ABI QS5 qRT-PCR detection system (ABI, United States) and SYBR Green Kit (Tiangen, Beijing, China). An ABI QS5 real-time PCR system was used under the following procedure: 95°C for 15 min, followed by 40 cycles of 95°C for 10 s, and 60°C for 32 s in a 20 μl volume. Each PCR assay was carried out in three biological replicates, of which each replicate corresponded to three technical repeats. Relative expression levels of the genes were calculated using the $2^{-\triangle\triangle Ct}$ method (Livak and Schmittgen, 2001).

RESULTS

Characterization and Phylogenetic Analysis of AhAP2/ERF Family Proteins in Cultivated Peanut

A total of 185 unigenes with the AP2 domain were characterized as AP2/ERF TFs in *A. hypogaea cv.* Tifrunner (**Supplementary Table S2**). Depending on the sequence characteristics, the AP2 domains, phylogenetic tree analysis, and the classification system established by the group of Yamaguchi-Shinozaki (Tamura et al., 2011) and Shinshi (Nakano et al., 2006), the *AP2/ERF* superfamily genes are mainly classified into AP2 (APETALA2), ERF (VI-X, ethylene-responsive-element-binding), DREB (I-V, dehydration-responsive-element-binding), RAV (related to ABI3/VP), and Soloist (few unclassified factors) subfamilies (**Figure 1** and **Supplementary Figure S1**). Among these, the 27 to encode two AP2 domains and the 4 to one AP2 domain together with one B3 domain were thus

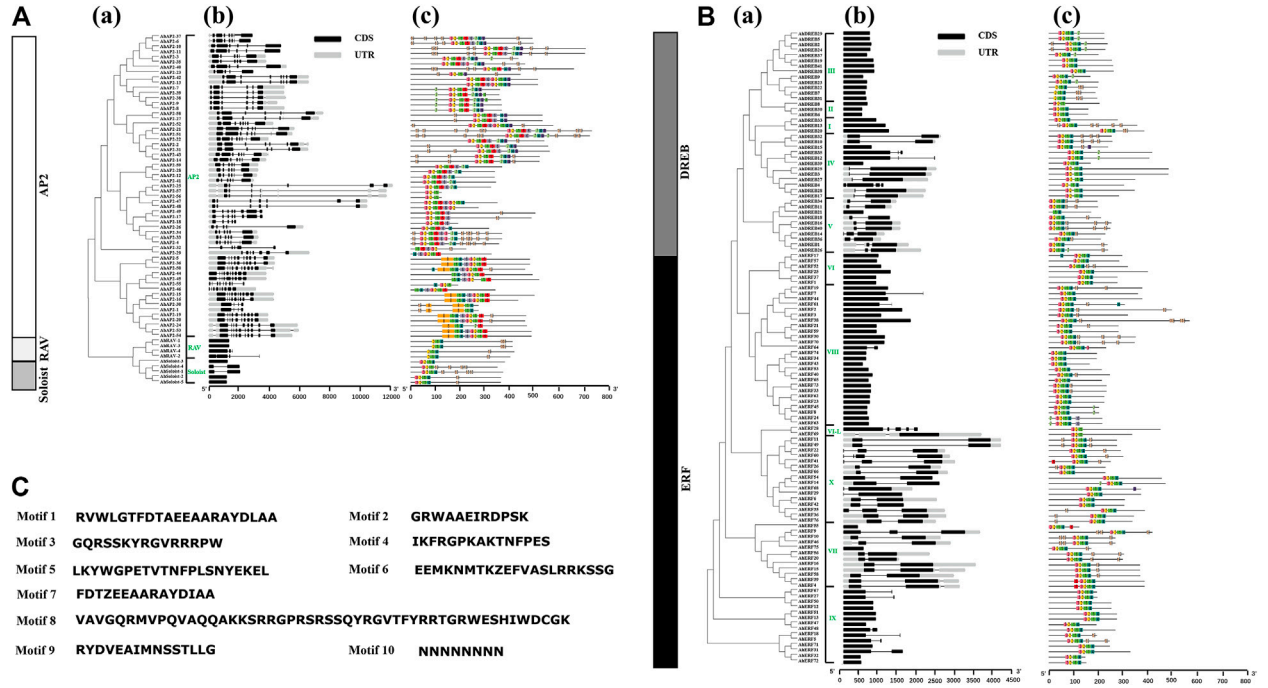


FIGURE 2 | Intron–exon structures and conserved motif analysis of *AhAP2/ERF* genes according to the phylogenetic relationship. **(A)** Related information of AP2, RAV, and Soloist subfamily. **(B)** Related information of ERF and DREB subfamily. **(C)** The amino acid composition of each motif. **(a)** The phylogenetic tree. **(b)** The exon–intron structure of *AP2/ERF* genes. **(c)** The distribution of conserved motifs in AP2/ERF proteins. Each conserved motif is represented by various-colored rectangles. Box length corresponds to motif length. Color blocks of different colors represent different family and group members.

assigned to the AP2 and RAV families, respectively. Based on the similarity of amino acid sequences with the AP2 domain, the 117 genes were further assigned to the ERF 76) and DREB 41) subfamilies, respectively. Thirty-two members with a single AP2 domain but were distinct from the ERF or DREB subfamily were classified into the AP2 subfamily (**Supplementary Table S2**). The remaining five genes with independent clades from others were identified as *Soloist* genes. Subsequently, all superfamily members were named according to the order on the chromosomes of each family member to distinguish from each other for the study (**Supplementary Table S3**).

To evaluate the phylogenic relationship and classification of the ERF and DREB subfamily, multiple-alignment analyses was performed on the protein sequences with the AP2 domain acquired from peanut (117), *Arabidopsis* (139), and rice (139), as suggested by Nakano et al. (2006) (**Supplementary Figures S2, S3**). The NJ phylogenetic tree divided the ERF and DREB subfamilies of peanut and Arabidopsis into 10 subgroups (DREB-I-V and ERF-VI to X) following the classification as described by Nakano et al. (2006) (**Supplementary Figure S4**). The phylogenetic tree of ERF and DREB subfamily proteins of peanut and rice also exhibited similar results (**Supplementary Figure S5**). Overall, current findings of the phylogenetic tree demonstrated that classification of the peanut ERF and DREB subfamily proteins is similar to the *Arabidopsis* and rice ERF family (**Supplementary Table S4**).

Molecular property analysis showed that MW of *AP2/ERF* superfamily members varied from 14.46 to 82.54 kDa. Most of *AP2/ERF* superfamily genes showed their localization in the nucleus. Moreover, the negative GRAVITY values suggested the globular hydrophilic nature of the AP2/ERF proteins. Interestingly, the members of the same families or clades shared similar physical properties, indicating the functions conservatively in the same clades and differentially among subfamilies.

Structure Analysis of *AhAP2/ERF*

Structural analysis of *AP2/ERF* genes is helpful for us to fully understand the conservative characteristics of peanut AP2/ERF protein and analyze its evolutionary differences. Numbers of introns varied among AP2/ERF subfamilies (**Figure 2**). All the *AP2* family genes contain 3–10 introns, whereas most members of *ERF*, *RAV*, and *Soloist* subfamilies have only 1–2 introns or do not possess introns. A similar phenomenon was discovered in *Arabidopsis thaliana* L. and *Cucumis sativus* L., where most *Arabidopsis* genes of the ERF family do not possess introns (Sakuma et al., 2002), and 83% of *CsERF* genes do not have intron (Ito et al., 2014). What is more, members of the ERF family clustered into one branch have a similar gene structure (**Figure 2**). There was some subgroup specificity: members of groups DREB-III, DREB-II, DREB-I, and ERF-VI did not possess introns, and the genes owning two introns were in the groups ERF-VII and ERF-X, possibly attributed to number changes of introns during evolution, whereas the number and position of the introns were relatively conserved in the same group of plant species.

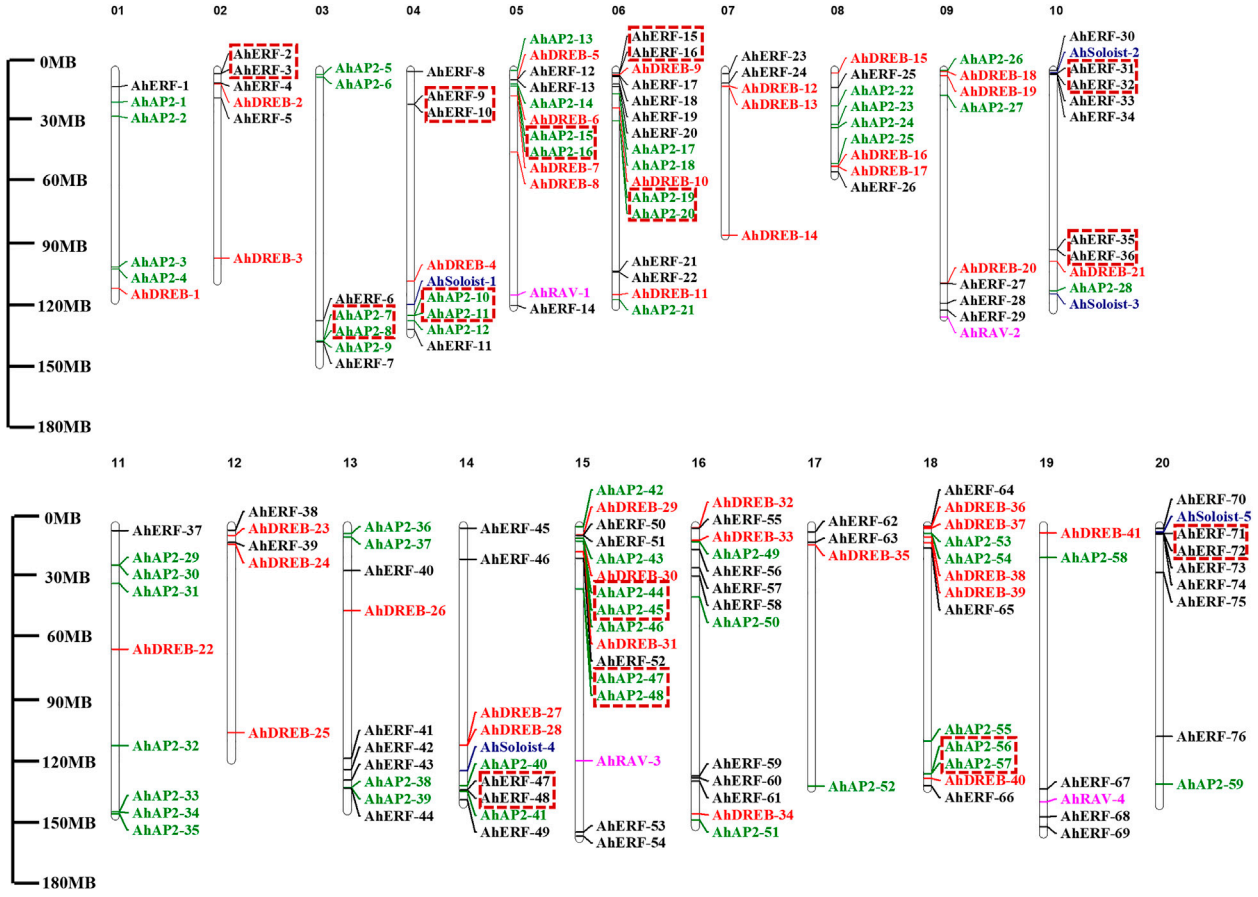


FIGURE 3 | Chromosome mapping and duplication of *AhAP2/ERF* genes. On the right of the chromosome is the gene name. Scale represents a 30-Mb chromosomal distance. Genes in tandem repeats are shown in the red box.

To provide evidence for the classification and the functional conversation of *AP2/ERF* superfamily genes among different groups, 10 conserved motifs (motifs 1–10) were analyzed by using MEME software (**Figure 2**). Among these, motifs 1 and 2 were dominantly present in the AP2 domain regions of all family members. The proteins of the same group showed identical numbers and arrangements of motifs, which are different among the various clades. For example, nine motifs were detected in AP2 and ERF families, and four motifs in RAV and Soloist families. Furthermore, the number and arrangement of motifs in the RAV (motifs 2, 1, 4, 10), ERF, and Soloist (motifs 3, 2, 1, 4) families showed high similarity. Motif 8 was only detected in few AP2 family members, signifying the meticulousness of the motif in the AP2 family. Remarkably, motifs in the same group showed great similarity, indicating the functional conservation in different groups. Comparing the intron–exon structure and conserved motif analysis, it is clear that the members of the same group showed great similarity of characteristics, indicating that most of *AhAP2/ERF* genes were highly conserved among groups.

Genome Distribution of *AP2/ERF* Genes

AP2/ERF genes showed random distribution on the 20 chromosomes of the peanut genome (**Figure 3** and **Supplementary Table S3**). Maximum numbers of *AhAP2/ERF* genes (16 genes) were located on Chromosomes 6 and 15, while Chr17 had the least number of genes (4 genes). Other chromosomes had a random number of allocated *AhAP2/ERF* genes (5–13 genes). Interestingly, most members are distributed at both ends of the chromosome. A similar study on *Arabidopsis* and other species showed consistent findings (Nakano et al., 2006; Guo et al., 2016). This location similarity of genes on chromosomes indicates functional consistency.

Duplication Events of *AP2/ERF* Genes and Synteny Analysis

Gene duplication events (segmental or tandem) play a significant role in the expansion and evolution of gene families in plant species (Baloglu, 2014). In total, 99 duplicated gene pairs, which were also named as homoeologous genes, were identified: 14 tandem and 85 segmental duplications (**Figure 4**). Segmental gene duplication mainly occurred in the *A. hypogaea cv.* Tifrunner genome rather than tandem duplication event (**Figure 4** and **Supplementary Table S5**). No duplication events occurred in group I of the ERF family, whereas more segmental duplication events occurred in other groups of the ERF family, implying that the biggest members of the families might

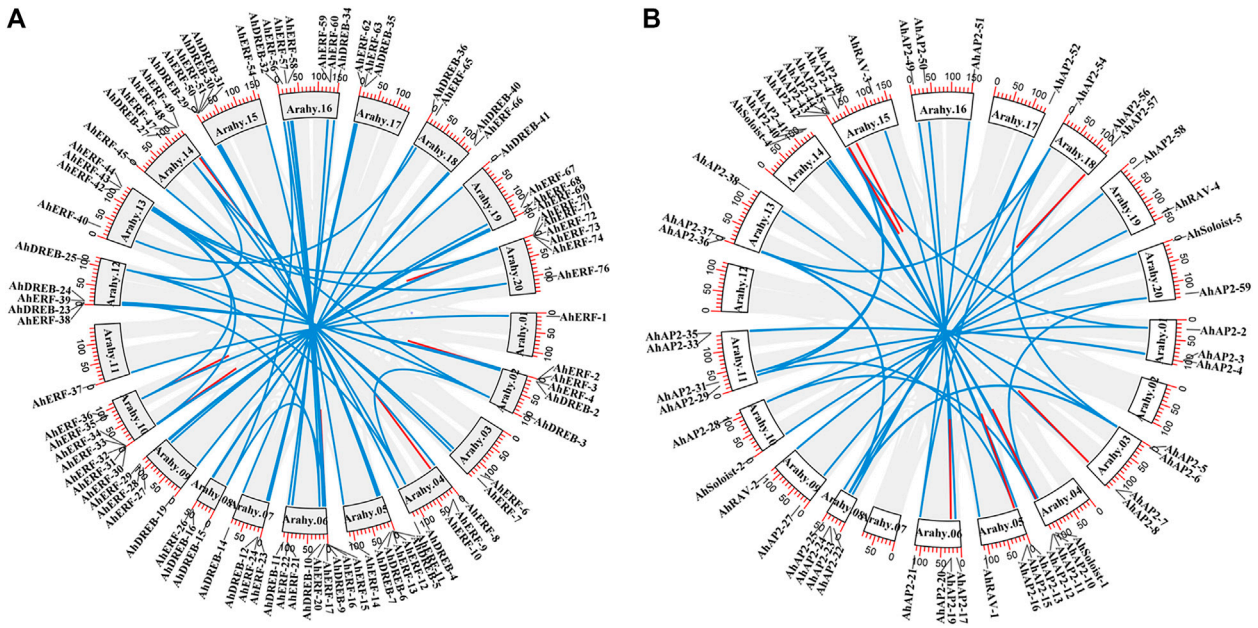


FIGURE 4 | Circos figures for chromosome locations with AP2/ERF duplication links. **(A)** ERF and DREB subfamily duplication links. **(B)** AP2, RAV, and Soloist subfamily duplication links. Blue and red lines indicated segmented and tandem duplicated gene pairs, respectively.

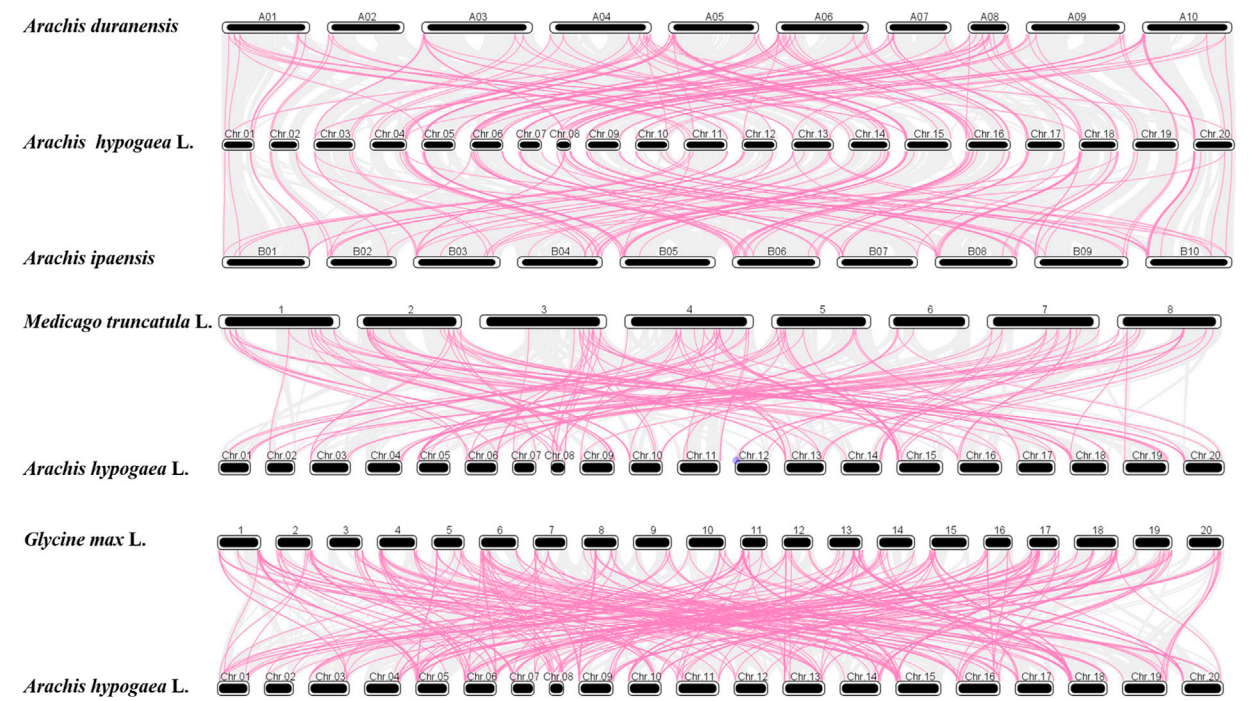


FIGURE 5 | Synteny of AP2/ERF genes in the different genomes of *A. duranensis*, *A. ipaensis*, *Medicago truncatula* L., *Glycine* max L., and *A. hypogaea* cv. Tifrunner.

arise from a higher frequency of segmental duplication, when adapting to various environmental shifts. In contrast, tandem duplication has a confined benefaction to the gene family expansion as compared to the segmental duplication. Similar studies on *Arabidopsis*, rice, sorghum (Wang et al., 2011), common bean (Kavas et al., 2015; 2016), and cucumber

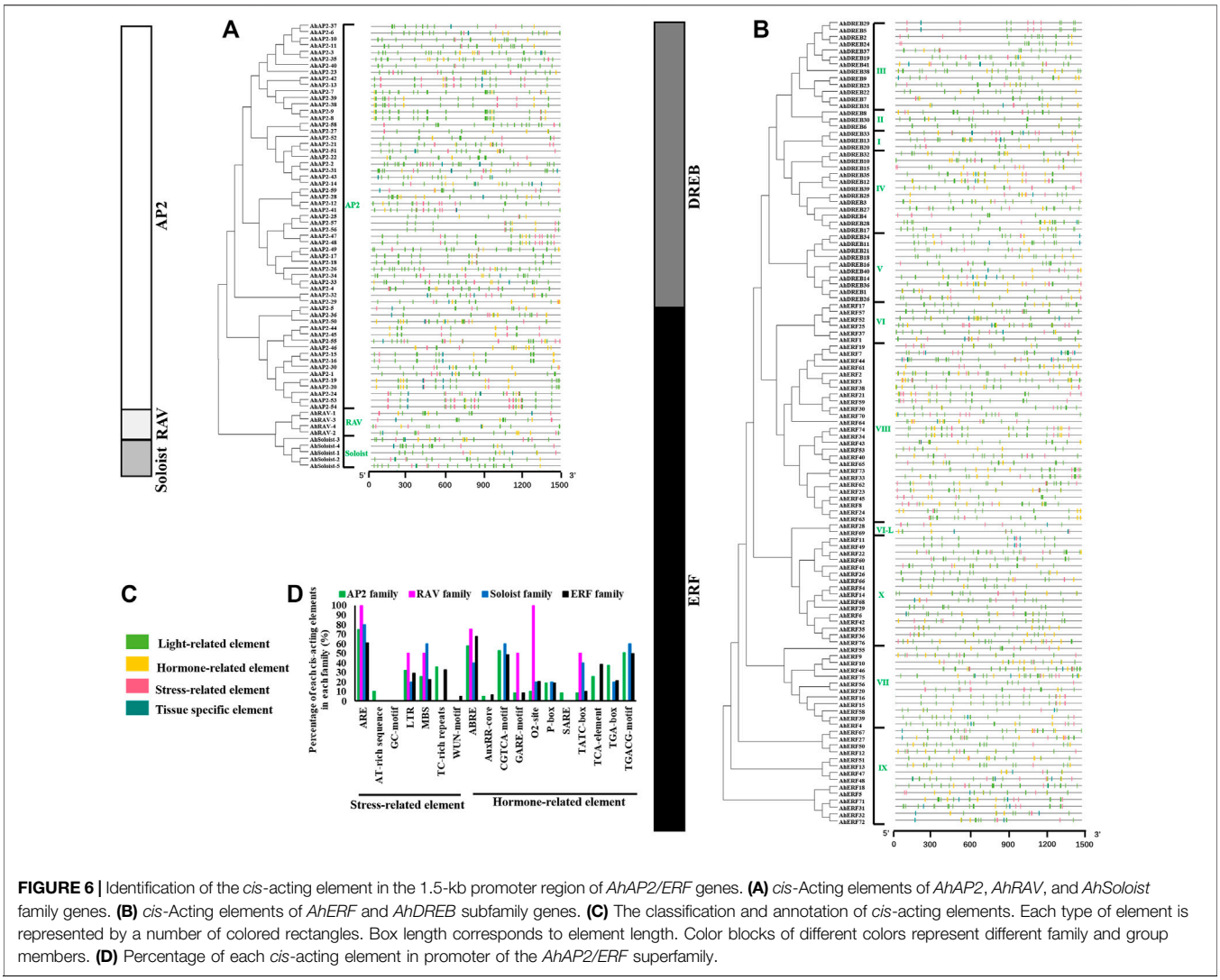


FIGURE 6 | Identification of the *cis*-acting element in the 1.5-kb promoter region of *AhAP2/ERF* genes. **(A)** *cis*-Acting elements of *AhAP2*, *AhRAV*, and *AhSoloist* family genes. **(B)** *cis*-Acting elements of *AhERF* and *AhDREB* subfamily genes. **(C)** The classification and annotation of *cis*-acting elements. Each type of element is represented by a number of colored rectangles. Box length corresponds to element length. Color blocks of different colors represent different family and group members. **(D)** Percentage of each *cis*-acting element in promoter of the *AhAP2/ERF* superfamily.

(Baloglu, 2014) showed compatible findings. In general, segmental duplication might be the main driving force for the *AP2/ERF* gene family expansion in peanut genome.

Moreover, comparative orthologous analysis was conducted among *A. hypogaea cv.* Tifrunner, *A. duranensis*, *A. ipaensis*, *Medicago truncatula* L., and *Glycine max* L. to characterize the evolutionary patterns of *AhAP2/ERF* genes with *Leguminosae* species (**Figure 5**). In total, 140, 133, 314, and 145 orthologous gene pairs were found with *A. duranensis*, *A. ipaensis*, *Medicago truncatula* L., and *Glycine max* L., respectively (**Supplementary Tables S6–S9**). The Ka/Ks for segmental duplication was 0.02–0.92 with an average of 0.28, while the ratio of tandem duplication ranged from 0.29 to 0.68 with an average of 0.42 (**Supplementary Table S5**). These segmental and tandem duplications may occur in ~3.82–43.68 Mya, respectively. In addition, the Ka/Ks ratio of ortholog gene pairs between *A. hypogaea cv.* Tifrunner and *Medicago truncatula* L. (0.25) and *A. hypogaea cv.* Tifrunner and *Glycine max* L. (0.24) were strongly subjected to pure selection (Lynch and Conery, 2000). The divergence times were 64.07 and 66.44 Mya for *Medicago truncatula* L. and *Glycine max* L., respectively.

Prediction Analysis of *cis*-Acting Elements With *AhAP2/ERF*

Specific *cis*-element motifs can be recognized by TFs and participate in gene expression regulation. In order to further study the potential regulatory mechanism of the *AP2/ERF* gene in diversified biological processes, especially in plant drought stress response, the 1.5 kb upstream sequence of the *AP2/ERF* gene translation start site was submitted to the PLANTCARE database to detect cis elements (Cui et al., 2018; Meng et al., 2020). A total of 56 known *cis*-elements (30 light-related elements, 11 hormone-related elements, 8 tissue-specific elements, and 7 stress-related elements) were detected (**Figure 6** and **Supplementary Table S10**). ABRE, AuxRR core, CGTCA motifs, GARE motifs, O2 site, P-box, TATC-box, TCA element, TGA element, and TGACG motifs involved in hormonal responses are found in 63.8, 5.9,

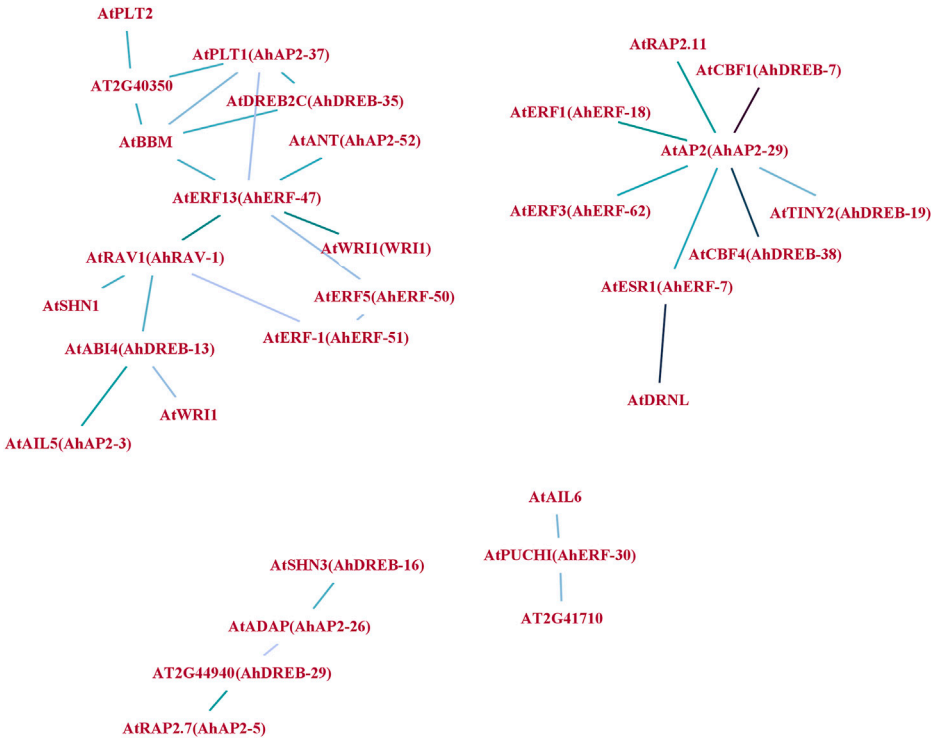


FIGURE 7 | Protein–protein interaction (PPI) network analysis of *AhAP2/ERF* TF proteins. Specific protein interactions between AP2/ERF transcription factors in peanut were determined using String (Szklarczyk et al., 2019). Strong associations are represented by thicker lines.

49.2, 9.2%,18.9, 18.4, 10.8, 32.4, 25.9, and 49.2% of *AhAP2/ERF* promoters, respectively. Meanwhile, there are a large number of stress-related elements, including MBS (drought inducibility), TC-rich repeats (defense and stress responsiveness), WUN motif (wound responsiveness), LTR (low-temperature responsiveness), ARE (essential for the anaerobic induction), and GC motif (anoxic specific inducibility) (**Supplementary Table S10**). Moreover, there was a divergence in the percentage of *cis*-acting elements in promoter regions of various families (**Figure 6D**). For example, all the RAV family members contained ARE and O2-site elements, whereas 74.6, 80.0, and 60.7% family members of AP2, Soloist, and ERF families possessed ARE elements, and 10.2, 20, and 20.5% of those families exhibited O2-site elements in the promoter region, respectively. Notably, MBS, an important *cis*-element related to the plant drought-inducibility process, was detected in the promoters of 25.4% of AP2, 50% of RAV, 60% of Soloist, and 22.2% of ERF family members. As a major hormone in plant response to drought stress, ABRE possessed 57.6% of AP2, 75% of RAV, 40% of Soloist, and 67.5% of ERF family members. TC-rich repeats, a *cis*-acting element involved in defense and stress responsiveness, were discovered only in the promoters of AP2 and ERF family members. WUN motif, a wound-responsive element, was only detected in the ERF family. The variants in the characterization of *cis*-acting elements implied the functional discrepancy in different families.

What is more, certain *ERF* members in groups III (*AhDREB5/29/41*), IV (*AhDREB15/27/39*), V (*AhDREB21/26*), VI (*AhERF25/52*), VII (*AhERF56*), VIII (*AhERF3/19/30/34/43/53/62/64/65/70/74*), and IX (*AhERF5/31/71/72*) possess a relatively large number of MBS elements in the promoter regions, implying the members' main role in the regulation of plant drought responsiveness and the function variations among groups (d). In other words, *cis*-acting elements in the same group showed great similarity, indicating the functional conservation in the same groups or clades (**Figure 6D**).

Interaction Network Analysis of *AhAP2/ERF* Proteins

To understand the synergy among peanut AP2/ERF TFs during their regulatory process, an interaction network was drawn using *Arabidopsis* ortholog genes (**Figure 7**). A sum of 31 gene pairs with a combined score over zero value was deliberated to have the interaction with others (**Supplementary Table S11**). *AhAP2-29*, *AhAP2-37*, *AhERF-47*, and *AhRAV-1* had more than three nodes and protein pairs and were involved in more powerful crossing networks, suggesting their core role in peanut. However, other members were only regulated by a few numbers of genes, indicating its less important role in transcriptional level regulation (**Figure 7**). Interestingly, proteins from different families showed complex interaction. For example, core members *AP2-39* or *RAV-3* interact with genes from *DREB* and *ERF* subfamilies, implying the exitance of functional

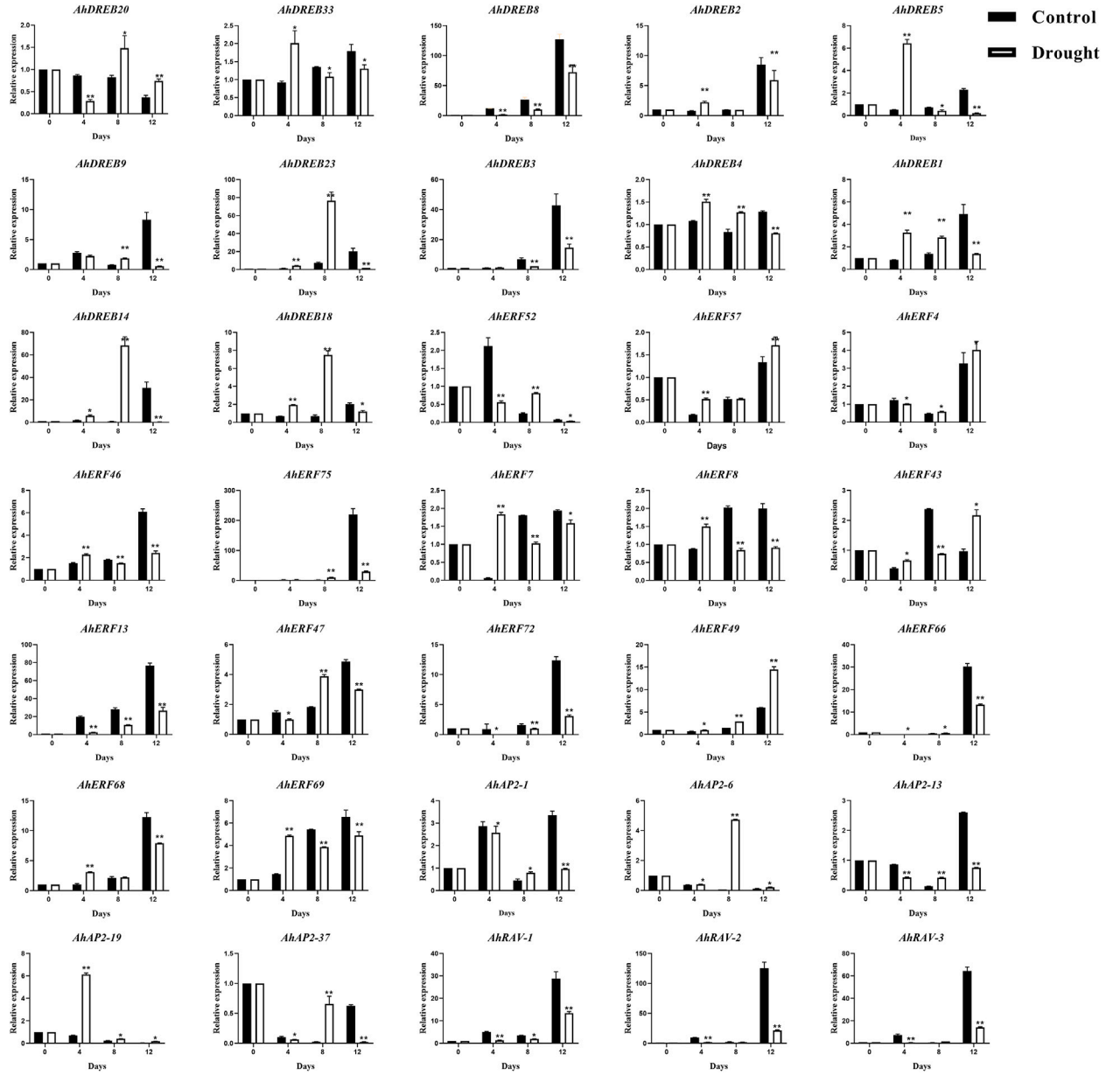


FIGURE 8 | qRT-PCR expression analysis of 35 selected *AhAP2/ERF* genes under drought stress conditions. The expression levels of the untreated (0 h) group were normalized to 1 as a control. Error bars were obtained from three biological replicates. Values are means ± standard errors (SEs) of three independent biological replicates (n = 3). Asterisks indicate a significant difference between the control group and treatment group at each time point as determined by Student's t-test (*p < 0.05; **p < 0.01).

interaction among subfamilies. These interrelationships will provide a reference for studying the regulatory functions of *AhAP2/ERF* genes in peanut.

Expression Profiles of *AhAP2/ERF* Genes Under Drought Stress Using qRT-PCR

For better knowing the possible regulatory roles of *AP2/ERF* family genes in peanut response to drought stress, 35 representative genes from AP2 (5), RAV (3), and ERF (27, members from each group) were selected to verify whether their expression levels would be induced under drought stress conditions by qRT-PCR (**Figure 8**), especially for the ones that possess MBS elements (drought-inducibility), ABRE, TC-rich elements, ARE, and WUN-motif elements in the promoter regions.

The treatment group with no watering was used to stimulate the expression of plant defense genes. The expression analysis of *AhAP2/ERF* genes responsive to drought stress could present useful information to identify their implied role as candidate

genes to mitigate drought stress severity. As shown in **Figure 8**, under the no watering condition, except *AhDREB-3*, *AhDREB8*, *AhERF-4*, *AhERF-13*, *AhERF-66*, *AhERF-72*, *AhERF-75*, *AhAP2-1*, *AhAP2-6*, *AhAP2-13*, *AhRAV-1*, *AhRAV-2*, and *AhRAV-3* exhibited downregulation, whereas the remaining genes were significantly upregulated and subsequently downregulated, at all time points compared with those at 0 h. Furthermore, the peak expression of *AhDREB-1*, *AhDREB-4*, *AhDREB-5*, *AhDREB-33*, *AhERF-7*, *AhERF-8*, *AhERF-69*, and *AhAP2-19* was discovered on the 4th day, but that of *AhDREB-14*, *AhDREB-18*, *AhDREB-20*, *AhDREB-23*, *AhERF-47*, *AhAP2-6*, and *AhAP2-37* was detected at the 8th day. Interestingly, the members in the same group showed semblable expression trends, which indicates the function consistency. Notably, collinear genes *AhDREB-9* and *AhDREB-23*, *AhAP-26* and *AhAP2-37*, *AhRAV-1*, and *AhRAV-3* showed highly similar expression patterns under drought stress treatment, indicating that their biological functions also have a certain similarity.

DISCUSSION

In plants, AP2/ERF TFs play diverse roles in multiple growth processes and work against environmental factors through transcriptional regulation (Sakuma et al., 2002; Nakano et al., 2006; Woo et al., 2010; Li A. et al., 2017; Yao et al., 2017). Therefore, studying the biological functions and molecular mechanisms of these proteins will facilitate obtaining a deeper understanding of the pathways adapting to environmental pressures during plant growth.

In this study, 185 *AP2/ERF* genes with at least one AP2 conserved domain were identified in the *A. hypogaea cv.* Tifrunner genome through genome-wide analysis. Similar to other plants, all putative *AP2/ERF* superfamily genes were identified as five subfamilies: AP2, ERF, DREB, RAV, and Soloist (**Supplementary Table S2**). Each subfamily has 59, 76, 41, 4, and 5 members, respectively. In different plants, the numbers of AP2/ERF proteins vary significantly depending on the genome size (eudicot or monocot) (**Supplementary Table S12**), which may result in gene evolution and duplication. On the other hand, the number of each subfamily follows the regular pattern: the number of the ERF subfamily is the largest, followed by DREB, AP2, and RAV or Soloist (**Supplementary Table S12**), suggesting that the composition of the AP2/ERF superfamily TFs is highly conserved in plants and may share a common ancestor before separation. Moreover, the largest number of ERF and DREB subfamilies strongly implies its main role in plant growth and development process. The differences in the values of molecular weight (14.46–82.54 kDa) and pI (4.55–10.58) of *AhAP2/ERF* suggest the putative differences in *AhAP2/ERF* (Lata et al., 2014). Subcellular localization predicted that *AhAP2/ERF* TFs are mainly localized to the nucleus and thus validates the posttranscriptional regulatory mechanism of the proteins (Karniely and Pines, 2005). Further analysis showed that the AP2/ERFs exhibit certain subfamily characteristics in intron/exon patterns, motif structures, and phylogenetic relationships (**Figures 1**, **2**). This high evolutionary conservation can be used as an important basis for subfamily classification.

Chromosomal mapping showed an uneven distribution of *AhAP2/ERF* genes on 20 chromosomes (**Figure 3**). There were hot regions or gene clusters on chr06 and chr15. Generally, tandem, large-scale chromosome segmental duplication and transposition were identified as the main evolutionary mechanisms that cause the expansion of the gene family (Lynch and Conery, 2000; Cannon et al., 2004). In total, 14 gene pairs showed tandem duplication and 85 gene pairs revealed segmental duplication, which sustain the overall 8.6% tandem duplication of AP2/ERF in *A. hypogaea cv.* Tifrunner (Zhuang et al., 2019) (**Figure 4**). The numbers of duplicated gene pairs vary between crops, such as the number of duplication pairs which is 90 in sunflower, 76 in grape, 51 in *Arabidopsis*, 41 in rice, and 11 in dark jute, all of which were lower than in peanut. Hence, this variation in *AP2/ERF* gene numbers in plants might be due to the different duplication events. The microsyntenic analysis of these *AP2/ERF* gene families across the *Leguminosae* family could provide valuable information about their evolution. Our findings demonstrated that a strong association between *AhAP2/ERF* genes of cultivated peanut and wild species was observed (**Figure 5**). Among them, there were an equal number of pairs of syntenic relationships in the genome of *A. hypogaea cv.* Tifrunner with *A. duranensis* and *A. ipaensis* (**Supplementary Tables S6, S7**). Notably, most *AhAP2/ERF* genes of *A. duranensis* and *A. ipaensis* might have more than one ortholog in cultivated peanut. These results suggested that cultivated peanut, an allotetraploid plant, likely contained twice the number of AP2/ERF observed for wild-type peanut. The mean Ka/Ks value of peanut with *Medicago truncatula* L. and *Glycine max* L. suggests a purifying selection of *AhAP2/ERF* genes that have undergone great selective constraint and substitution elimination by natural selection (**Supplementary Tables S8, S9**).

Abiotic stresses do a great harm to the regular growth in peanut at early stages; thus, seedlings at 5 weeks of age were used for drought tolerance (Passioura, 1983; Pierret et al., 2007; Songsri et al., 2008). As an important organ for plants to absorb water and mineral elements, roots directly experience soil drought, and thus, the expression pattern of *AhAP2/ERF* genes under drought stress in roots is essential to clarifying functional divergence (Passioura, 1983; Songsri et al., 2008). The prediction of peanut AP2/ERF protein function by constructing a protein interaction network (Letovsky and Kasif, 2003) proposed the interaction among AP2, ERF, and RAV families, thus implying its interactive function in response to various stresses. For example, *AhAP2-29* showed a strong interaction with *AtRAP2.11*, *AtCBF1*, *AtTINY2*, and *AtESR1*, which are members of the AP2/ERF family that participate in the stress tolerance in *Arabidopsis* (Banno et al., 2001; Novillo et al., 2004; Wei et al., 2005; Min et al., 2012), implying the potential function of AhAP2-29 in peanut response to drought stress. The same phenomenon is discovered in the AP2/ERF family members, including *AhAP2-37*, *AhERF47*, and *AhRAV-1* which are known to function against stress in *Arabidopsis*, thus indicating that there are strong and complex interactions of *AhAP2*, *AhERF*, and *AhRAV* members in peanut response to

drought stress. However, under no-watering condition, the expression of *RAVs* was extremely downregulated, perhaps suggesting the indirect role in peanut against drought stress, which is in-line with the previous studies in other crops (Hu et al., 2004; Sohn et al., 2006; Je et al., 2010; Li X.-J. et al., 2015; Wei et al., 2018).

The promoter sequence possesses vital information about gene functional components (i.e., *cis*-acting elements) and reflects potential function of the gene (Kabir et al., 2021). In this study, four distinct types of *cis*-acting elements were found, of which hormone-related and stress-related elements are ones having close relationships with plant stress conditions. Moreover, ABA, as a stress signal, is essential during plant growth and development. It integrates various stress signals and controls downstream stress responses to make plants adapt to various stress environments through uninterrupted adjustment (Tuteja, 2007). Promoter analysis showed that almost all AhAP2/ERF members have ABA response elements, especially ones from AP2, ERF, and DREB subfamilies. Other considerable elements which are related to their function in peanut were TC-rich elements, MBS, and other hormone-related and stress-related elements, which interact in a way. For *ERF*, *DREB*, and *AP2* subfamily genes, it appears that, except *AhDREB-3*, *AhERF-13*, *AhERF-66*, and *AhERF-72*, one owns any *cis*-acting elements of ABRE, TC-rich, MBS, or ARE which may be upregulated by the no-watering treatment, implying the main role of *cis*-acting elements of the promoter in peanut response to drought stress. However, *AhRAV* genes were all downregulated during drought regardless if they are in the promoter region. For the complex interaction of *ERF*, *DREB*, and *AP2* (**Figure 7**), *AhRAV* may be strongly regulated by other subfamilies. These results suggest that *AhAP2*, *AhERF*, and *AhDREB* genes may play pivotal roles in response to drought stress.

CONCLUSION

All in all, a total of 185 *AP2/ERF* genes were identified in the *A. hypogaea cv.* Tifrunner genome and divided into AP2 (59), ERF (76), DREB (41), RAV (4), and Soloist 5) subfamilies. Members in the same family or group shared great similarity of exon–intron structure and conserved motifs. Segmental duplication contributed to the expansion of *AhAP2/ERF* genes, and these duplication pair genes had evolved under strong purifying selection. *cis*-Element analysis suggested that the expression of *AhAP2/ERF* can be regulated by hormones and various environmental factors. Protein interaction predicted complex interaction relationships among or within groups of ERF, DREB, and AP2 members. The expression profile under drought stress by qRT-PCR showed that some *AhAP2/ERF* were significantly upregulated, indicating their potential roles in response to drought 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

MC, SH, and MH conceived the study and carried out the bioinformatic analysis. MC, PC, JG, and HL collected the plant materials and performed the experiments. PD, WD, and BH participated in handling figures and tables. MC and ZZ drafted the manuscript. MH, LS, and XZ revised the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (No. 31871663), the China Agriculture Research System (CARS-13), the Major Technology Research and Development of Henan Province, China (201300111000) and the Henan Provincial Agriculture Research System, China (S2012-5).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.750761/full#supplementary-material

REFERENCES

Allen, M. D., Yamasaki, K., Ohme-Takagi, M., Tateno, M., and Suzuki, M. (1998). A Novel Mode of DNA Recognition by a β-Sheet Revealed by the Solution Structure of the GCC-Box Binding Domain in Complex with DNA. *EMBO J.* 17, 5484–5496. doi:10.1093/emboj/17.18.5484

Alonso, J. M., Stepanova, A. N., Solano, R., Wisman, E., Ferrari, S., Ausubel, F. M., et al. (2003). Five Components of the Ethylene-Response Pathway Identified in a Screen for Weak Ethylene-Insensitive Mutants in *Arabidopsis*. *Proc. Natl. Acad. Sci.* 100, 2992–2997. doi:10.1073/pnas.0438070100

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

Aukerman, M. J., and Sakai, H. (2003). Regulation of Flowering Time and floral Organ Identity by a microRNA and its APETALA2-Like Target Genes. *Plant Cell* 15, 2730–2741. doi:10.1105/tpc.016238

Bailey, T. L., Bodén, 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

Baloglu, M. C. (2014). Genome-Wide In Silico Identification and Comparison of Growth Regulating Factor (GRF) Genes in Cucurbitaceae Family. *Plant Omics* 7, 260–270. doi:10.1097/01.tp.0000332556.64365.e7

Banno, H., Ikeda, Y., Niu, Q.-W., and Chua, N.-H. (2001). Overexpression of *Arabidopsis* ESR1 Induces Initiation of Shoot Regeneration. *Plant Cell* 13, 2609–2618. doi:10.2307/387152210.1105/tpc.010234

Bertioli, D. J., Cannon, S. B., Froenicke, L., Huang, G., Farmer, A. D., Cannon, E. K. S., 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

Boutilier, K., Offringa, R., Sharma, V. K., Kieft, H., Ouellet, T., Zhang, L., et al. (2002). Ectopic Expression of BABY BOOM Triggers a Conversion from Vegetative to Embryonic Growth. *Plant Cell* 14, 1737–1749. doi:10.1105/tpc.001941

Boyer, J. S. (1982). Plant Productivity and Environment. *Science* 218, 443–448. doi:10.1126/science.218.4571.443

Brand, Y., and Hovav, R. (2010). Identification of Suitable Internal Control Genes for Quantitative Real-Time PCR Expression Analyses in Peanut (*Arachis hypogaea*). *Peanut Sci.* 37, 12–19. doi:10.3146/PS09-014.1

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

Cardwell, K. F., and Henry, S. H. (2008). Risk of Exposure to and Mitigation of Effect of Aflatoxin on Human Health: A West African Example. *J. Toxicol. Toxin Rev.* 23, 217–247. doi:10.1081/TXR-200027817

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, N., Yang, Q., Su, M., Pan, L., Chi, X., Chen, M., et al. (2012). Cloning of Six ERF Family Transcription Factor Genes from Peanut and Analysis of Their Expression during Abiotic Stress. *Plant Mol. Biol. Rep.* 30, 1415–1425. doi:10.1007/s11105-012-0456-0

Chou, K.-C., and Shen, H.-B. (2008). Cell-PLoc: A Package of Web Servers for Predicting Subcellular Localization of Proteins in Various Organisms. *Nat. Protoc.* 3, 153–162. doi:10.1038/nprot.2007.494

Chuck, G., Meeley, R. B., and Hake, S. (1998). The Control of Maize Spikelet Meristem Fate by the APETALA2-Like Gene Indeterminate Spikelet1. *Genes Develop.* 12, 1145–1154. doi:10.1101/gad.12.8.1145

Chuck, G., Meeley, R., and Hake, S. (2008). Floral Meristem Initiation and Meristem Cell Fate Are Regulated by the maize AP2 Genes Ids1 and Sid1. *Development* 135, 3013–3019. doi:10.1242/dev.024273

Chuck, G., Muszynski, M., Kellogg, E., Hake, S., and Schmidt, R. J. (2002). The Control of Spikelet Meristem Identity by the Branched Silkless1 Gene in Maize. *Science* 298, 1238–1241. doi:10.1126/science.1076920

Cui, M., Wang, C., Zhang, W., Pervaiz, T., Haider, M. S., Tang, W., et al. (2018). Characterization of Vv-miR156: *Vv-SPL* Pairs Involved in the Modulation of Grape Berry Development and Ripening. *Mol. Genet. Genomics* 293, 1333–1354. doi:10.1007/s00438-018-1462-1

Du, D., Hao, R., Cheng, T., Pan, H., Yang, W., Wang, J., et al. (2013). Genome-Wide Analysis of the AP2/ERF Gene Family in *Prunus Mume*. *Plant Mol. Biol. Rep.* 31, 741–750. doi:10.1007/s11105-012-0531-6

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., Lepiniec, L., et al. (2010). MYB Transcription Factors in *Arabidopsis*. *Trends Plant Sci.* 15, 573–581. doi:10.1016/j.tplants.2010.06.005

El Ouakfaoui, S., Schnell, J., Abdeen, A., Colville, A., Labbé, H., Han, S., et al. (2010). Control of Somatic Embryogenesis and Embryo Development by AP2 Transcription Factors. *Plant Mol. Biol.* 74, 313–326. doi:10.1007/s11103-010-9674-8

El-Sharkawy, I., Sherif, S., Mila, I., Bouzayen, M., and Jayasankar, S. (2009). Molecular Characterization of Seven Genes Encoding Ethylene-Responsive Transcriptional Factors during Plum Fruit Development and Ripening. *J. Exp. Bot.* 60, 907–922. doi:10.1093/jxb/ern354

Farombi, E. O. (2006). Aflatoxin Contamination of Foods in Developing Countries: Implications for Hepatocellular Carcinoma and Chemopreventive Strategies. *Afr. J. Biotech.* 5, 1–14. doi:10.1016/j.enzmictec.2005.08.024

Feller, A., Machemer, K., Braun, E. L., and Grotewold, E. (2011). Evolutionary and Comparative Analysis of MYB and bHLH Plant Transcription Factors. *Plant J.* 66, 94–116. doi:10.1111/j.1365-313X.2010.04459.x

Feng, K., Hou, X.-L., Xing, G.-M., Liu, J.-X., Duan, A.-Q., Xu, Z.-S., et al. (2020). Advances in AP2/ERF Super-Family Transcription Factors in Plant. *Crit. Rev. Biotechnol.* 40, 750–776. doi:10.1080/07388551.2020.1768509

Finet, C., Fourquin, C., Vinauger, M., Berne-Dedieu, A., Chambrier, P., Paindavoine, S., et al. (2010). Parallel Structural Evolution of Auxin Response Factors in the Angiosperms. *Plant J.* 63, 952–959. doi:10.1111/j.1365-313X.2010.04292.x

Finn, R. D., Tate, J., Mistry, J., Coggill, P. C., Sammut, S. J., Hotz, H. R., et al. (2008). The Pfam Protein Families Database. *Nucleic Acids Res.* 36, D281–D288. doi:10.1093/nar/gkm960

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

Fu, M., Kang, H. K., Son, S.-H., Kim, S.-K., and Nam, K. H. (2014). A Subset of Arabidopsis RAV Transcription Factors Modulates Drought and Salt Stress Responses Independent of ABA. *Plant Cel Physiol* 55, 1892–1904. doi:10.1093/pcp/pcu118

Gao, C., Sun, J., Wang, C., Dong, Y., Xiao, S., Wang, X., et al. (2017). Genome-Wide Analysis of Basic/Helix-Loop-Helix Gene Family in Peanut and Assessment of its Roles in Pod Development. *PLoS ONE* 12, e0181843. doi:10.1371/journal.pone.0181843

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S. e., Wilkins, M. R., Appel, R. D., et al. (2005). Protein Identification and Analysis Tools on the ExPASy Server. *The Proteomics Protoc. handbook*, 571–607. doi:10.1385/1-59259-890-0:571

Ghorbani, R., Zakipour, Z., Alemzadeh, A., and Razi, H. (2020). Genome-Wide Analysis of AP2/ERF Transcription Factors Family in *Brassica Napus*. *Physiol. Mol. Biol. Plants* 26, 1463–1476. doi:10.1007/s12298-020-00832-z

Girardi, C. L., Rombaldi, C. V., Dal Cero, J., Nobile, P. M., Laurens, F., Bouzayen, M., et al. (2013). Genome-Wide Analysis of the AP2/ERF Superfamily in Apple and Transcriptional Evidence of ERF Involvement in Scab Pathogenesis. *Scientia Horticulturae* 151, 112–121. doi:10.1016/j.scienta.2012.12.017

Guo, B., Wei, Y., Xu, R., Lin, S., Luan, H., Lv, C., et al. (2016). Genome-Wide Analysis of *APETALA2/Ethylene-Responsive Factor (AP2/ERF)* Gene Family in Barley (*Hordeum Vulgare* L.). *PLoS ONE* 11, e0161322. doi:10.1371/journal.pone.0161322

Gutterson, N., and Reuber, T. L. (2004). Regulation of Disease Resistance Pathways by AP2/ERF Transcription Factors. *Curr. Opin. Plant Biol.* 7, 465–471. doi:10.1016/j.pbi.2004.04.007

Hirota, A., Kato, T., Fukaki, H., Aida, M., and Tasaka, M. (2007). The Auxin-Regulated AP2/EREBP GenePUCHIIs Required for Morphogenesis in the Early Lateral Root Primordium ofArabidopsis. *Plant Cell* 19, 2156–2168. doi:10.1105/tpc.107.050674

Hu, L., and Liu, S. (2011). Genome-Wide Identification and Phylogenetic Analysis of the *ERF* Gene Family in Cucumbers. *Genet. Mol. Biol.* 34, 624–634. doi:10.1590/S1415-47572011005000054

Hu, Y. X., Wang, Y. H., Liu, X. F., and Li, J. Y. (2004). *Arabidopsis* RAV1 Is Down-Regulated by Brassinosteroid and May Act as a Negative Regulator during Plant Development. *Cell Res* 14, 8–15. doi:10.1038/sj.cr.7290197

Huang, Y., Liu, Y., Zhang, M., Chai, M., He, Q., Jakada, B. H., et al. (2020). Genome-Wide Identification and Expression Analysis of the ERF Transcription Factor Family in Pineapple (*Ananas Comosus* (L.) Merr.). *Peer J.* 8, e10014. doi:10.7717/peerj.10014

Irish, V. F., and Sussex, I. M. (1990). Function of the Apetala-1 Gene during Arabidopsis floral Development. *Plant Cell* 2, 741–753. doi:10.1105/tpc.2.8.741

Ito, T. M., Polido, P. B., Rampim, M. C., Kaschuk, G., and Souza, S. G. H. (2014). Genome-Wide Identification and Phylogenetic Analysis of the AP2/ERF Gene Superfamily in Sweet orange (*Citrus Sinensis*). *Genet. Mol. Res.* 13, 7839–7851. doi:10.4238/201410.4238/2014.september.26.22

Jayaprakash, A., Thanmalagan, R. R., Roy, A., Arunachalam, A., and Lakshmi, P. (2019). Strategies to Understand *Aspergillus flavus* Resistance Mechanism in *Arachis hypogaea* L. *Curr. Plant Biol.* 20, 100123. doi:10.1016/j.cpb.2019.100123

Je, B. I., Piao, H. L., Park, S. J., Park, S. H., Kim, C. M., Xuan, Y. H., et al. (2010). RAV-Like1 Maintains Brassinosteroid Homeostasis via the Coordinated Activation of BRI1 and Biosynthetic Genes in Rice. *Plant Cell* 22, 1777–1791. doi:10.4161/psb.5.11.1335710.1105/tpc.109.069575

Jin, J., Zhang, H., Kong, L., Gao, G., and Luo, J. (2014). PlantTFDB 3.0: A portal for the Functional and Evolutionary Study of Plant Transcription Factors. *Nucl. Acids Res.* 42, D1182–D1187. doi:10.1093/nar/gkt1016

Jofuku, K. D., den Boer, B. G., Van Montagu, M., and Okamuro, J. K. (1994). Control of Arabidopsis Flower and Seed Development by the Homeotic Gene APETALA2. *Plant Cell* 6, 1211–1225. doi:10.1105/tpc.6.9.1211

Jofuku, K. D., Omidyar, P. K., Gee, Z., and Okamuro, J. K. (2005). Control of Seed Mass and Seed Yield by the floral Homeotic Gene APETALA2. *Proc. Natl. Acad. Sci.* 102, 3117–3122. doi:10.1073/pnas.0409893102

Kabir, S. M. T., Hossain, M. S., Bashar, K. K., Honi, U., Ahmed, B., Emdad, E. M., et al. (2021). Genome-Wide Identification and Expression Profiling of AP2/ERF Superfamily Genes under Stress Conditions in Dark Jute (*Corchorus Olitorius* L.). *Ind. Crops Prod.* 166, 113469. doi:10.1016/j.indcrop.2021.113469

Karniely, S., and Pines, O. (2005). Single Translation-Dual Destination: Mechanisms of Dual Protein Targeting in Eukaryotes. *EMBO Rep.* 6, 420–425. doi:10.1038/sj.embor.7400394

Kavas, M., Kizildogan, A., Gökdemir, G., and Baloglu, M. C. (2015). Genome-Wide Investigation and Expression Analysis of AP2-ERF Gene Family in Salt Tolerant Common Bean. *EXCLI J.* 14, 1187–1206. doi:10.17179/excli2015-600

Kavas, M., Baloğlu, M. C., Atabay, E. S., Ziplar, U. T., Daşgan, H. Y., and Ünver, T. (2016). Genome-Wide Characterization and Expression Analysis of Common Bean bHLH Transcription Factors in Response to Excess Salt Concentration. *Mol. Genet. Genomics.* 291, 129–143. doi:10.1007/s00438-015-1095-6

Kim, M. J., Ruzicka, D., Shin, R., and Schachtman, D. P. (2012). The *Arabidopsis* AP2/ERF Transcription Factor RAP2.11 Modulates Plant Response to Low-Potassium Conditions. *Mol. Plant* 5, 1042–1057. doi:10.1093/mp/sss003

Kitomi, Y., Ito, H., Hobo, T., Aya, K., Kitano, H., and Inukai, Y. (2011). The Auxin Responsive AP2/ERF Transcription Factor Crown Rootless5 Is Involved in crown Root Initiation in rice through the Induction of OsRR1, a Type-A Response Regulator of Cytokinin Signaling. *Plant J.* 67 (3), 472–484. doi:10.1111/j.1365-313x.2011.04610.x

Krizek, B., and Beth, A. (2009). Aintegumenta and Aintegumenta-LIKE6 Act Redundantly to Regulate Arabidopsis Floral Growth and Patterning. *Plant Physiol.* 150, 1916–1929. doi:10.1104/pp.109.141119

Lata, C., Mishra, A. K., Muthamilarasan, M., Bonthala, V. S., Khan, Y., and Prasad, M. (2014). Genome-Wide Investigation and Expression Profiling of AP2/ERF Transcription Factor Superfamily in Foxtail Millet (*Setaria Italica* L.). *PLoS ONE* 9, e113092. doi:10.1371/journal.pone.0113092

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

Letovsky, S., and Kasif, S. (2003). Predicting Protein Function from Protein/protein Interaction Data: A Probabilistic Approach. *Bioinformatics* 19, i197–i204. doi:10.1093/bioinformatics/btg1026

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, A., Yu, X., Cao, B. B., Peng, L. X., Gao, Y., Feng, T., et al. (2017a). LkAP2L2, an AP2/ERF Transcription Factor Gene of Larix Kaempferi, with Pleiotropic Roles in Plant branch and Seed Development. *Russ. J. Genet.* 53 (12), 1335–1342. doi:10.1134/S1022795417120079

Li, C.-W., Su, R.-C., Cheng, C.-P., Sanjaya, S. J., You, S.-J., Hsieh, T.-H., et al. (2011). Tomato RAV Transcription Factor Is a Pivotal Modulator Involved in the AP2/EREBP-Mediated Defense Pathway. *Plant Physiol.* 156, 213–227. doi:10.1104/pp.111.174268

Li, H., Wang, Y., Wu, M., Li, L., Li, C., Han, Z., et al. (2017b). Genome-Wide Identification of AP2/ERF Transcription Factors in Cauliflower and Expression Profiling of the ERF Family under Salt and Drought Stresses. *Front. Plant Sci.* 8, 946. doi:10.3389/fpls.2017.00946

Li, M.-Y., Xu, Z.-S., Huang, Y., Tian, C., Wang, F., and Xiong, A.-S. (2015b). Genome-wide Analysis of AP2/ERF Transcription Factors in Carrot (*Daucus Carota* L.) Reveals Evolution and Expression Profiles under Abiotic Stress. *Mol. Genet. Genomics* 290, 2049–2061. doi:10.1007/s00438-015-1061-3

Li, P., Chai, Z., Lin, P., Huang, C., Huang, G., Xu, L., et al. (2020). Genome-Wide Identification and Expression Analysis of AP2/ERF Transcription Factors in Sugarcane (*Saccharum Spontaneum* L.). *BMC Genomics* 21, 685. doi:10.1186/s12864-020-07076-x

Li, X.-J., Li, M., Zhou, Y., Hu, S., Hu, R., Chen, Y., et al. (2015a). Overexpression of Cotton *RAV1* Gene in *Arabidopsis* Confers Transgenic Plants High Salinity and Drought Sensitivity. *PLoS ONE* 10, e0118056. doi:10.1371/journal.pone.0118056

Li, X., Duan, X., Jiang, H., Sun, Y., Tang, Y., Yuan, Z., et al. (2006). Genome-Wide Analysis of Basic/Helix-Loop-Helix Transcription Factor Family in rice and *Arabidopsis*. *Plant Physiol.* 141, 1167–1184. doi:10.1104/pp.106.080580

Licausi, F., Giorgi, F. M., Zenoni, S., Osti, F., Pezzotti, M., and Perata, P. (2010). Genomic and Transcriptomic Analysis of the AP2/ERF Superfamily in *Vitis vinifera*. *BMC Genomics* 11, 719. doi:10.1186/1471-2164-11-719

Licausi, F., Ohme-Takagi, M., and Perata, P. (2013). APETALA2/Ethylene Responsive Factor (AP2/ERF) Transcription Factors: Mediators of Stress Responses and Developmental Programs. *New Phytol.* 199, 639–649. doi:10.1111/nph.12291

Liu, C., Wu, Y., and Wang, X. (2012). bZIP Transcription Factor OsbZIP52/RISBZ5: A Potential Negative Regulator of Cold and Drought Stress Response in rice. *Planta* 235, 1157–1169. doi:10.1007/s00425-011-1564-z

Liu, S., Wang, X., Wang, H., Xin, H., Yang, X., Yan, J., et al. (2013). Genome-Wide Analysis of *ZmDREB* Genes and Their Association with Natural Variation in Drought Tolerance at Seedling Stage of *Zea mays* L. *Plos Genet.* 9, e1003790. doi:10.1371/journal.pgen.1003790

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

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

Maes, T., Van de Steene, N., Zethof, J., Karimi, M., D'Hauw, M., Mares, G., et al. (2001). Petunia Ap2-Like Genes and Their Role in Flower and Seed Development. *Plant Cell* 13, 229–244. doi:10.1105/tpc.13.2.229

Magnani, E., Sjölander, K., and Hake, S. (2004). From Endonucleases to Transcription Factors: Evolution of the AP2 DNA Binding Domain in Plants[W]. *The Plant Cell* 16, 2265–2277. doi:10.1105/tpc.104.023135

Manning, K., Tör, M., Poole, M., Hong, Y., Thompson, A. J., King, G. J., et al. (2006). A Naturally Occurring Epigenetic Mutation in a Gene Encoding an SBP-Box Transcription Factor Inhibits Tomato Fruit Ripening. *Nat. Genet.* 38, 948–952. doi:10.1038/ng1841

Mao, X., Zhang, H., Qian, X., Li, A., Zhao, G., and Jing, R. (2012). TaNAC2, a NAC-Type Wheat Transcription Factor Conferring Enhanced Multiple Abiotic Stress Tolerances in *Arabidopsis*. *J. Exp. Bot.* 63, 2933–2946. doi:10.1093/jxb/err462

Meng, J., Yang, J., Peng, M., Liu, X., and He, H. (2020). Genome-Wide Characterization, Evolution, and Expression Analysis of the Leucine-Rich Repeat Receptor-Like Protein Kinase (LRR-RLK) Gene Family in Medicago Truncatula. *Life* 10, 176. doi:10.3390/life10090176

Nakano, T., Suzuki, K., Fujimura, T., and Shinshi, H. (2006). Genome-Wide Analysis of the ERF Gene Family in *Arabidopsis* and Rice. *Plant Physiol.* 140, 411–432. doi:10.1104/pp.105.073783

Nakashima, K., Takasaki, H., Mizoi, J., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2012). NAC Transcription Factors in Plant Abiotic Stress Responses. *Biochim. Biophys. Acta (Bba) - Gene Regul. Mech.* 1819, 97–103. doi:10.1016/j.bbagrm.2011.10.005

Nimitr, V., Patcharin, S., Chutipong, A., Sanun, J., and Aran, P. (2003). Effect of Water Stress on Yield and Agronomic Characters of Peanut (*Arachis hypogaea* L.). *Songklanakar J. Sci. Tech.* 25, 283–288.

Novillo, F., Alonso, J. M., Ecker, J. R., and Salinas, J. (2004). CBF2/DREB1C Is a Negative Regulator of CBF1/DREB1B and CBF3/DREB1A Expression and Plays a central Role in Stress Tolerance in *Arabidopsis*. *Proc. Natl. Acad. Sci.* 101, 3985–3990. doi:10.1073/pnas.0303029101

Okamuro, J. K., Caster, B., Villarroel, R., Van Montagu, M., and Jofuku, K. D. (1997). The AP2 Domain of APETALA2 Defines a Large New Family of DNA Binding Proteins in *Arabidopsis*. *Proc. Natl. Acad. Sci.* 94, 7076–7081. doi:10.1073/pnas.94.13.7076

Pan, Y.-B., Burner, D. M., and Legendre, B. L. (2000). An Assessment of the Phylogenetic Relationship Among Sugarcane and Related Taxa Based on the Nucleotide Sequence of 5S rRNA Intergenic Spacers. *Genetica* 108, 285–295. doi:10.1023/A:1004191625603

Passioura, J. B. (1983). Roots and Drought Resistance. *Agric. Water Manage.* 7, 265–280. doi:10.1016/0378-3774(83)90089-6

Pierret, A., Doussan, C., Capowiez, Y., Bastardie, F., and Pagès, L. (2007). Root Functional Architecture: A Framework for Modeling the Interplay between Roots and Soil. *Vadose Zone J.* 6, 269–281. doi:10.2136/vzj2006.0067

Pietsch, C., Sreenivasulu, N., Wobus, U., and Röder, M. S. (2009). Linkage Mapping of Putative Regulator Genes of Barley Grain Development Characterized by Expression Profiling. *BMC Plant Biol.* 9, 4. doi:10.1186/1471-2229-9-4

Rademacher, E. H., Möller, B., Lokerse, A. S., Llavata-Peris, C. I., van den Berg, W., and Weijers, D. (2011). A Cellular Expression Map of the Arabidopsis AUXIN Response Factor Gene Family. *Plant J.* 68, 597–606. doi:10.1111/j.1365-313X.2011.04710.x

Reddy, T. Y., Reddy, V. R., and Anbumozhi, V. (2003). Physiological Responses of Groundnut (*Arachis Hypogea* L.) to Drought Stress and its Amelioration: A Critical Review. *Plant Growth Regul.* 41, 75–88. doi:10.1023/A:1027353430164

Riechmann, J. L., and Meyerowitz, E. M. (1998). The AP2/EREBP Family of Plant Transcription Factors. *Biol. Chem.* 379, 633–646. doi:10.1515/bchm.1998.379.6.633

Rushton, P. J., Somssich, I. E., Ringler, P., and Shen, Q. J. (2010). WRKY Transcription Factors. *Trends Plant Sci.* 15, 247–258. doi:10.1016/j.tplants.2010.02.006

Sakuma, Y., Liu, Q., Dubouzet, J. G., Abe, H., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2002). DNA-Binding Specificity of the ERF/AP2 Domain of Arabidopsis DREBs, Transcription Factors Involved in Dehydration- and Cold-Inducible Gene Expression. *Biochem. Biophysical Res. Commun.* 290, 998–1009. doi:10.1006/bbrc.2001.6299

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* 13, 2498–2504. doi:10.1101/gr.1239303

Sharoni, A. M., Nuruzzaman, M., Satoh, K., Shimizu, T., Kondoh, H., Sasaya, T., et al. (2011). Gene Structures, Classification and Expression Models of the AP2/EREBP Transcription Factor Family in rice. *Plant Cel Physiol* 52, 344–360. doi:10.1093/pcp/pcq196

Silva, G. F. F. E., Silva, E. M., da Silva Azevedo, M., Guivin, M. A. C., Ramiro, D. A., Figueiredo, C. R., et al. (2014). MicroRNA156-Targeted SPL/SBP Box Transcription Factors Regulate Tomato Ovary and Fruit Development. *Plant J.* 78, 604–618. doi:10.1111/tpj.12493

Singh, K., Foley, R. C., and Luis, O. S. (2002). Transcription Factors in Plant Defense and Stress Responses. *Curr. Opin. Plant Biol.* 5, 430–436. doi:10.1016/s1369-5266(02)00289-3

Soares, V. L. F., Rodrigues, S. M., de Oliveira, T. M., de Queiroz, T. O., Lima, L. S., Hora-Júnior, B. T., et al. (2011). Unraveling New Genes Associated with Seed Development and Metabolism in Bixa Orellana L. By Expressed Sequence Tag (EST) Analysis. *Mol. Biol. Rep.* 38 (2), 1329–1340. doi:10.1007/s11033-010-0234-8

Sohn, K. H., Lee, S. C., Jung, H. W., Hong, J. K., and Hwang, B. K. (2006). Expression and Functional Roles of the Pepper Pathogen-Induced Transcription Factor RAV1 in Bacterial Disease Resistance, and Drought and Salt Stress Tolerance. *Plant Mol. Biol.* 61, 897–915. doi:10.1007/s11103-006-0057-0

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

Songsri, P., Jogloy, S., Vorasoot, N., Akkasaeng, C., Patanothai, A., and Holbrook, C. C. (2008). Root Distribution of Drought-Resistant Peanut Genotypes in Response to Drought. *J. Agron. Crop Sci.* 194, 92–103. doi:10.1111/j.1439-037X.2008.00296.x

Soni, P., Gangurde, S. S., Ortega-Beltran, A., Kumar, R., Parmar, S., Sudini, H. K., et al. (2020). Functional Biology and Molecular Mechanisms of Host-Pathogen Interactions for Aflatoxin Contamination in Groundnut (*Arachis hypogaea* L.) and maize (*Zea mays* L.). *Front. Microbiol.* 11, 227. doi:10.3389/fmicb.2020.00227

Sun, L., Hu, R., Shen, G., and Zhang, H. (2013). Genetic Engineering Peanut for Higher Drought- and Salt-Tolerance. *Food Nutr. Sci.* 04, 1–7. doi:10.4236/fns.2013.46A001

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING V11: Protein-Protein Association Networks with Increased Coverage, Supporting Functional Discovery in Genome-Wide Experimental Datasets. *Nucleic Acids Res.* 47, D607–D613. doi:10.1093/nar/gky1131

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol. Biol. Evol.* 28, 2731–2739. doi:10.1093/molbev/msr121

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol. Bio. Evol.* 30, 2725–2729. doi:10.1093/molbev/mst197

Terol, J., Nueda, M. J., Ventimilla, D., Tadeo, F., and Talon, M. (2019). Transcriptomic Analysis of Citrus Clementina Mandarin Fruits Maturation Reveals a MADS-Box Transcription Factor that Might Be Involved in the Regulation of Earliness. *BMC Plant Biol.* 19, 47. doi:10.1186/s12870-019-1651-z

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weighting, Position-specific gap Penalties and Weight Matrix Choice. *Nucl. Acids Res.* 22, 4673–4680. doi:10.1093/nar/22.22.4673

Trevaskis, B., Bagnall, D. J., Ellis, M. H., Peacock, W. J., and Dennis, E. S. (2003). MADS Box Genes Control Vernalization-Induced Flowering in Cereals. *Proc. Natl. Acad. Sci.* 100, 13099–13104. doi:10.1073/pnas.1635053100

Tsutsui, T., Yamaji, N., and Feng Ma, J. (2011). Identification of a Cis-Acting Element of ART1, a C2H2-Type Zinc-Finger Transcription Factor for Aluminum Tolerance in rice. *Plant Physiol.* 156, 925–931. doi:10.1104/pp.111.175802

Tuteja, N. (2007). Abscisic Acid and Abiotic Stress Signaling. *Plant Signaling Behav.* 2, 135–138. doi:10.4161/psb.2.3.4156

Ulm, R., Baumann, A., Oravecz, A., Mate, Z., Adam, E., Oakeley, E. J., et al. (2004). Genome-Wide Analysis of Gene Expression Reveals Function of the bZIP Transcription Factor HY5 in the UV-B Response of Arabidopsis. *Proc. Natl. Acad. Sci.* 101, 1397–1402. doi:10.1073/pnas.0308044100

Vahdati, K., and Lotfi, N. (2013). "Abiotic Stress Tolerance in Plants with Emphasizing on Drought and Salinity Stresses in walnut," in *Abiotic Stress-Plant Responses and Applications in Agriculture*. Editors K Vahdati and C Leslie (Rijeka: InTech), 307–365. doi:10.5772/56078

Voorrips, R. E. (2002). MapChart: Software for the Graphical Presentation of Linkage Maps and QTLs. *J. Hered.* 93, 77–78. doi:10.1093/jhered/93.1.77

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, 1–16. doi:10.1007/s13205-020-2123-8

Wan, L., Wu, Y., Huang, J., Dai, X., Lei, Y., Yan, L., et al. (2014). Identification of ERF Genes in Peanuts and Functional Analysis of AhERF008 and AhERF019 in Abiotic Stress Response. *Funct. Integr. Genomics* 14, 467–477. doi:10.1007/s10142-014-0381-4

Wang, J., Zhou, J., Zhang, B., Vanitha, J., Ramachandran, S., and Jiang, S.-Y. (2011). Genome-Wide Expansion and Expression Divergence of the Basic Leucine Zipper Transcription Factors in Higher Plants with an Emphasis on SorghumF. *J. Integr. Plant Biol.* 53, 212–231. doi:10.1111/j.1744-7909.2010.01017.x

Wei, G., Pan, Y., Lei, J., and Zhu, Y.-X. (2005). Molecular Cloning, Phylogenetic Analysis, Expressional Profiling and *In Vitro* Studies of TINY2 from *Arabidopsis thaliana*. *BMB Rep.* 38, 440–446. doi:10.5483/BMBRep.2005.38.4.440

Wei, Y., Chang, Y., Zeng, H., Liu, G., He, C., and Shi, H. (2018). RAV Transcription Factors Are Essential for Disease Resistance against Cassava Bacterial Blight via Activation of Melatonin Biosynthesis Genes. *J. Pineal Res.* 64, e12454. doi:10.1111/jpi.12454

Woo, H. R., Kim, J. H., Kim, J., Kim, J., Lee, U., Song, I.-J., et al. (2010). The RAV1 Transcription Factor Positively Regulates Leaf Senescence in *Arabidopsis*. *J. Exp. Bot.* 61, 3947–3957. doi:10.1093/jxb/erq206

Wuitschick, J. D., Lindstrom, P. R., Meyer, A. E., and Karrer, K. M. (2004). Homing Endonucleases Encoded by Germ Line-Limited Genes in *Tetrahymena Thermophila* Have APETELA2 DNA Binding Domains. *Eukaryot. Cel* 3, 685–694. doi:10.1128/EC.3.3.685-694.2004

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 ImprovementF. *J. Integr. Plant Biol.* 53, 570–585. doi:10.1111/j.1744-7909.2011.01062.x

Yao, Y., He, R. J., Xie, Q. L., Zhao, X. H., Deng, X. M., He, J. B., et al. (2017). Ethylene Response Factor 74 (ERF74) Plays an Essential Role in Controlling a Respiratory Burst Oxidase Homolog D (RbohD)-

dependent Mechanism in Response to Different Stresses in Arabidopsis. *NEW PHYTOL.* 213, 1667–1681. doi:10.1111/nph.14278

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

Zhu, Y., Liu, X., Gao, Y., Li, K., and Guo, W. (2021). Transcriptome-Based Identification of AP2/ERF Family Genes and Their Cold-Regulated Expression during the Dormancy Phase Transition of Chinese Cherry Flower Buds. *Scientia Horticulturae* 275, 109666. doi:10.1016/j.scienta.2020.109666

Zhuang, J., Chen, J.-M., Yao, Q.-H., Xiong, F., Sun, C.-C., Zhou, X.-R., et al. (2011). Discovery and Expression Profile Analysis of AP2/ERF Family Genes from *Triticum aestivum*. *Mol. Biol. Rep.* 38, 745–753. doi:10.1007/s11033-010-0162-7

Zhuang, W., Chen, H., Yang, M., Wang, J., Pandey, M. K., Zhang, C., et al. (2019). The Genome of Cultivated Peanut Provides Insight into Legume Karyotypes, Polyploid Evolution and Crop Domestication. *Nat. Genet.* 51, 865–876. doi:10.1038/s41588-019-0402-2

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.

Copyright © 2021 Cui, Haider, Chai, Guo, Du, Li, Dong, Huang, Zheng, Shi, Zhang and Han. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

GLOSSARY

ABRE abscisic acid responsiveness element

Ala alanine

AP2/ERF APETALA2/ethylene response element-binding factor

Asp aspartate

AuxRR auxin responsiveness

ARE anaerobic response regulatory elements

ARF auxin response factors

bHLH basic helix–loop–helix

bZIP basic leucine zipper

Cs Cucumis sativus

DREB dehydration-responsive-element-binding

Elu glutamic acid

GRF growth-regulating factor

GARE gibberellin-responsive element

KDa kilodalton

LTR low-temperature response elements

MST monosaccharide transporter

MEGA Molecular Evolutionary Genetics Analysis

MEME Multiple Expectation Maximization for Motif Elicitation

Mya million years ago

NAC NAM, ATAF, and CUC

NCBI National Center for Biotechnology Information

PEG polyethylene glycol

pI isoelectric point

qRT-PCR quantitative real-time

RAV related to ABI3/VP

SBP SQUAMOSA promoter-binding protein

SMART imple Modular Architecture Research Tool

TFs transcription factors

Val valine

WUN wound.

frontiers in Genetics

ORIGINAL RESEARCH
published: 05 January 2022
doi: 10.3389/fgene.2021.799805

Stress Conditions Modulate the Chromatin Interactions Network in *Arabidopsis*

*Vikash Kumar Yadav[1,2], Swadha Singh[1,3†], Amrita Yadav[1†], Neha Agarwal[1], Babita Singh[1], Siddhi Kashinath Jalmi[4], Vrijesh Kumar Yadav[1], Vipin Kumar Tiwari[1,2], Verandra Kumar[5], Raghvendra Singh[6] and Samir Vishwanath Sawant[1,2]**

[1]CSIR-National Botanical Research Institute, Lucknow, India, [2]Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, India, [3]School of Natural Sciences, University of California, Merced, Merced, CA, United States, [4]Department of Botany, Goa University, Taleigão, India, [5]Department of Botany, Manyawar Kanshiram Government Degree College, Aligarh, India, [6]CSIR-Centre for Cellular & Molecular Biology, Hyderabad, India

Stresses have been known to cause various responses like cellular physiology, gene regulation, and genome remodeling in the organism to cope and survive. Here, we assessed the impact of stress conditions on the chromatin-interactome network of *Arabidopsis thaliana*. We identified thousands of chromatin interactions in native as well as in salicylic acid treatment and high temperature conditions in a genome-wide fashion. Our analysis revealed the definite pattern of chromatin interactions and stress conditions could modulate the dynamics of chromatin interactions. We found the heterochromatic region of the genome actively involved in the chromatin interactions. We further observed that the establishment or loss of interactions in response to stress does not result in the global change in the expression profile of interacting genes; however, interacting regions (genes) containing motifs for known TFs showed either lower expression or no difference than non-interacting genes. The present study also revealed that interactions preferred among the same epigenetic state (ES) suggest interactions clustered the same ES together in the 3D space of the nucleus. Our analysis showed that stress conditions affect the dynamics of chromatin interactions among the chromatin loci and these interaction networks govern the folding principle of chromatin by bringing together similar epigenetic marks.

Keywords: stress, chromatin-chromatin interactions, genome organization, epigenetic state, Hi-C, heterochromatin, QTL

OPEN ACCESS

Edited by:
Vijay Gahlaut, Institute of Himalayan Bioresource Technology (CSIR), India

Reviewed by:
Wen-Feng Nie, Yangzhou University, China
Junghoon Park, Konkuk University, South Korea

****Correspondence:***
Samir Vishwanath Sawant
samirsawant@nbri.res.in

†These authors have contributed equally to this work.

Specialty section:
This article was submitted to Plant Genomics, a section of the journal Frontiers in Genetics

Received: *22 October 2021*
Accepted: *15 November 2021*
Published: *05 January 2022*

Citation:
Yadav VK, Singh S, Yadav A, Agarwal N, Singh B, Jalmi SK, Yadav VK, Tiwari VK, Kumar V, Singh R and Sawant SV (2022) Stress Conditions Modulate the Chromatin Interactions Network in Arabidopsis. *Front. Genet. 12:799805. doi: 10.3389/fgene.2021.799805*

INTRODUCTION

The haploid *A. thaliana* genome contains approximately 125 million base pairs of DNA packaged into five chromosomes (Kaul et al., 2000). That makes a total 250 million base pairs of DNA in a single diploid cell of *A. thaliana*, which spans a total length of ~8.5 cm. This stretch of DNA is approximately 16,000 times larger than the diameter of *A. thaliana* nucleus (Dittmer et al., 2007). Thus, the DNA packaging in the nucleus is accomplished in a very organized way which prevents the

Abbreviations: ES, epigenetic state; GWAS, genome-wide association study; HT, high temperature; NC, native condition; QTL, quantitative trait loci; SA, salicylic acid; TFs, transcription factors.

DNA from becoming an unmanageable tangle. Interestingly, after being packaged so compactly it manages in such a way that their distal regulatory elements remain accessible to their target gene for their regulation. Chromatin-chromatin interaction has been identified as an important mechanism for such regulation (Li et al., 2012; Sun et al., 2020; Yadav et al., 2021). A distal element can specifically interact with its target gene situated on the same or different chromosome by looping (Dean, 2011). Loop formation is thus an integral part of chromatin organization which facilitates interactions between distal genomic elements (Deng et al., 2012; Sandhu et al., 2012).

Several high throughput approaches such as Formaldehyde-Assisted Isolation of Regulatory Elements (FAIR) (Simon et al., 2012), Chromatin Immunoprecipitation Sequencing (ChIP-Seq) (Furey, 2012), and DNaseI-Seq (Song and Crawford, 2010) are available as standard experimental methods for the identification of regulatory elements. However, their primary limitations are that they cannot determine the precise association of distal regulatory elements with target genes and spatial conformation required for their optimal activity. Previously, cytogenetic techniques and microscopic observation have been used to study chromosomal organization but over the last 20 years, our knowledge about chromosomal architecture enhances with the advancement of high-resolution techniques. There are several techniques which have been used widely to identify local and global chromatin interactome network, for example, chromosome conformation capture (3C) (Dekker et al., 2002), circular chromosome conformation capture (4C) (Zhao et al., 2006), chromosome conformation capture carbon copy (5C) (Dostie et al., 2006), high-throughput chromosome conformation capture (Hi-C) (Lieberman-Aiden et al., 2009), DNase Hi-C (Ma et al., 2018), Capture Hi-C (Eijsbouts et al., 2019), INT-Hi-C (Yadav et al., 2021), and chromatin interaction analysis using paired-end tag (ChIA-PET) (Fullwood and Ruan, 2009).

Three-dimensional chromatin organization is necessary for many biological processes including transcriptional regulation, replication, and repair (Ling et al., 2006; Zhao et al., 2006). A series of publications on *A. thaliana* showed interest of the plant community in chromatin organization and its function (Moissiard et al., 2012; Grob et al., 2013; Grob et al., 2014; Feng et al., 2014; Sequeira-Mendes et al., 2014; Wang et al., 2015, 2016; Zhu et al., 2015; Liu et al., 2016; Bi et al., 2017; Zhang et al., 2019; Yadav et al., 2021). Chromatin architecture of wild-type and mutant (atmorc6-1) *A. thaliana* showed a similar pattern although the interaction frequency varies among different chromatin regions in wild-type and mutant (Moissiard et al., 2012). Further, it is reported that chromosomes interact with each other via pericentromeric and heterochromatic regions (Feng et al., 2014). There exists a strong correlation between chromosomal architecture and epigenetic landscape (Grob et al., 2014; Yadav et al., 2021). The importance of chromatin looping was identified in the b1 locus of maize which is required for both paramutation and its high expression (Louwers et al., 2009a). Similarly, distal regulatory element is required for gene activation of FLC locus in *A. thaliana* (Crevillén et al., 2012). Thus, it is interesting and important to identify these distal regulatory elements in plants and their regulation. Thus, global mapping of chromatin interactions in *A. thaliana* is likely to uncover the genome architecture and its regulation.

Plants being sessile organisms always faced various stress conditions (both biotic and abiotic), however, very few efforts have been made to understand how stress conditions influence the chromatin interactions (Sun et al., 2020; Li et al., 2021; Liang et al., 2021). Thus, our present study assesses the impact of biotic and abiotic stresses on plant chromatin interaction networks. We used Hi-C to identify chromatin interactions and the impact of stress conditions on chromatin interactome. Our study revealed thousands of chromatin interactions in native condition (NC), heat treatment (HT, abiotic stress), and salicylic acid (SA, mimic biotic stress) treated *A. thaliana* and the impact of stresses on these interaction dynamics. We also investigate the correlation between epigenetic state (ES), chromatin interaction network, and gene expression. Our study will help to understand how the stress conditions affect the chromatin organization in plants which may directly or indirectly affect the genome regulation, and hence the organism response to the external environment.

MATERIALS AND METHODS

Plant Material

Seeds of *A. thaliana* (Columbia-0 ecotype) (ABRC; https://abrc.osu.edu/) were germinated and grown for 6 weeks on solrite under the long-day conditions (22°C, 16-h light and 8-h dark), after seeds have been stratified on soil (solrite) supplemented with water at 4°C for 4 days. For SA (Sigma) treatment 2 mM SA was sprayed on the aerial part of the plant material 2 days before nuclei were harvested. Plants grown as mentioned were exposed to 40°C for 1 h for the high-temperature treatment. The aerial tissue of native condition (NC) and treated plants (SA and HT) were used for preparing Hi-C and RNA-Seq libraries.

Fixation of Plant Tissue and Nuclei Preparation

The aerial tissue of 6-week-old NC, SA, and HT treated plants were cross-linked for 1 h separately by adding 37% formaldehyde to a final concentration of 1% in extraction buffer (2 M Hexylene glycol, 20 mM PIPES, pH 7.0, 10 mM $MgCl_2$, 5 mM ß-mercaptoethenol). After cross-linking the remaining formaldehyde was sequestered by adding 1/16 volume of 2 M ice-cold glycine for 10 min. The remaining solution was decanted, and tissue was rinsed three times with ice-cold milli-Q (MQ). Eventually, tissues were dried using paper towels and frozen in liquid N_2. Nuclei were isolated from the cross-linked samples as described (Bowler et al., 2004; Louwers et al., 2009b). The quality of nuclei was checked using fluorescent microscopy by DAPI staining and DNA quality and quantity were estimated using agarose gel electrophoresis and Quant-iT assay.

Hi-C Libraries Preparation and Sequencing

A total of 6 Hi-C libraries, 2 for NC, 2 for SA and HT treated samples were prepared according to Lieberman-Aiden et al.

(2009) and van Berkum et al. (2010) with some modifications adapted for plant samples as described in Louwers et al. (2009b). See **Supplementary File S4** for a detailed experimental procedure. Amplified Hi-C libraries were sequenced on a HiSeq2500 sequencer obtaining paired-end (100 × 2 bp) reads.

RNA-Seq and Data Analysis

The 6 RNA libraries for three conditions, two for native condition Col-0 (NC), two for HT and two for SA treated samples were prepared. Briefly, total RNA was isolated using Spectrum™ Plant Total RNA Kit (Sigma) and libraries were prepared with the Illumina standard protocol. RNA-seq reads were aligned to *A. thaliana* reference genome (TAIR10) using TopHat with default parameters (Trapnell et al., 2009). Normalized FPKM (fragments per kilobase per million mapped reads) counts for each gene were calculated using Cufflink (Trapnell et al., 2012). The differentially expressed genes (up- and down-regulated gene) were determined by log2 fold change between untreated (NC) and treated (HT and SA) samples.

Mapping and Filtering Uninformative Reads

For Hi-C analysis, reads were filtered based on the quality score using FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/index.html). Each end of paired reads was aligned separately to the *A. thaliana* reference genome (TAIR 10) using Bowtie v1.1.0 (Swarbreck et al., 2007; Langmead et al., 2009). Mapped reads were used for further downstream analysis. Read pairs that were not representative of true interactions like continuous genomic fragments, self-ligation, or re-ligation products were omitted using HOMER command line–removePEbg–removeSelfLigation. We had considered only those reads in which one or both of the paired-end reads have HindIII restriction site within the fragment length estimated from 3′ end of the reads (Heinz et al., 2010).

Normalizing Hi-C Data

Hi-C data was normalized to avoid the biasness due to its mapping ability, variable number of restriction sites in a region, or technical artifacts (inaccessibility of restriction enzyme to the DNA) and linear distance between interacting regions. The expected number of reads in any given genomic region is calculated based on the number of reads in all other regions of the genome. The expected number of reads between any two regions depends on both the linear distance and sequencing depth of the library. The expected number of reads was calculated with the following equation (Heinz et al., 2010).

$$eij = f(i - j)\frac{(n^{*}i)(n^{*}j)}{N^{*}}$$

where f is the expected number of reads as a function of distance, N* is the total number of reads, and n* is the estimated number of interacting reads at each region i and j. HOMER used simple hill climbing optimization to calculate inferred total reads through the difference between the observed and the expected number of reads.

Generation of Interaction Matrices

The interaction matrix is the simplified way to represent the Hi-C data where interaction frequencies between any two loci can be visualized. To create the contact matrix, the genome was divided into 200 kb bin size (Moissiard et al., 2012). The interaction matrix was generated based on the frequency of interacting reads between the two bins. Contact matrix corresponds to the number of interacting reads between locus i and j. The row and column correspond to coordinates of genomic regions and the corresponding value provides interaction information between each locus. We had generated normalized and correlation interactions matrices. Interaction matrices were normalized assuming that each bin has an equal chance of interaction with all other bins in the genome and computed by ratios between the total observed and the expected number of reads in a given bin size. The interaction matrix reveals which parts of chromosomes are positioned close together or apart from each other in the nucleus. A correlation matrix is based on the Pearson correlation coefficient and considers how each bin interacts with all other bins.

Data Visualization

To visualize high-resolution interaction data, we generated heat maps with MeV (v4.9) using interaction matrices (Saeed et al., 2003). Circos (v0.69) was used to visualize significant *cis* and *trans* interaction networks (Krzywinski et al., 2009).

Identification of Significant Interaction and Annotation

Significant interactions were identified based on the premise behind the enrichment of observed interacting reads over the expected. It searches the genome for a pair of loci that have more interacting reads (observed) than would be expected by chance. For two potentially interacting loci, HOMER model their expected read count using the cumulative binomial distribution, where it calculates the expected read count possibly mapped between the genomic loci and the number of observed read count between the genomic loci. We had identified the significant interactions at a resolution of 1 Kb with *p*-value cut-off 0.05 at default parameter. To extend this, we annotate these significant interactions to explore what these coordinates represent in *A. thaliana* genome using TAIR10 annotation.

Expression Profile of Interacting Genes

To check the effect of interactions on the expression profile of interacting genes, we analyzed the expression profile of interacting genes with the control sets of genes. For control, we choose an equal number of random genes that are identified as non-interacting in our study. The expression of both interacting and non-interacting sets was extracted from the RNA-seq data of the respective condition (**Supplementary File S5**).

Motifs Identification

For the identification of conserved motifs, we considered only those interacting sequences in which at least one partner was a

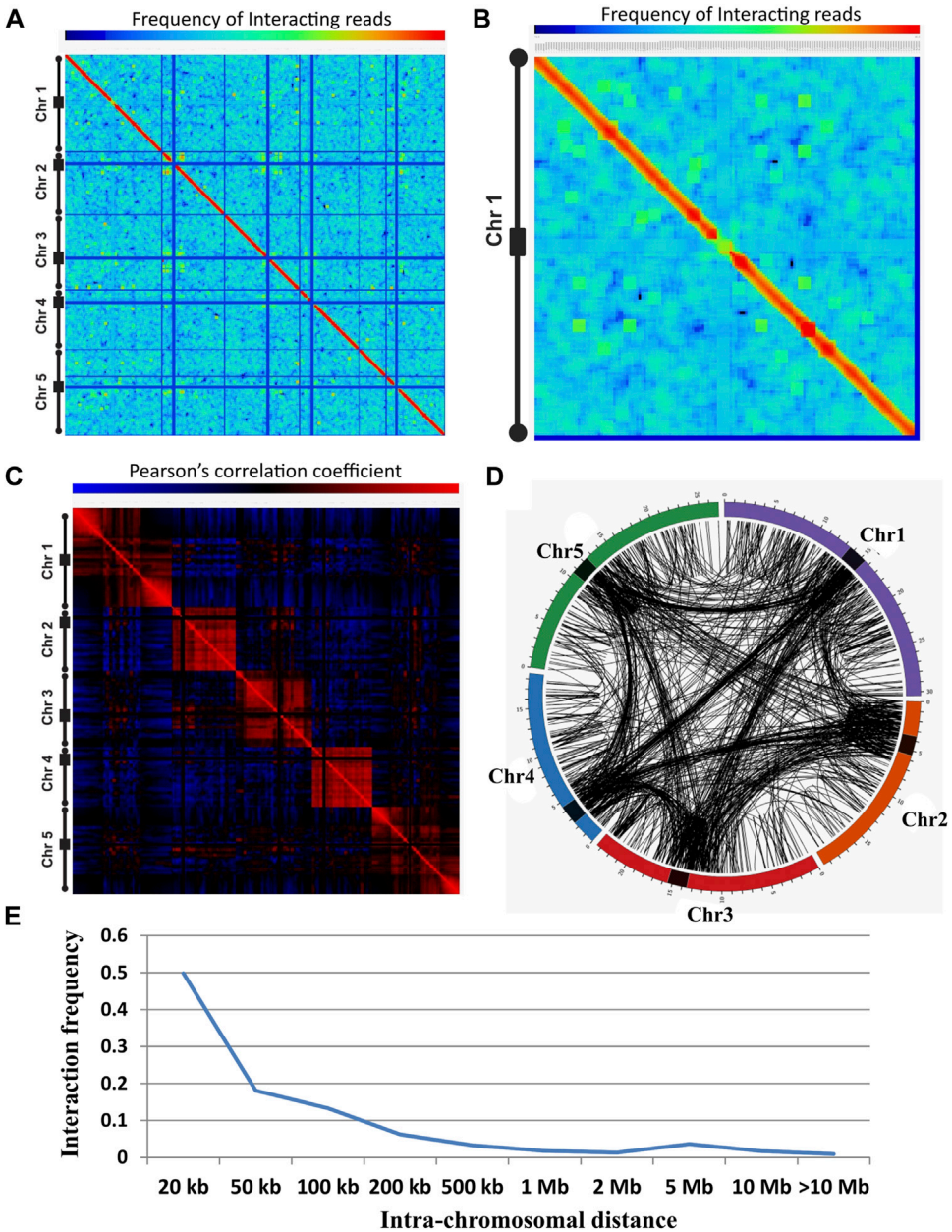


FIGURE 1 | Chromatin interactions landscape in *A. thaliana*. **(A)** Genome-wide normalized interaction matrix at 200 kb resolution for *Arabidopsis* genome (NC), showing intense red diagonal representing the enrichment of interacting reads in close proximity. Blue line represents reads that are less enriched in the centromeric and telomeric region of the genome. The color bar ranging from blue to red represents the lower to higher enrichment of interacting reads. The five different chromosomes are indicated with a black line, in which a rectangular box represents the pericentromeric and circles represent the telomeric region of the chromosome. **(B)** Normalized interaction matrix for chromosome 1 showing several spots of enriched interacting reads in the genome indicating the presence of long-distance chromatin-chromatin interactions. **(C)** Genome-wide correlation interaction matrix representing the correlation among the interacting region of HT library. The red color represents the positive correlation and the blue color represents the negative correlation between the two regions. Correlation matrix suggests two distinct regions of the genome showing positively correlated interacting regions and negatively correlated non-interacting regions. **(D)** Circos representing the genome-wide identified significant interaction of NC library. Interactions were represented in the chromosomes through the black line connecting the two points (interacting regions). Spans that link the regions within the chromosome represent *cis* interactions while spans that link the regions between the chromosomes represent *trans* interactions. The outermost colored circle is the graphical representation of *Arabidopsis* chromosomes and the black rectangular box on it represents the centromeric region of each chromosome. Chromosome numbers are indicated after the chromosome (Chr) abbreviation. **(E)** Plot showing the relationship of interactions frequency with the linear physical distance along the chromosome for NC library. Intra-chromosomal interaction frequency decreases with increasing linear distance on the chromosomes.

protein-coding gene. Sequences corresponding to these regions in all three conditions (NC, HT, and SA) were subjected to motif identification through MEME (http://meme.nbcr.net/meme/tools/meme). One kb region of interacting sequences was extracted from TAIR10. Total 1485, 1028 and 1196 non-redundant sequences of HT, SA, and NC were subjected to motif prediction. MEME, v4.11.2, was deployed using ZOOPS model, with motif width 6–10 bases, Evalue 0.001, and maximum numbers of motifs to return were 10 (Bailey et al., 2009). Predicted motifs were further annotated with the help of the STAMPS tool using the AGRIS database (The *Arabidopsis* Gene Regulatory Information Server, http://*Arabidopsis*.med.ohio-state.edu/) and identified the binding site of known TFs in the interacting regions (Mahony and Benos, 2007; Yilmaz et al., 2011).

Association of Interacting Regions with Epigenetic States

To identify whether the captured interacting regions have any preferential distribution in previously reported 9 ES (Sequeira-Mendes et al., 2014), the coordinates of interacting regions were mapped onto the coordinates of 9 different ES of *A. thaliana* to identify the ES in the interacting regions. One region may fall into more than one ES as these predefined ES have overlapping regions.

GWAS Enrichment

We mapped the publicly available GWAS hit (Atwell et al., 2010) to the interacting and non-interacting region (control). Background frequency was calculated as total unique SNP per bp of the genome.

RESULTS

Stress Condition Modulate Chromatin Interactome

The present study aims at understanding the dynamics of chromatin interaction during stress conditions, thus HT representing abiotic stress and SA mimic biotic stress was selected to capture chromatin interactions. We captured the chromatin interactions in NC and in HT and SA treated *A. thaliana* (Col-0) using Hi-C. A total of ~262 million paired reads were obtained, and of these ~87 million for NC (43.5 and 43.4 million for biological replicates 1 and 2), ~81 million for HT (39.7 and 41.6 million for biological replicates 1 and 2), and ~94 million for SA (55.8 and 38.6 million for biological replicates 1 and 2). To increase the depth of data we combined the biological replicates in the subsequent analysis. Heat map at a resolution of 200 kb effectively shows the interactions within chromosome arms, between arms, and between the chromosomes, and exhibits distinct substructure in the form of an intense diagonal (**Figure 1A**). The intense diagonal indicates that the majorities of interacting reads are a short distance within the 200 kb (**Figure 1A**; **Supplementary Figure S1**). Further, we observed fewer interacting reads between the centromere and the rest of the genome in NC, HT, and SA conditions. To have a closer inspection of chromosomal regions that interact with each other, we plotted a normalized interaction matrix for individual chromosomes. It showed many blocks enriched in interacting reads over a long distance within and between the arms of the same chromosome (**Figure 1B**; **Supplementary Figure S2**). These enriched blocks are found across all chromosomes and in all conditions with variations in positioning and intensity of blocks at different conditions (**Supplementary Figure S2**).

We plotted the correlation matrix to identify how each locus interacts with all other loci on a chromosome. The correlation heat map showed two distinct types of compartments within a chromosome; one compartment showed correlated regions (red) and the other represented non-correlated regions (blue) (**Figure 1C**; **Supplementary Figure S3**). This correlation plot indicates there are a preferential enrichment and depletion of interacting reads among many chromatin loci. Strong correlation was found along the diagonal similar to the interaction matrices showing enrichment of reads in the neighborhood. The correlation matrix showed less correlation among the centromere and with the rest of the genome indicating that interacting reads between the centromere and another part of the genome are scanty (**Figure 1C**; **Supplementary Figure S3**).

We identified potential statistically significant (p-value ≤ 0.05) chromatin interactions at a resolution of 1 kb. Our analysis revealed a total of 3635, 5320, and 3309 statistically significant interactions for NC, HT, and SA libraries, respectively (**Supplementary Table S1**). Based on the positioning of these interacting loci on a chromosome, these interactions were designated as *cis* (intra-chromosomal, both interacting loci found on the same chromosome including interactions between homologous chromosomes) or *trans* (inter-chromosomal, both interacting loci located on a different chromosome) interactions. In NC out of 3635 total significant interactions, 2698 were *cis* interactions and 937 were *trans* interactions. For HT and SA treated, out of a total 5320 and 3309 significant interactions 3731 and 2288 were assigned as *cis*, and 1589 and 1021 as *trans,* respectively (**Supplementary Table S1** and **Supplementary File S3**). For the *cis* interactions, the median distances between interactions were 20,118, 24,392, and 25,770 bps, for NC, HT, and SA conditions, respectively, indicating the captured interactions were indeed both short and long range.

Further, the Circos plot of intra- and inter-chromosomal contacts indicates that the interactions were enriched in the centromeric regions for all five chromosomes (**Figure 1D**; **Supplementary Figure S4**). In general, the visual pattern of the chromosomal interaction profile of NC did not change on HT and SA treatment although the frequency of interactions between the chromatin loci may vary due to stress conditions (**Figure 1D**; **Supplementary Figure S4**). We found that telomere preferentially interacts with the other telomere and centromere of the same or different chromosomes (**Figure 1D**; **Supplementary Figure S4**).

We also identified that the number of intra-chromosomal contact probability decreases as a function of genomic distance in

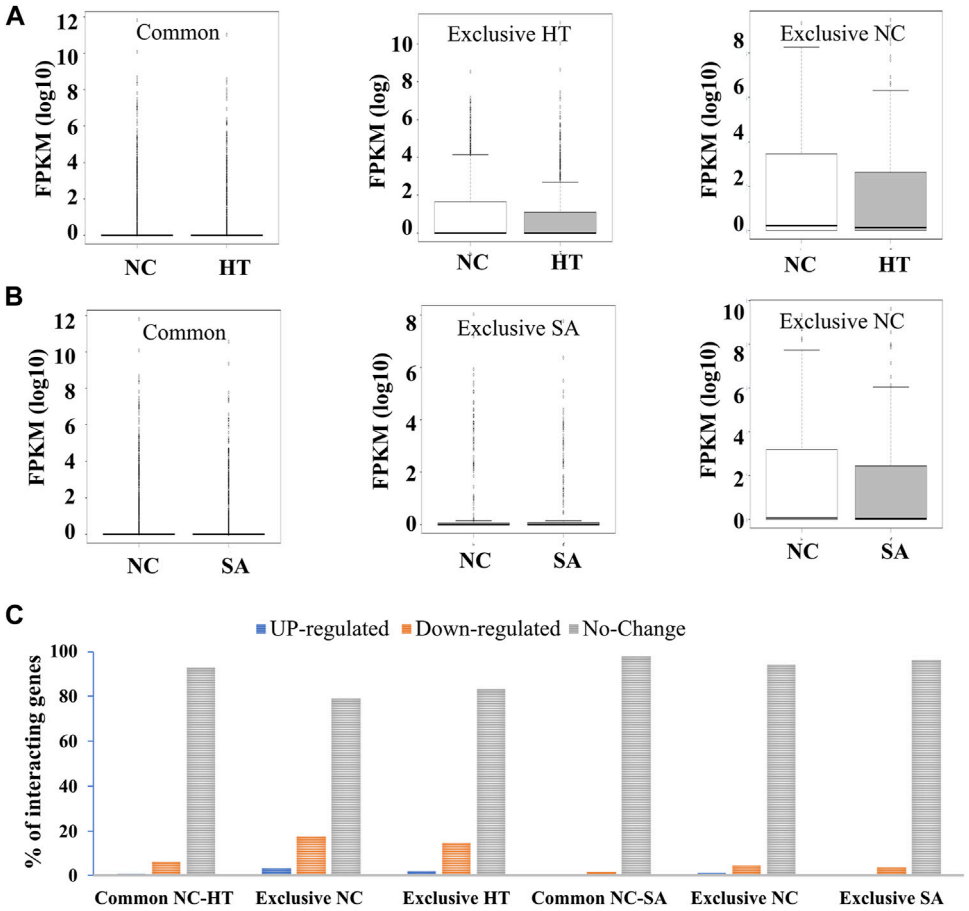


FIGURE 2 | Expression profiling of common and uniquely interacting genes in control and stress conditions. Enrichment and depletion of interaction leads to common and unique interacting genes in HT as well as in SA libraries in comparison to the NC. These common and uniquely interacting genes in **(A)** NC vs. HT and **(B)** NC vs. SA do not show any significant change in the expression profile in control and treated samples (*p*-value >0.005) indicating that these interactions may not directly regulate the expression of interacting genes. **(C)** Bar chart showing the percentage distribution of up-regulated, down-regulated, and unchanged expression of interacting genes in the common and exclusive interacting regions.

the base pair along the linear chromosome (Dekker et al., 2002; Lieberman-Aiden et al., 2009), and it was similar in all three conditions (**Figure 1E**; **Supplementary Figure S4**). Next, we identified the common interaction through overlapping interacting regions in all three conditions. A total of 2531, 2413, and 2422 interactions were shared between NC and HT, between NC and SA, and between HT and SA, respectively. On comparison of all three conditions, 2059 interactions were common, 750, 2426, and 533 interactions were exclusive to NC, HT, and SA conditions, respectively (**Supplementary Figure S5**). The result indicates that the stress conditions affect the dynamics of chromatin interactions.

Impact of Chromatin Interactome on Gene Expression

Our analysis revealed common and exclusive interactions among all the interactions identified in three conditions, namely, NC, SA, and HT (**Supplementary Figure S5**). The intriguing question was whether these changes in interactions affect the expression of genes involved in those interactions. We thus compared the expression profiles of common and uniquely interacting genes involved in the interactions in untreated (NC) and treated (HT and SA) conditions. We found no significant difference in the expression profile of common and uniquely interacting genes in NC and SA and HT conditions (**Figures 2A,B**). Further, we analyzed the percentage distribution of up-regulated, down-regulated, and unchanged expression of interacting genes in the common and exclusive interacting genes (untreated vs. treated). This revealed that most of the genes involved in the interactions showed no change in the expression profile (**Figure 2C**). These results indicate that *Arabidopsis* responds to different treatment (HT and SA) by either enrichment or by depletion of chromatin contacts, but this change in contacts may not lead to the uniform change in global expression profile of interacting genes. Since we did not observe any significant difference in the expression pattern of common and uniquely interacting genes at high resolution (1 kb), we identify interacting regions at lower resolution (200 kb) to explore the difference in the expression profile of larger interacting blocks. We identified

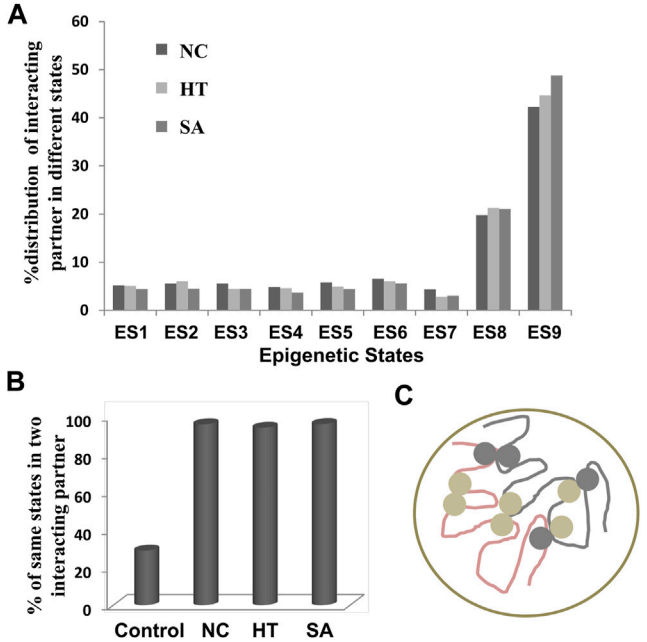


FIGURE 3 | Distribution of interacting regions into various epigenetic states. **(A)** Mapping of interacting regions on different epigenetic states as defined previously (Sequeira-Mendes et al., 2014). Interacting regions were highly enriched in state 8 and 9 which are the marks for heterochromatic regions and the least represented in state 7 which is exclusively associated with the intragenic regions. **(B)** The interactions frequency of two interacting regions among the NC, HT, and SA libraries sharing the same ES is significantly high over the control sets of interacting regions (p-value <0.00001) indicating that interactions were firmly associated with the epigenetic states of interacting regions. **(C)** Cartoon representing the interactions among the same ES is more frequent than the different ES (light and dark grey circles represent the different ES).

399 interaction blocks shared between NC and SA, and 58 and 34 interaction blocks unique to NC and SA, respectively. Similarly, 440 interaction blocks were common in NC and HT, and 17 and 131 were unique to NC and HT, respectively. We next calculated the cumulative expression of all genes in common and unique blocks in treated and untreated conditions. However, even at a lower resolution of 200 kb, we did not observe any significant change in the expression of genes involved in the common interaction or unique interactions in different conditions (**Supplementary Figure S6**). Our results thus indicate that enrichment or depletion of chromatin contacts at lower or higher resolution does not directly influence the global gene expression. We hypothesize possibly that change in the contacts and thus gene expression may be limited to a few specific cells, and thus the analysis of global gene expression in entire seedlings does not reveal any change in the expression.

Enrichment of Heterochromatin-Related Epigenetic Signature in Interacting Regions

Sequeira-Mendes and colleagues defined *A. thaliana* genome into nine ES based on various epigenetic marks (histone variants, histone marks, and CG methylation) (Sequeira-Mendes et al., 2014). Each ES has sets of enriched epigenetic marks and a few characteristic features. We mapped the coordinates of identified interacting regions to the ES to classify the identified interactions into various ES. The ES one to seven represent poorly in interactions, always less than 10% in all the conditions (**Figure 3A**). The distribution of interactions into various ES in all three conditions showed an almost similar pattern. We found significant enrichment of interacting regions in the ES8 (more than 20%) and 9 (more than 40%), these ES, marks for the heterochromatic region of the genome and enriched in epigenetic marks such as mCG, H3K9me2, and H3K27me1 (Sequeira-Mendes et al., 2014). The ES9 is almost twice enriched than ES8 in all three conditions in our analysis, although both represent the heterochromatic region but differ in the genomic position (**Figure 3A**). ES8 preferentially co-localizes with intergenic regions (AT-rich) while ES9 corresponds to heterochromatic pericentromeric regions (GC-rich). Both ES8 and 9 are highly enriched with the transposable element (TE); however, the enrichment of TE is comparatively higher in ES9 compared to ES8. The distribution of interacting elements in epigenetic ES8 and 9 indicates that identified interactions are mainly represented in the heterochromatin region. Since our chromatin interaction data over-represent TEs, we also analyzed a separate set of chromatin interactions excluding TEs. We identified that the distribution pattern of interacting regions without TEs into various ES, are more or less similar; however, the overall percentage of interacting regions that fall into ES8 and 9 was decreased but still higher than other ES (**Supplementary Figure S7A**). We identified that TEs involved in interactions showed more than 95% enrichment in ES8 and 9 (**Supplementary Figure S7B**). So, the above results indicate that not only the interacting TE but other interacting regions were also highly enriched in the ES8 and 9.

Next, we examined whether there are any preferential interactions among these ES. We identified the ES of both the interacting partners in NC, HT, and SA conditions and compared them with the randomly generated interactions used as a control. We found that interacting partners are always likely to be in the same ES which is always more than 90% instances in all the three conditions analyzed (**Figures 3B,C**). This enrichment of the same ES in the interacting partners was significantly higher (p-value < 0.00001) than the control set selected which showed merely 30% enrichment of the same ES (**Figure 3B**). The result indicates that chromatin interactions govern the folding principle of chromatin by bringing together similar epigenetic marks.

Conserved Motifs in the Interacting Regions Impart in Suppression of Genes Expression

We used MEME (Bailey et al., 2009) to identify the top 10 conserved motifs in those interacting sequences in which at least one partner was a protein-coding gene, in all three conditions (**Supplementary Figure S8**). Predicted motifs were further annotated for a *cis*-regulatory element with STAMP (Mahony and Benos, 2007) using the AGRIS database (The *Arabidopsis* Gene Regulatory Information Server, http://

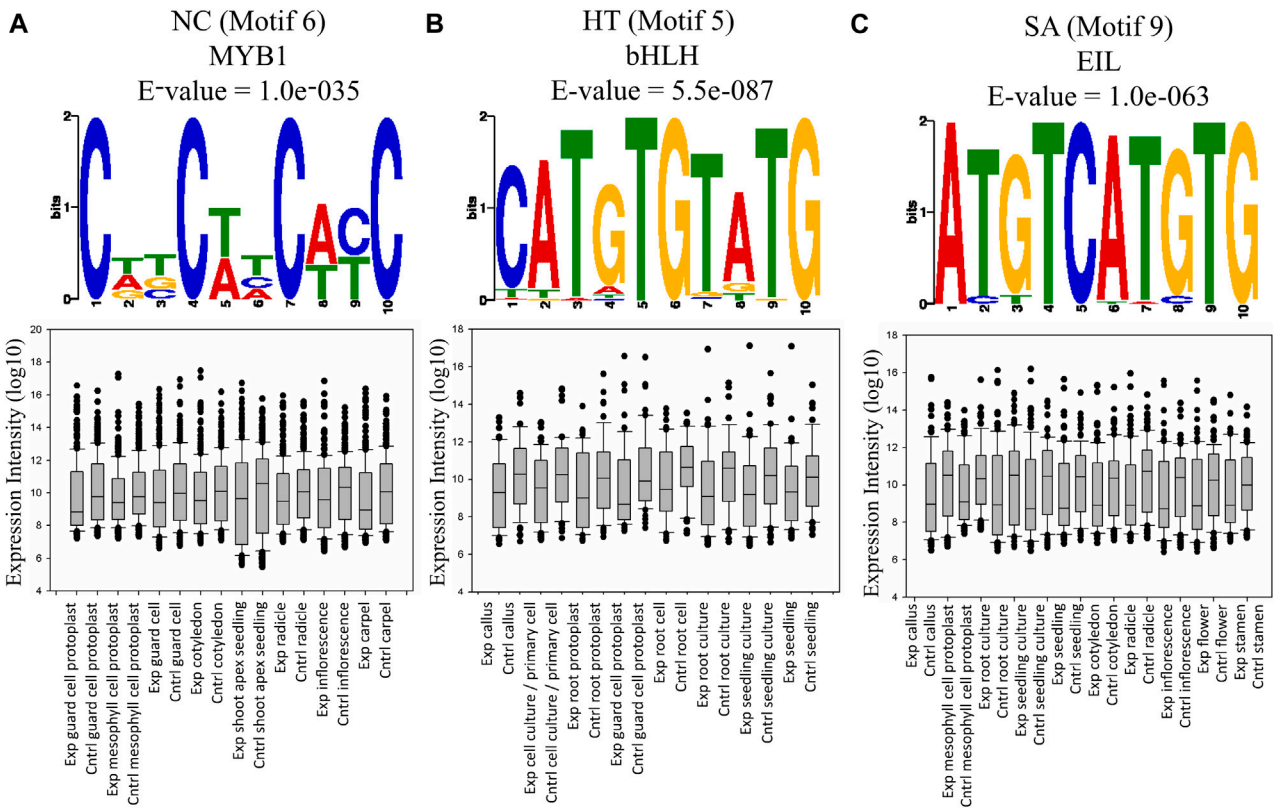


FIGURE 4 | The expression profile of interacting protein-coding genes having the binding site for known TFs. In this figure, some of the conditions represented are showing a statistically significant difference over the control (*t*-test *p*-value < 0.05). Interacting genes containing motifs for different TFs showed significantly lower expression in different tissues or plant parts. **(A)** Motif 6 of NC. **(B)** Motif 5 of HT. **(C)** Motif 9 of SA.

Arabidopsis.med.ohio-state.edu/). The motif containing AAGCTT was conserved in the top two positions among all the conditions (**Supplementary Figure S9**). The conservation of AAGCTT is expected since it is a HindIII site that is used for preparing the Hi-C libraries and thus it further validates the quality of our Hi-C library. Predicted motifs were annotated into 9 *cis*-regulatory elements in NC, HT, and SA libraries, respectively. Among these four regulatory elements, MYB1, ERE, SORLIP5, and LS7 were exclusively found in NC, while EIL1 and OBP-1_4_5 were exclusively present in SA, and LFY and AG_v3 were found in the HT library (**Figure 4**; **Supplementary Figure S8**). RAV1 element which provides the binding site for the TF Related-to-ABI3/VP1 (RAV) was identified in the NC and HT conditions, while the LFY and PRHA *cis*-regulatory elements have the binding site for Orphan and Homeobox TFs were found in HT and SA treated libraries. The very high conservation of some of the known TF binding sites in the interacting regions indicates the probable role of these elements in co-regulation of interacting genes. To explore this possibility, Motif RAV1-A and MYB1 for NC, Motif AtMYC2, PRHA, and LFY in HT and motif PRHA, and AtMYC2 and EIL1 in SA were selected further since these motifs are known to interact with known TFs (**Supplementary Figure S9**). We identified a set of genes interacting with each motif and also selected a random set of non-interacting genes as a control. The expression profiles of the interacting and non-interacting genes were retrieved from 69 cell and tissue conditions using the GENVESTIGATOR database. We identified for motif 4 and 6, there were 3 and 23 distinct expression profiles in NC, respectively, similarly for motif 5 and 8, 46 and 1 profiles in HT and motif 6, 7, and 9 in SA, 12, 58, and 37 profiles were the expressions of interacting genes was lower (p-value < 0.05) than the control dataset (**Supplementary Figure S9**). Thus, irrespective of the motif they are interacting, interacting genes showed either lower expression or no difference than non-interacting genes in specific cell and tissue conditions. Thus, results indicate that either these conserved motifs bind to TFs which are largely suppressor or TFs may activate genes specifically in restricted cell type or temporally or conditionally which cannot be identified by global expression profiling in set parameters.

Interacting Regions Exclude Quantitative Trait Loci Associated with Phenotypic Diversity

Further, we ask a question whether the chromatin interacting region has a functional role in controlling phenotypic diversity in

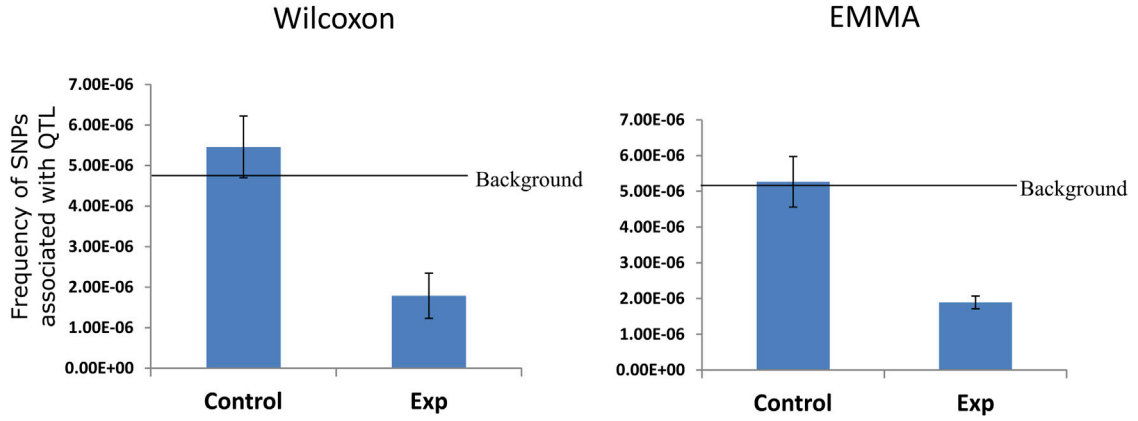


FIGURE 5 | Distribution of SNP associated with QTL in interacting regions. Mapping of publicly available SNP (Atwell et al., 2010) associated with quantitative trait on interacting and non-interacting regions (control), showing more than two times depletion of GWAS hit in the chromatin interacting region over the non-interacting regions. The selected control region is not biased since genome-wide GWAS SNP frequency (background) is similar in the background and control region.

A. thaliana. We made use of GWAS data (Atwell et al., 2010) which revealed using EMMA and Wilcoxon test 615 SNPs and 567 unique SNPs respectively associated with the quantitative trait loci for several phenotypic traits in natural accessions of *A. thaliana*. We calculated the frequency of association of SNPs in interacting genome; we also consider the same length of a non-interacting genome as a control and also the entire *A. thaliana* genome to calculate background frequency of associated SNP. The frequency of associated SNP in the non-interacting genome is marginally higher than the background frequency (**Figure 5**). However, interestingly the frequency of associated SNPs identified using EMMA and Wilcoxon test was almost two times lower in the interacting regions of *A. thaliana* genome (**Figure 5**). The result indicates that the SNPs associated with phenotypic diversity in the natural population of *A. thaliana* are preferentially excluded from the portion of the genome involved in chromatin interactions.

DISCUSSION

The establishment of Hi-C protocol (Lieberman-Aiden et al., 2009) leads to extensive work resulting in the identification of genome-wide chromatin interaction networks in human (Sandhu et al., 2012; Jin et al., 2013), *Drosophila* (Sexton et al., 2012), yeast (Duan et al., 2010), and plants (Moissiard et al., 2012; Grob et al., 2013; Grob et al., 2014; Feng et al., 2014; Wang et al., 2015; Liu et al., 2016). In the present work for the first time, we analyzed the effect of either hormonal treatment or abiotic stress on chromatin interaction networks.

As previously reported, we also observed that the nuclear architecture of *A. thaliana* is significantly different from others as its genome is not partitioned into the larger interactive topological domain (Moissiard et al., 2012; Feng et al., 2014). This spectacular feature could be explained due to the lack of CTCF in *Arabidopsis* genome (Heger et al., 2012). As previously reported, we have also identified that *A. thaliana* genome can be partitioned into two broad groups based on the correlation of interacting reads (**Figure 1C**) (Lieberman-Aiden et al., 2009; Sexton et al., 2012; Yadav et al., 2021). The correlated region has a higher frequency of interacting reads than the non-correlated region of the genome (Grob et al., 2014). Further, our analysis of correlation heat map showing the visual transition of chromatin compartment in HT condition (**Figure 1C**; **Supplementary Figure S3**). Similar observations have recently been reported for *Arabidopsis* and rice (Sun et al., 2020; Liang et al., 2021), suggesting that the compartment transition might be conserved among different species. We have uncovered several interactions in the heterochromatic and euchromatic regions that established physical communication with the different regions of the genome, which will provide a new mechanistic way of gene regulation in plants (**Figure 1**). Our interaction data suggest that the interactions along the chromosome are more frequent (~69–74%) than those between the chromosomes (~26–31%) (**Supplementary Table S1**) and interaction frequency is higher for interacting pairs of loci that lie close together in a linear chromosome (**Figure 1E**), which is the general feature of spatial chromatin organization (Lieberman-Aiden et al., 2009; Sanyal et al., 2012; Feng et al., 2014; Grob et al., 2014). We observed that heat stress induces the genome-wide long-range chromatin interactions (>20 Kb, ~55 vs. ~50% in HT vs. NC) but decreases short-range chromatin interactions (<20 kb, ~45 vs. ~50% in HT vs. NC; Supplementary File 3) which is contrasting to the previous observation in *Arabidopsis* (Sun et al., 2020). However, in rice it is reported that the short-distance interactions were decreased during heat stress (Liang et al., 2021). This variation could be technical or because of different development stages of plant material and heat stress condition used in both studies.

We found an increase in *trans* interactions (~29.9 vs. ~25.8% in HT vs. NC) after heat stress similar to a previous study in rice (Liang et al., 2021). Further, we also identified that the frequency of inter-chromosomal interaction is quite high in the pericentromeric regions (**Figure 1D**) which is a general feature

of inter-chromosomal contact (Lieberman-Aiden et al., 2009; Feng et al., 2014; Grob et al., 2014; Wang et al., 2015; Nützmann et al., 2020). We assume the enrichment of inter-chromosomal interactions near the centromeric regions facilitated the heterochromatin-mediated chromatin fiber concretion at the centromere, providing structural constraints vital for the genome organization (Mizuguchi et al., 2014). Since we found substantial inter-chromosomal interactions in the pericentromeric regions which are generally enriched with crucial heterochromatin marks, it is not surprising that we observed enrichment of heterochromatic regions in the interactions. These results are in line with previously reported studies in *A. thaliana* (Moissiard et al., 2012; Feng et al., 2014; Grob et al., 2014; Yadav et al., 2021). Interactions between the heterochromatic regions and genes possibly assist the cell or organ-specific clustering of interacting genes hence regulating the expression in specific tissue type (Fransz et al., 2002; Nützmann et al., 2020; Yadav et al., 2021). Recently, Nutzmann and colleagues demonstrated that dynamic clustering of biosynthetic gene cluster with chromatic arm (euchromatic region) and pericentromeric (heterochromatic region) in root and leaf, respectively, define the transcriptional active and repressed states (Nützmann et al., 2020). Furthermore, intra- and inter-chromosomal contacts between the pericentromeric regions of the different chromosomes are high which is enriched with the TE, thus these interactions may be playing an important role in keeping transposons silent and thus maintaining the integrity of the genome (Grob et al., 2014). We also identified interactions among the telomere of different chromosomes and telomere with the centromere as reported earlier (Moissiard et al., 2012; Feng et al., 2014; Grob et al., 2014) and which is in concordance with previous DNA FISH assay (Fransz et al., 2002; Schubert et al., 2012).

Uniquely, we were interested in addressing whether a stress condition alters the number of significantly identified interactions and changes in the interactions influence the expression of genes associated with it. Our results revealed that a stress condition either biotic (as mimicked by SA) or abiotic does result in the change in the chromatin interactions (**Supplementary Figure S5**). The stress condition leads to establishment of 896 and 2789 new interactions and loss of 1222 and 1104 interactions in SA and HT treatment, respectively, as compared to native conditions (**Supplementary Figure S5**). The chromosomal reorganization in response to heat stress is consistent with the previous studies in *Arabidopsis* and rice (Sun et al., 2020; Liang et al., 2021). Our data show that the chromosomal contact increased in response to heat stress compared with the control (**Supplementary Table S1**). Similar observations have been reported for *Arabidopsis* seedling demonstrating that in general the chromosomal contact is enhanced between different regions along the chromosomes (Sun et al., 2020).

Our results at 1 Kb or 200 Kbs revealed that whether the interactions are enriched or depleted due to stress conditions at any loci does not affect the global gene expression associated with interactions (**Figure 2**; **Supplementary Figure S6**). This result brings us to an important question that if not for the accommodating change in the expression of genes associated with interactions, then why was there so much change in the interactions profile after stress treatment? The possibility could be the chromosomal organization is prone to random variation which is unlikely caused by essentially biological processes hence cannot directly correlate with the transcriptional state of the cell (Nagano et al., 2013; Grob et al., 2014). It should also be noted that we are likely to underestimate the impact of interactions on gene expression because captured interactions are not happening in all the cells and tissues and might be representing a particular type of cell and tissue. Since we analyzed the expression in the seedlings this does not directly reflect those finer changes. However, it does not remain a limitation when working with the animal cell lines or cell/tissue specific analysis where the expression is directly correlated with the interactions (Chepelev et al., 2012; Dixon et al., 2012; Jin et al., 2013; Yadav et al., 2021). Thus, further detailed studies are needed to address such questions as to why chromatin interactions change after encountering stress conditions.

Mapping of interactions on previously identified ES (Sequeira-Mendes et al., 2014) in *A. thaliana* revealed that the interactions are highly enriched in repressive heterochromatin marks (H3K9me2, H3K27me1) is also in concordance with the recent report (Liu et al., 2016). It is also noteworthy that as reported previously (Feng et al., 2014) that H3K9me2 and H3K27me1 marks are highly enriched in interacting regions, our analysis also showed enrichment of ES8 and ES9 (**Figure 3A**) in the interacting *A. thaliana* genome which is distinctly enriched with H3K9me2 and H3K27me1 (Sequeira-Mendes et al., 2014). This further confirms that genes involved in the interactions are enriched with heterochromatin. Further, the ES of the interacting partners is likely to be the same (**Figure 3B**) which further confirms the folding principle of chromatin.

The possibility of interactions involving chromatin loops and regulating gene expression by either activation or suppression leads us to identify conserved *cis*-regulatory elements in the interacting region. The interacting regions showed significant conservation of binding sites of some of the characterized TFs (**Supplementary Figure S8**). These TFs play a major role in various plant growth and developmental processes, like bHLH in the regulation of a multiplicity of transcriptional programs (Toledo-Ortiz et al., 2003), RAV-1 was known to be a negative regulator of growth and its level repressed by hormones involved in abiotic stress (Fu et al., 2014). MYB TFs family participates in the regulation of various biological processes like responses to abiotic and biotic stress, metabolism, development, and differentiation (Ambawat et al., 2013). EIL family is involved in ethylene signaling in plants (Solano et al., 1998). This intriguing possibility is that these motifs might bind to suppressor or activator conditionally and regulate their expression by looping. It is also noteworthy to mention that the identified interacting genes are enriched from the heterochromatic and pericentromeric region and the control sets of genes are selected randomly and from the gene pools so this lower expression in different tissue could be technical instead of biological.

The previous studies (Moissiard et al., 2012; Feng et al., 2014; Grob et al., 2014; Liu et al., 2016) pointed out a folding principle that partitioned *A. thaliana* genome into highly interacting silenced heterochromatin region and less interacting active euchromatic region. We asked whether the interacting heterochromatin region is at all responsible for functional phenotypic diversity in *A. thaliana* occurs naturally in ecotypes. Interestingly, we observed that always QTLs associated with phenotypic diversity in a natural *A. thaliana* population are selectively excluded from the portion of the genome involved in the chromatin interactions (**Figure 5**). This is in a way not surprising considering most of the euchromatin and actively transcribing portion of the genome is excluded from the interaction. This raised very important questions about the evolution, functionality, and importance of these chromatin interacting regions that need to be addressed in the future.

DATA AVAILABILITY STATEMENT

Sequence data generated for this study have been submitted to the NCBI under the BioProject ID: PRJNA317803 (Hi-C) and PRJNA317804 (RNA-seq).

AUTHOR CONTRIBUTIONS

S.V.S. and Vi.K.Y. conceptualized and designed the experiments. Vi.K.Y. performed the experiments. Vi.K.Y., S.S., A.Y., and R.S. performed the bioinformatic data analysis. B.S., Vr.K.Y., N.A., V.K.T., and V.K. helped in the standardization of Hi-C. Vi.K.Y. and S.V.S. analyzed and interpreted the data and wrote the manuscript. S.K.J. edited the manuscript and helped in data analysis. All authors have read and agreed to the published version of the manuscript.

FUNDING

This work was supported by a grant of Council of Scientific and Industrial Research, India (EMPOWER Project Code, OLP-079). V.K.Y. thanks University Grant Commission (UGC), New Delhi for providing the fellowship. Institution manuscript number allotted to this manuscript is CSIR-NBRI_MS/2021/11/04.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.799805/full#supplementary-material

Supplementary file 1: Supplementary Table S1. | Identified significant interactions at 1Kb resolution and its classification based on the location of interacting regions on chromosomes.

Supplementary file 2: Supplementary Figures S1-S9. S1. | Genome-wide normalized interaction matrix of A. thaliana for HT and SA library. **S2**. Chromosome-wise contact matrices. **S3**. Genome-wide correlation interaction matrix. **S4**. Visualization of statistically significant interactions. **S5**. Venn diagram showing the common and unique interactions among different libraries. **S6**. Effect of interactions on gene expression at lower resolution (200kb). **S7**. Distribution of interacting regions into various ES. **S8**. Motifs identification and annotation. **S9**. Comparison of expression profile of interacting protein coding genes having the binding site for know TFs as identified through AGRIS annotation of motifs.

Supplementary file 3. | Identification of significant interactions at 1 Kb resolution using HOMER.

Supplementary file 4. | Hi-C libraries preparation and sequencing.

Supplementary file 5. | Expression data of interacting genes in NC, HT and SA condition RNA-seq.

REFERENCES

Ambawat, S., Sharma, P., Yadav, N. R., and Yadav, R. C. (2013). MYB Transcription Factor Genes as Regulators for Plant Responses: An Overview. *Physiol. Mol. Biol. Plants* 19, 307–321. doi:10.1007/s12298-013-0179-1

Atwell, S., Huang, Y. S., Vilhjálmsson, B. J., Willems, G., Horton, M., Li, Y., et al. (2010). Genome-wide Association Study of 107 Phenotypes in *Arabidopsis thaliana* Inbred Lines. *Nature* 465, 627–631. doi:10.1038/nature08800

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

Bi, X., Cheng, Y.-J., Hu, B., Ma, X., Wu, R., Wang, J.-W., et al. (2017). Nonrandom Domain Organization of the Arabidopsis Genome at the Nuclear Periphery. *Genome Res.* 27, 1162–1173. doi:10.1101/gr.215186.116

Bowler, C., Benvenuto, G., Laflamme, P., Molino, D., Probst, A. V., Tariq, M., et al. (2004). Chromatin Techniques for Plant Cells. *Plant J.* 39, 776–789. doi:10.1111/j.1365-313X.2004.02169.x

Chepelev, I., Wei, G., Wangsa, D., Tang, Q., and Zhao, K. (2012). Characterization of Genome-wide Enhancer-Promoter Interactions Reveals Co-expression of Interacting Genes and Modes of Higher Order Chromatin Organization. *Cell Res.* 22, 490–503. doi:10.1038/cr.2012.15

Crevillén, P., Sonmez, C., Wu, Z., and Dean, C. (2012). A Gene Loop Containing the floral Repressor FLC Is Disrupted in the Early Phase of Vernalization. *EMBO J.* 32, 140–148. doi:10.1038/emboj.2012.324

Dean, A. (2011). In the Loop: Long Range Chromatin Interactions and Gene Regulation. *Brief. Funct. Genomics* 10, 3–10. doi:10.1093/bfgp/elq033

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing Chromosome Conformation. *Science* 295, 1306–1311. doi:10.1126/science.1067799

Deng, W., Lee, J., Wang, H., Miller, J., Reik, A., Gregory, P. D., et al. (2012). Controlling Long-Range Genomic Interactions at a Native Locus by Targeted Tethering of a Looping Factor. *Cell* 149, 1233–1244. doi:10.1016/j.cell.2012.03.051

Dittmer, T. A., Stacey, N. J., Sugimoto-Shirasu, K., and Richards, E. J. (2007). LITTLE NUCLEIGenes Affecting Nuclear Morphology inArabidopsis Thaliana. *Plant Cell* 19, 2793–2803. doi:10.1105/tpc.107.053231

Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., et al. (2012). Topological Domains in Mammalian Genomes Identified by Analysis of Chromatin Interactions. *Nature* 485, 376–380. doi:10.1038/nature11082

Dostie, J., Richmond, T. A., Arnaout, R. A., Selzer, R. R., Lee, W. L., Honan, T. A., et al. (2006). Chromosome Conformation Capture Carbon Copy (5C): a Massively Parallel Solution for Mapping Interactions between Genomic Elements. *Genome Res.* 16, 1299–1309. doi:10.1101/gr.5571506

Duan, Z., Andronescu, M., Schutz, K., McIlwain, S., Kim, Y. J., Lee, C., et al. (2010). A Three-Dimensional Model of the Yeast Genome. *Nature* 465, 363–367. doi:10.1038/nature08973

Eijsbouts, C. Q., Burren, O. S., Newcombe, P. J., and Wallace, C. (2019). Fine Mapping Chromatin Contacts in Capture Hi-C Data. *BMC Genomics* 20, 77. doi:10.1186/s12864-018-5314-5

Feng, S., Cokus, S. J., Schubert, V., Zhai, J., Pellegrini, M., and Jacobsen, S. E. (2014). Genome-wide Hi-C Analyses in Wild-type and Mutants Reveal High-Resolution Chromatin Interactions in Arabidopsis. *Mol. Cel.* 55, 694–707. doi:10.1016/j.molcel.2014.07.008

Fransz, P., De Jong, J. H., Lysak, M., Castiglione, M. R., and Schubert, I. (2002). Interphase Chromosomes in Arabidopsis Are Organized as Well Defined Chromocenters from Which Euchromatin Loops Emanate. *Proc. Natl. Acad. Sci.* 99, 14584–14589. doi:10.1073/pnas.212325299

Fu, M., Kang, H. K., Son, S.-H., Kim, S.-K., and Nam, K. H. (2014). A Subset of Arabidopsis RAV Transcription Factors Modulates Drought and Salt Stress Responses Independent of ABA. *Plant Cel. Physiol.* 55, 1892–1904. doi:10.1093/pcp/pcu118

Fullwood, M. J., and Ruan, Y. (2009). ChIP-based Methods for the Identification of Long-Range Chromatin Interactions. *J. Cel. Biochem.* 107, 30–39. doi:10.1002/jcb.22116

Furey, T. S. (2012). ChIP-seq and beyond: New and Improved Methodologies to Detect and Characterize Protein-DNA Interactions. *Nat. Rev. Genet.* 13, 840–852. doi:10.1038/nrg3306

Grob, S., Schmid, M. W., and Grossniklaus, U. (2014). Hi-C Analysis in Arabidopsis Identifies the KNOT, a Structure with Similarities to the Flamenco Locus of Drosophila. *Mol. Cel.* 55, 678–693. doi:10.1016/j.molcel.2014.07.009

Grob, S., Schmid, M. W., Luedtke, N. W., Wicker, T., and Grossniklaus, U. (2013). Characterization of Chromosomal Architecture in Arabidopsis by Chromosome Conformation Capture. *Genome Biol.* 14, R129. doi:10.1186/gb-2013-14-11-r129

Heger, P., Marin, B., Bartkuhn, M., Schierenberg, E., and Wiehe, T. (2012). The Chromatin Insulator CTCF and the Emergence of Metazoan Diversity. *Proc. Natl. Acad. Sci.* 109, 17507–17512. doi:10.1073/pnas.1111941109

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., et al. (2010). Simple Combinations of Lineage-Determining Transcription Factors Prime Cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol. Cel.* 38, 576–589. doi:10.1016/j.molcel.2010.05.004

Jin, F., Li, Y., Dixon, J. R., Selvaraj, S., Ye, Z., Lee, A. Y., et al. (2013). A High-Resolution Map of the Three-Dimensional Chromatin Interactome in Human Cells. *Nature* 503, 290–294. doi:10.1038/nature12644

Kaul, S., Koo, H. L., Jenkins, J., Rizzo, M., Rooney, T., Tallon, L. J., et al. (2000). Analysis of the Genome Sequence of the Flowering Plant *Arabidopsis thaliana*. *Nature* 408, 796–815. doi:10.1038/35048692

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

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

Li, G., Ruan, X., Auerbach, R. K., Sandhu, K. S., Zheng, M., Wang, P., et al. (2012). Extensive Promoter-Centered Chromatin Interactions Provide a Topological Basis for Transcription Regulation. *Cell* 148, 84–98. doi:10.1016/j.cell.2011.12.014

Li, L., Luo, H., Lim, D.-H., Han, L., Li, Y., Fu, X.-D., et al. (2021). Global Profiling of RNA-Chromatin Interactions Reveals Co-regulatory Gene Expression Networks in Arabidopsis. *Nat. Plants* 7, 1364–1378. doi:10.1038/s41477-021-01004-x

Liang, Z., Zhang, Q., Ji, C., Hu, G., Zhang, P., Wang, Y., et al. (2021). Reorganization of the 3D Chromatin Architecture of rice Genomes during Heat Stress. *BMC Biol.* 19, 53. doi:10.1186/s12915-021-00996-4

Lieberman-Aiden, E., Van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., et al. (2009). Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science* 326, 289–293. doi:10.1126/science.1181369

Ling, J. Q., Li, T., Hu, J. F., Vu, T. H., Chen, H. L., Qiu, X. W., et al. (2006). CTCF Mediates Interchromosomal Colocalization between Igf2/H19 and Wsb1/Nf1. *Science* 312, 269–272. doi:10.1126/science.1123191

Liu, C., Wang, C., Wang, G., Becker, C., Zaidem, M., and Weigel, D. (2016). Genome-wide Analysis of Chromatin Packing in *Arabidopsis thaliana* at Single-Gene Resolution. *Genome Res.* 26, 1057–1068. doi:10.1101/gr.204032.116

Louwers, M., Bader, R., Haring, M., Van Driel, R., De Laat, W., and Stam, M. (2009a). Tissue- and Expression Level-specific Chromatin Looping at Maizeb1Epialleles. *Plant Cell* 21, 832–842. doi:10.1105/tpc.108.064329

Louwers, M., Splinter, E., van Driel, R., de Laat, W., and Stam, M. (2009b). Studying Physical Chromatin Interactions in Plants Using Chromosome Conformation Capture (3C). *Nat. Protoc.* 4, 1216–1229. doi:10.1038/nprot.2009.113

Ma, W., Ay, F., Lee, C., Gulsoy, G., Deng, X., Cook, S., et al. (2018). Using DNase Hi-C Techniques to Map Global and Local Three-Dimensional Genome Architecture at High Resolution. *Methods* 142, 59–73. doi:10.1016/j.ymeth.2018.01.014

Mahony, S., and Benos, P. V. (2007). STAMP: a Web Tool for Exploring DNA-Binding Motif Similarities. *Nucleic Acids Res.* 35, W253–W258. doi:10.1093/nar/gkm272

Mizuguchi, T., Fudenberg, G., Mehta, S., Belton, J.-M., Taneja, N., Folco, H. D., et al. (2014). Cohesin-dependent Globules and Heterochromatin Shape 3D Genome Architecture in *S. pombe*. *Nature* 516, 432–435. doi:10.1038/nature13833

Moissiard, G., Cokus, S. J., Cary, J., Feng, S., Billi, A. C., Stroud, H., et al. (2012). MORC Family ATPases Required for Heterochromatin Condensation and Gene Silencing. *Science* 336, 1448–1451. doi:10.1126/science.1221472

Nagano, T., Lubling, Y., Stevens, T. J., Schoenfelder, S., Yaffe, E., Dean, W., et al. (2013). Single-cell Hi-C Reveals Cell-To-Cell Variability in Chromosome Structure. *Nature* 502, 59–64. doi:10.1038/nature12593

Nützmann, H.-W., Doerr, D., Ramírez-Colmenero, A., Sotelo-Fonseca, J. E., Wegel, E., Di Stefano, M., et al. (2020). Active and Repressed Biosynthetic Gene Clusters Have Spatially Distinct Chromosome States. *Proc. Natl. Acad. Sci. USA* 117, 13800–13809. doi:10.1073/pnas.1920474117

Saeed, A. I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., et al. (2003). TM4: A Free, Open-Source System for Microarray Data Management and Analysis. *Biotechniques* 34, 374–378. doi:10.2144/03342mt01

Sandhu, K. S., Li, G., Poh, H. M., Quek, Y. L. K., Sia, Y. Y., Peh, S. Q., et al. (2012). Large-Scale Functional Organization of Long-Range Chromatin Interaction Networks. *Cel. Rep.* 2, 1207–1219. doi:10.1016/j.celrep.2012.09.022

Sanyal, A., Lajoie, B. R., Jain, G., and Dekker, J. (2012). The Long-Range Interaction Landscape of Gene Promoters. *Nature* 489, 109–113. doi:10.1038/nature11279

Schubert, V., Berr, A., and Meister, A. (2012). Interphase Chromatin Organisation in Arabidopsis Nuclei: Constraints versus Randomness. *Chromosoma* 121, 369–387. doi:10.1007/s00412-012-0367-8

Sequeira-Mendes, J., Aragüez, I., Peiró, R., Mendez-Giraldez, R., Zhang, X., Jacobsen, S. E., et al. (2014). The Functional Topography of the Arabidopsis Genome Is Organized in a Reduced Number of Linear Motifs of Chromatin States. *Plant Cell* 26, 2351–2366. doi:10.1105/tpc.114.124578

Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., et al. (2012). Three-dimensional Folding and Functional Organization Principles of the Drosophila Genome. *Cell* 148, 458–472. doi:10.1016/j.cell.2012.01.010

Simon, J. M., Giresi, P. G., Davis, I. J., and Lieb, J. D. (2012). Using Formaldehyde-Assisted Isolation of Regulatory Elements (FAIRE) to Isolate Active Regulatory DNA. *Nat. Protoc.* 7, 256–267. doi:10.1038/nprot.2011.444

Solano, R., Stepanova, A., Chao, Q., and Ecker, J. R. (1998). Nuclear Events in Ethylene Signaling: A Transcriptional cascade Mediated by ETHYLENE-INSENSITIVE3 and ETHYLENE-RESPONSE-FACTOR1. *Genes Dev.* 12, 3703–3714. doi:10.1101/gad.12.23.3703

Song, L., and Crawford, G. E. (2010). DNase-Seq: A High-Resolution Technique for Mapping Active Gene Regulatory Elements across the Genome from Mammalian Cells. *Cold Spring Harb. Protoc.* 2010, pdb.prot5384. doi:10.1101/pdb.prot5384

Sun, L., Jing, Y., Liu, X., Li, Q., Xue, Z., Cheng, Z., et al. (2020). Heat Stress-Induced Transposon Activation Correlates with 3D Chromatin Organization Rearrangement in Arabidopsis. *Nat. Commun.* 11, 1–13. doi:10.1038/s41467-020-15809-5

Swarbreck, D., Wilks, C., Lamesch, P., Berardini, T. Z., Garcia-Hernandez, M., Foerster, H., et al. (2007). The Arabidopsis Information Resource (TAIR): Gene

Structure and Function Annotation. *Nucleic Acids Res.* 36, D1009–D1014. doi:10.1093/nar/gkm965

Toledo-Ortiz, G., Huq, E., and Quail, P. H. (2003). The Arabidopsis Basic/Helix-Loop-Helix Transcription Factor Family[W]. *Plant Cell* 15, 1749–1770. doi:10.1105/tpc.013839

Trapnell, C., Pachter, L., and Salzberg, S. L. (2009). TopHat: Discovering Splice Junctions with RNA-Seq. *Bioinformatics* 25, 1105–1111. doi:10.1093/bioinformatics/btp120

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

van Berkum, N. L., Lieberman-Aiden, E., Williams, L., Imakaev, M., Gnirke, A., Mirny, L. A., et al. (2010). Hi-C: A Method to Study the Three-Dimensional Architecture of Genomes. *J. Visualized Experiments* 39, 1–7. doi:10.3791/1869

Wang, C., Liu, C., Roqueiro, D., Grimm, D., Schwab, R., Becker, C., et al. (2015). Genome-wide Analysis of Local Chromatin Packing in Arabidopsis Thaliana. *Genome Res.* 25, 246–256. doi:10.1101/gr.170332.113

Wang, J., Zhou, Y., Li, X., Meng, X., Fan, M., Chen, H., et al. (2016). Genome-Wide Analysis of the Distinct Types of Chromatin Interactions in Arabidopsis Thaliana. *Plant Cel. Physiol.* 58, pcw194. doi:10.1093/pcp/pcw194

Yadav, V. K., Santos-González, J., and Köhler, C. (2021). INT-Hi-C Reveals Distinct Chromatin Architecture in Endosperm and Leaf Tissues of Arabidopsis. *Nucleic Acids Res.* 49, 4371–4385. doi:10.1093/NAR/GKAB191

Yilmaz, A., Mejia-Guerra, M. K., Kurz, K., Liang, X., Welch, L., and Grotewold, E. (2011). AGRIS: the Arabidopsis Gene Regulatory Information Server, an Update. *Nucleic Acids Res.* 39, D1118–D1122. doi:10.1093/nar/gkq1120

Zhang, H., Zheng, R., Wang, Y., Zhang, Y., Hong, P., Fang, Y., et al. (2019). The Effects of Arabidopsis Genome Duplication on the Chromatin Organization and Transcriptional Regulation. *Nucleic Acids Res.* 47, 7857–7869. doi:10.1093/nar/gkz511

Zhao, Z., Tavoosidana, G., Sjölinder, M., Göndör, A., Mariano, P., Wang, S., et al. (2006). Circular Chromosome Conformation Capture (4C) Uncovers Extensive Networks of Epigenetically Regulated Intra- and Interchromosomal Interactions. *Nat. Genet.* 38, 1341–1347. doi:10.1038/ng1891

Zhu, B., Zhang, W., Zhang, T., Liu, B., and Jiang, J. (2015). Genome-wide Prediction and Validation of Intergenic Enhancers in Arabidopsis Using Open Chromatin Signatures. *Plant Cell* 27, 2415–2426. doi:10.1105/tpc.15.00537

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.

Copyright © 2022 Yadav, Singh, Yadav, Agarwal, Singh, Jalmi, Yadav, Tiwari, Kumar, Singh and Sawant. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genetics

REVIEW
published: 09 February 2022
doi: 10.3389/fgene.2022.832153

Genomic Selection: A Tool for Accelerating the Efficiency of Molecular Breeding for Development of Climate-Resilient Crops

Neeraj Budhlakoti[1†], Amar Kant Kushwaha[2†], Anil Rai[1], K K Chaturvedi[1], Anuj Kumar[1], Anjan Kumar Pradhan[3], Uttam Kumar[4], Rajeev Ranjan Kumar[1], Philomin Juliana[4], D C Mishra[1] and Sundeep Kumar[3]**

[1]ICAR- Indian Agricultural Statistics Research Institute, New Delhi, India, [2]ICAR- Central Institute for Subtropical Horticulture, Lucknow, India, [3]ICAR- National Bureau of Plant Genetic Resources, New Delhi, India, [4]Borlaug Institute for South Asia (BISA), Ludhiana, India

OPEN ACCESS

Edited by:
Vijay Gahlaut, Institute of Himalayan Bioresource Technology (CSIR), India

Reviewed by:
Aditya Pratap, Indian Institute of Pulses Research (ICAR), India
Upendra Kumar, Chaudhary Charan Singh Haryana Agricultural University, India

***Correspondence:**
D C Mishra
dwij.mishra@gmail.com
Sundeep Kumar
sundeep.kumar@icar.gov.in

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Genomics, a section of the journal Frontiers in Genetics

Received: *09 December 2021*
Accepted: *10 January 2022*
Published: *09 February 2022*

Citation:
Budhlakoti N, Kushwaha AK, Rai A, Chaturvedi KK, Kumar A, Pradhan AK, Kumar U, Kumar RR, Juliana P, Mishra DC and Kumar S (2022) Genomic Selection: A Tool for Accelerating the Efficiency of Molecular Breeding for Development of Climate-Resilient Crops. Front. Genet. 13:832153. doi: 10.3389/fgene.2022.832153

Since the inception of the theory and conceptual framework of genomic selection (GS), extensive research has been done on evaluating its efficiency for utilization in crop improvement. Though, the marker-assisted selection has proven its potential for improvement of qualitative traits controlled by one to few genes with large effects. Its role in improving quantitative traits controlled by several genes with small effects is limited. In this regard, GS that utilizes genomic-estimated breeding values of individuals obtained from genome-wide markers to choose candidates for the next breeding cycle is a powerful approach to improve quantitative traits. In the last two decades, GS has been widely adopted in animal breeding programs globally because of its potential to improve selection accuracy, minimize phenotyping, reduce cycle time, and increase genetic gains. In addition, given the promising initial evaluation outcomes of GS for the improvement of yield, biotic and abiotic stress tolerance, and quality in cereal crops like wheat, maize, and rice, prospects of integrating it in breeding crops are also being explored. Improved statistical models that leverage the genomic information to increase the prediction accuracies are critical for the effectiveness of GS-enabled breeding programs. Study on genetic architecture under drought and heat stress helps in developing production markers that can significantly accelerate the development of stress-resilient crop varieties through GS. This review focuses on the transition from traditional selection methods to GS, underlying statistical methods and tools used for this purpose, current status of GS studies in crop plants, and perspectives for its successful implementation in the development of climate-resilient crops.

Keywords: GS, climate change, STGS, MTGS, abiotic stress, biotic stress, GEBV, climate-resilient crops

INTRODUCTION

Sustainable food production is the utmost requirement for food and nutritional security. Based on reports, 821 million people are point below nourishment level; i.e., 151 million children under 5 years are stunted; in terms of micronutrients, two billion people are not able to meet the requirement for living a healthy life, globally. To meet these demands, the production and supply system has to be

sound. It has been projected that production has to be increased by 60% by 2050, amid different challenges related to the production system posed by climate change (WHO/FAO, 2015), which is further projected to worsen by an increase in the price of food to the extent of 1–29% by 2050. The development of climate-resilient varieties through conventional approaches of hybridization and selection is input-intensive (labor, land, and time), limiting the realized genetic gain. Improvement in the genetic gain as per the Lush equation (Lush, 1943) can be secured through i) better intensity of selection *via* accurate and high-throughput phenotyping and ii) having a broad genetic base representing diverse eco-geography in breeding program. The advancement in genomics approaches leads to the availability of huge resources like genome sequence information, transcriptome, and proteome that have paved the way to hasten the identification of target genes mitigating the effects of climate change (Varshney et al., 2018). This sequence of information also leads to the identification of several mutant loci at the nucleotide level which might be associated with characters of complex nature like yield in general and under different circumstances of stress, which are otherwise very difficult to decipher. Genomic selection emerged as an important tool which can utilize such information for modeling the crop yield for effective and rapid selection under different environmental conditions to meet the production challenges in a climate-changing world.

Changes brought about by climate change have affected the phenology of different crop species leading to a detrimental effect on production and productivity. Different stresses, viz., heat, cold, drought, and flood, are specific manifestations of climate change. Genetic improvement of crops based on phenotypic selection has been successfully achieved through traditional breeding. However, in recent past, genomics led to the identification of several underlying genes/QTLs providing tolerance to these specific conditions, which have been utilized in marker-assisted selection (MAS). MAS is an indirect selection process, where individuals for a particular trait of interest are selected based on the known markers linked to it (Fernando and Grossman, 1989). This method has been efficiently used in the past for selection of individuals in plant breeding to increase the selection accuracy compared to the traditional phenotype-based selection process (Mohan et al., 1997). In cereals, MAS resulted in a number of varieties, viz., Improved Pusa Basmati1 (Gopalakrishnan et al., 2008), Pusa Basmati 1728 (Singh et al., 2017a), Pusa Basmati 1637 (Singh et al., 2017b), Pusa Samba 1850 (Krishnan et al., 2019), Improved Samba Mahsuri (Madhavi et al., 2016), and Swarna-Sub1 (Neeraja et al., 2007) in rice, HUW510 in wheat (Vasistha et al., 2017), and HHB67-Improved in pearl millet (Rai et al., 2008). C214 in chickpea (Varshney et al., 2014a), JTN5503 and DS880 in soybean (Arelli et al., 2006, 2009), and JL24 and TAG24 in groundnut (Varshney et al., 2014b) have been derived using MAS. However, MAS is practically feasible only if the trait of interest is associated with one or very few major genes, and it is impractical or irrelevant for quantitative traits (i.e., polygenic traits that are governed by few hundreds of minor genes) (Bernardo, 2008), which most of the stress tolerance–related traits are based on. To overcome this issue, a new selection tool called genomic selection (GS) was proposed that can facilitate selection for such traits, by means of net genetic merit of an individual obtained using the effects of dense markers distributed across the genome (Meuwissen et al., 2001). In this approach, the individual effect of each marker is estimated, and the additive sum of all the marker effects is used for calculation of the genomic-estimated breeding values (GEBV) of each individual. In the current scenario of climate change, GS is a promising tool for improving the genetic gain of individuals under the breeding program (Yuan et al., 2019). The basic process of any genomic selection process starts with the creation of training population, i.e., individuals having both genotypic and phenotypic information, and this information is used to build a model, where the phenotype is used as a response and genotype as a predictor. The information from the developed model is later used to estimate the GEBV of breeding population, i.e., individuals having only genotypic information. The basic process of GS is also explained in **Figure 1**.

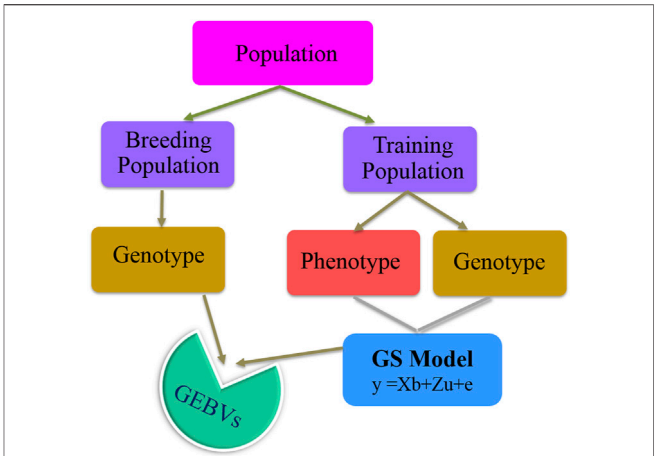


FIGURE 1 | Basic schema of the genomic selection process.

The major advantage of using GS is that it allows for a drastic reduction in the duration of the breeding cycle as compared to traditional breeding and also minimizes the cost associated with extensive phenotyping, thereby subsequently accelerating genetic gains and ensuring food and nutritional security (Heffner et al., 2010). However, there are certain factors such as the size of training and breeding populations, genetic diversity of breeding population, heritability of the underlying trait, influence of genotype–environment (GxE) interaction, density of markers, and genetic relationship between training population and breeding population or selection candidates, which may influence the genomic prediction's accuracy (De Roos et al., 2009; Lorenzana and Bernardo, 2009; Luan et al., 2009; Daetwyler et al., 2010; Clark et al., 2011; Howard et al., 2014). Hence, successful implementation of GS in breeding programs requires careful consideration of all these factors. Apart from these factors, there are certain limitations of genomic selection. Changes in gene frequencies and epistatic interactions drastically affect the estimates of GEBV. Most of the models used to estimate GEBV ignore the effect of epistasis which plays a prime role

especially in cross pollinated plants (Heffner et al., 2009). The rate of declination of selection response is more in GS than pedigree based selection, which can be minimized through the addition of new markers to the model (Nakaya and Isobe, 2012). However, the cost of implementation of GS is more than that of the traditional breeding program.

The choice of models is an important factor in implementing GS, and several parametric and non-parametric genomic prediction models are available for this purpose. One of the most common and widely used parametric genomic selection model is the best linear unbiased prediction (BLUP). It is a mixed model–based whole-genome regression approach that is used to estimate the marker effects, and the same has been successfully applied to predict complex traits (Habier et al., 2009, 2013; de los Campos et al., 2013). In general, it was observed that the performance of parametric models found to be efficient only for traits with additive genetic architectures. For traits that are highly affected by epistatic or non-additive interactions, it becomes challenging to use parametric models (Moore and Williams, 2009). Epistatic interactions play a key role in explaining genetic variation for quantitative traits. Hence, ignoring such type of information in the prediction model might result in lower genomic prediction accuracies (Cooper et al., 2002). Due to these factors, it is not always advisable to practice simple linear or parametric models. Gianola et al. (2006) first used non-parametric and semiparametric methods for modeling the complex genetic architecture. Subsequently, several statistical methods were implemented to model both additive and epistatic effects for genomic selection (Xu, 2007; Cai et al., 2011). For a detailed comparison of various parametric, non-parametric and semiparametric methods in different settings of population size and trait heritability, one can refer to Howard et al. (2014) and Budhlakoti et al. (2020c). Recently, some semiparametric (Legarra and Reverter, 2018) and advanced approaches (Tanaka, 2018; Budhlakoti et al., 2020a, 2020b; Majumdar et al., 2020; Sehgal et al., 2020; Tanaka, 2020; Mishra et al., 2021) have also been proposed and implemented in context to genomic selection. In the next section, few most commonly used methods for genomic selection studies have been discussed.

STATISTICAL MODEL FOR GENOMIC SELECTION

The process of selecting the suitable individuals in GS starts with a simple linear model sometimes also called least-squares regression or ordinary least-squares regression (OLS):

$$Y = 1_n\mu + X\beta + \varepsilon$$

where $\mathbf{Y} = n \times 1$ vectors of observations, μ is the mean, $\boldsymbol{\beta} = p \times 1$ vectors of marker effects, $\varepsilon = n \times 1$ vectors of random residual effects, $\mathbf{X}$ = design matrix of order $n \times p$ (where each row represents the genotype/individuals/lines (n) and each column corresponds to the marker (p)), and $\varepsilon \sim N(0, \sigma_e^2)$.

One major problem in linear models using several thousands of genome-wide markers is that the number of markers (p) exceeds the number of observations (n), i.e., genotype/individuals/lines, and this creates the problem of over-parameterization (large "p" and small "n" problem (p >> n)). Using a subset of significant markers can be an alternative for dealing with the large "p" and small "n" problem. Meuwissen et al. (2001) used a modification of the least-squares regression for GS. They performed least-squares regression analysis on each marker separately with the following model:

$$Y = X_j\beta_j + \varepsilon$$

where $X_j = j^{th}$ column of the design matrix of the markers and β_j = genetic effect of the j^{th} marker.

Markers with significant effects are selected using the log likelihood of this model, and those are further used for estimation of breeding values. However, it has to be noted that some key information may be lost by selection based on the subset of markers.

Hence, an efficient solution for the over-parameterization problem in linear models is using ridge regression (RR), which is a penalized regression–based approach (Meuwissen et al., 2001). It also solves the problems of multicollinearity at the same time (i.e., correlated predictors, e.g., SNP, or markers). RR shrinks the coefficients of correlated predictors equally toward zero and solves the regression problem using ℓ2 penalized least squares. Here, the goal is to derive an estimator of parameter β with a smaller variance than the least-squares estimator. Similar to RR, the least absolute shrinkage and selection operator (LASSO) (Tibshirani, 1996; Usai et al., 2009) is another variant of penalized regression, which uses the ℓ_1 penalized least-squares criterion to obtain a sparse solution. However, sometimes LASSO may not work well with highly correlated predictors (e.g., SNPs in high linkage disequilibrium) (Ogutu et al., 2012). The elastic net (ENET) is an extension of the LASSO that is robust to extreme correlations among the predictors (Friedman et al., 2010), and it is a compromise between ℓ1 penalty (LASSO) and ℓ2 penalty (RR) (Zou and Hastie, 2005).

The RR model considers that each marker contributes to equal variance, which is not the case for all traits. Therefore, the variance of the markers based on the trait's genetic architecture has to be modeled. For this purpose, several Bayesian models have been proposed where it is assumed that there is some prior distribution of marker effects. Furthermore, inferences about model parameters are obtained on the basis of posterior distributions of marker effects. There are several variants of Bayesian models for genomic prediction such as Bayes A, Bayes B, Bayes Cπ, and Bayes Dπ (Meuwissen et al., 2001; Habier et al., 2011) and other derivatives, e.g., Bayesian LASSO and Bayesian ridge regression (BRR). Besides the marker-based models, the best linear unbiased prediction (BLUP) (Henderson et al., 1959) is one of the most commonly used genomic prediction methods. There are many variants of BLUP available for this purpose, e.g., genomic BLUP (GBLUP), single-step GBLUP (ssGBLUP), ridge regression BLUP (RRBLUP), and

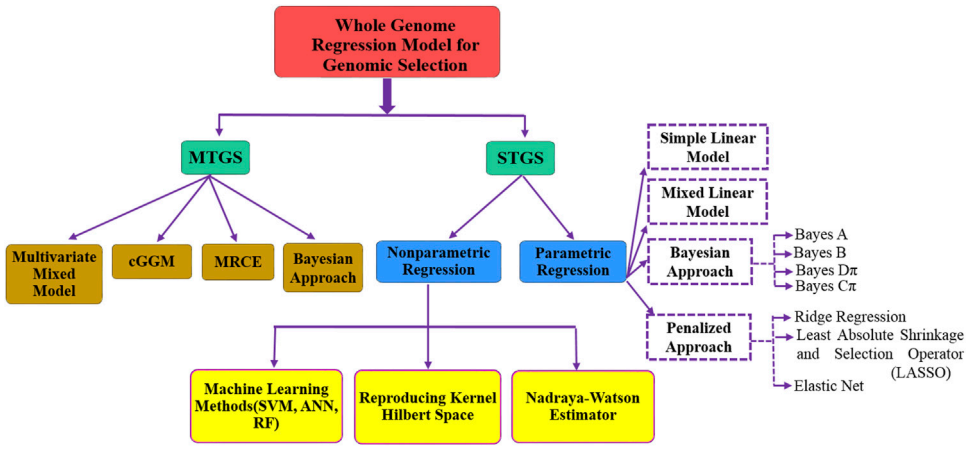


FIGURE 2 | Overall summary of the most commonly used models in genomic selection.

GBLUP with linear ridge kernel regression (rrGBLUP), of which GBLUP is very frequently used. The GBLUP uses the genomic relationships calculated using markers instead of the conventional BLUP which uses the pedigree relationships to obtain the GEBV of the lines or individuals (Meuwissen et al., 2001).

The genomic prediction models discussed so far perform well for traits with additive genetic architecture, but their performance becomes very poor in case of epistatic genetic architectures. Hence, Gianola et al. (2006) first used non-parametric and semiparametric methods for modeling the complex genetic architecture. Subsequently, several statistical methods were implemented to model both additive and epistatic effects for genomic selection (Xu, 2007; Cai et al., 2011; Legarra and Reverter, 2018). There are several non-parametric methods that have been studied in relation to genomic selection, e.g., NW (Nadaraya–Watson) estimator (Gianola et al., 2006), RKHS (reproductive kernel Hilbert space) (Gianola et al., 2006), SVM (support vector machine) (Maenhout et al., 2007; Long et al., 2011), ANN (artificial neural network) (Gianola et al., 2011), and RF (random forest) (Holliday et al., 2012), among them SVM, NN, and RF are based on the machine learning approach.

Methods discussed earlier in this section are based on genomic information where information is available for a single trait, i.e., single-trait genomic selection (STGS). As the performance of STGS-based methods may be affected significantly in case of pleiotropy, i.e., one gene linked to multiple traits, a mutation in a pleiotropic gene may have an effect on several traits simultaneously. It was observed that low heritability traits can borrow information from correlated traits and consequently achieve higher prediction accuracy. However, STGS-based methods consider the information of each trait independently. Hence, we may lose crucial information which may ultimately result in poor genomic prediction accuracy. Nowadays, as we are receiving data on multiple traits, so multi-trait genomic selection (MTGS)-based methods may provide more accurate GEBV and subsequently a higher prediction accuracy. Several MTGS-based methods have been studied in relation to GS, e.g., multivariate mixed model approach (Jia and Jannink, 2012; Klápště et al., 2020), Bayesian multi-trait model (Jia and Jannink, 2012; Cheng et al., 2018), MRCE (multivariate regression with covariance estimation) (Rothman et al., 2010), and cGGM (conditional Gaussian graphical model) (Chiquet et al., 2017). Jia and Jannink (2012) presented three multivariate linear models (i.e., GBLUP, Bayes A, and Bayes Cπ) and compared them to univariate models, and a detailed comparison of various STGS- and MTGS-based methods has also been studied by Budhlakoti et al. (2019c). A brief structure of different STGS- and MTGS-based methods used in GS studies is given in **Figure 2**.

GS: IMPLICATIONS IN CROP IMPROVEMENT

GS in Cereals

Cereals are an important part of our daily diet as they contribute about 50% of the total dietary energy supply (WHO/FAO, 2003). Wheat, rice, maize, and barley are the major cereal crops, which are being grown on arable land all over the world amounting to a total of 2,817 million tonnes of production (FAO). Production of these crops is being challenged by calamities created by a change in climatic pattern (Reynolds, 2010), and over that, it is being complicated by the rising demand of increasing population (Tester and Langridge, 2010; Furbank and Tester, 2011). To meet the challenges, the production system has to be efficient and sustainable with lower pressure on the ecosystem. High-yielding, resource-efficient crop varieties are an integral component of such production systems which can address the challenges. But the development of such variety is a painstaking endeavor as most of the crop productivity traits are under the control of a complex genetic system (most genes are of minor effect) with the complication of low heritability and high order of epitasis (Mackay, 2001). Though conventional selection methods have resulted in a number of varieties but the genetic gain per unit time is not as much rewarding as GS, it provides an opportunity to hasten the cycle of selection (Bernardo and Yu, 2007; Lorenz

et al., 2011). The potential of GS can be assessed from the fact that it has the ability to select high breeding value individuals rapidly from early-generation populations without the need of extensive phenotyping. This has been shown effectively in cereal crops in the recent past. Wheat, rice, maize, and barley are the first candidate crops where the effectiveness of GS has been studied. GS in these crops leads to the identification of different models which were able to efficiently predict the performance of traits under question and filter out the important breeding material. In the following section, the role of GS in cereal crops has been discussed.

Grain Yield and Related Traits' Improvement

Grain yield is a major trait which is affected directly or indirectly by other traits including thousand grain weight, number of tillers bearing panicle, number of grains per panicle, number of filled grains per panicle etc. Genomic prediction for these traits utilizing different types of training populations and models have been evaluated. The variations in the accuracies of genomic prediction have been attributed to the heritability of the trait, training population, and models used. The genomic prediction accuracy for a very complex and physiological trait–like distribution of weight to the individual grain in the panicle in rice (Yabe et al., 2018) ranged from 0.28 to 0.78 for grain yield in maize (Rio et al., 2019). For the improvement of accuracy, the role of training population also has a significant effect, and it has been reported that prediction based on the training set developed using North Carolina mating design II (0.60) was found at par with that of full diallel matings (0.58) and superior to that of test cross (0.10) (Fristche-Neto et al., 2018). Similarly, better prediction accuracies for grain yield were observed in recombinant inbred lines and doubled haploid populations compared to natural populations (Liu et al., 2018). The accuracy of GS for grain yield is also highly influenced by the size of training populations and genetic relationships between the training and breeding populations (Lozada et al., 2019; Lozada and Carter, 2020). Longin et al. (2014) reported that GS followed by one cycle of phenotypic selection has been reported to facilitate identification of superior parental lines with better combining ability and high annual genetic gain for grain yield in wheat than simple phenotypic selection. However scheme had not considered the cost and time involved in production and nursery screening of these lines, and thus, additional schemes like GSrapid have been proposed which have better selection gain and have been recommended for utilization in a hybrid breeding program of different cereal crops (Marulanda et al., 2016). GS could also be potentially used in the prediction of the performance of a large number of hybrid combinations (VanRaden, 2008; Crossa et al., 2017). The earlier GS studies on cereals started with wheat where the DArT marker system was used (Crossa et al., 2010, 2011; Heffner et al., 2011; Burgueño et al., 2012; Pérez-Rodríguez et al., 2012). However, later, other genome-wide SNP platforms became the routine marker in genomic selection owing to their own advantages (Poland et al., 2012; Zhao et al., 2012). Detailed information on GS studies for grain yield and related traits in major cereals, pulses, oilseeds, and horticultural crops with the details of statistical models, marker platforms, types of populations used, and the prediction accuracies of statistical models are listed in **Table 1**.

Biotic Stress Tolerance

With the change in weather patterns, emergence/resurgence of new races and biotypes of pathogens and insects is being reported globally (Juarez et al., 2013; Váry et al., 2015; Fones et al., 2020). Hence, identification of resistance genes in the germplasm and their incorporation into the breeding program are required to develop biotic stress–tolerant varieties. MAS has proved to be efficient in breeding for qualitative resistance, but for quantitative resistance which is governed by many genes with smaller effects, MAS has not been so effective. GS has proved its role in improving tolerance against biotic stresses in cereals which are quantitatively controlled, though it has been applied to a very limited extent. Most of the studies on the utility of GS for biotic stress tolerance have been reported from wheat, for a wide array of diseases including three types of rusts, *Fusarium* head blight, septoria tritici blotch, powdery mildew, tan spot, and *Stagonospora nodorum* blotch. The genomic prediction accuracies for these diseases ranged from 0.14 to 0.85 (Rutkoski et al., 2012; Daetwyler et al., 2014; Mirdita et al., 2015; Juliana et al., 2017; Sarinelli et al., 2019). In rice, GS has been utilized to identify blast-tolerant lines (Huang et al., 2019). In maize, GS has been successfully utilized to select lines from natural populations for tolerance to *Stenocarpella maydis* causing ear rot (dos Santos et al., 2016) and from biparental populations for superior yield under heavy infestation of *Striga* (Badu-Apraku et al., 2019). In case of barley, markers and prediction models were utilized for *Fusarium* head blight severity, and the prediction accuracy was quite higher, i.e., 0.72, than that of conventional phenotyping (Lorenz et al., 2012; Sallam and Smith, 2016).

Abiotic Stress Tolerance

The occurrence of drought, high-temperature stress during crop growth stages, flood, etc., is at surge due to climate change, causing significant crop losses (Qin et al., 2011). With the 1°C increase in global temperature, yield reduction has been predicted up to 6.4% in wheat (Liu et al., 2016). The sustainable and economic options under such situations to cover the losses are changing cropping patterns or developing abiotic stress–tolerant varieties. Identification of tolerant genotypes from the germplasm and their utilization in the breeding program become a prime requirement for development of such varieties (Baenziger, 2016). The major issue in breeding for abiotic stress tolerance is their complex inheritance, low heritability, and high environmental effect on them (Bernardo, 2008).

Conventional breeding methods for abiotic stresses suffer from limitations of accuracy and reproducibility. Though molecular markers have been utilized to identify and transfer yield QTLs under abiotic stress conditions (Ribaut and Ragot, 2007; Almeida et al., 2013), but it may not be effective as QTL from limited genetic resources explain little variation for grain yield under stress and are also highly influenced by the genetic background (Semagn et al., 2013) as well as the environment and

TABLE 1 | Genomic prediction for grain yield and related traits in different crops (i.e. Cereals, Pulses, Oilseeds and Horticultural crops).

Crop	Model	Genotyping Techniques	Population type	Trait		Prediction accuracy (PA)	Reference
A. Cereals							
i) Maize	GBLUP	Taqman (ABI 2002)	F1 from half diallel and test crosses	Grain yield (GY)		0.58	Zhao et al. (2012)
	GBLUP	Affymetrix®	F1 from test crosses (TC), North Carolina design II (NCII), and full diallel (FD)	GY		0.10 (TC) 0.58(NCII) 0.60(FD)	Fristche-Neto et al. (2018)
	RRBLUP	55 K SNP array	Natural population (NP), recombinant inbred line (RIL), double haploid (DH), and F2:3	GY		RIL&DH (0.41) > F2:3 (0.36) > NP(0.40)	Liu et al. (2018)
	GBLUP			100 kernel weight		F2:3 (0.77) > RIL&DH (0.65) > NP(0.48)	
	Bayes A, Bayes B, Bayes C, LASSO, and RKHS					()	
	GBLUP and multigroup GBLUP	Genotyping by sequencing (GBS) 50 K Illumina® 600 K and Affymetrix® Axiom	TC	GY		0.78	Rio et al. (2019)
				Yield index (YI)		0.73	
	RRBLUP and BSSV (Bayesian stochastic search variable)	DArTSeq™ and Illumina HiSeq2000	Inbred lines	Ear rot	Proportion of rotten kernel	0.87	dos Santos et al. (2016)
					Ear rot incidence	0.24–0.56	
	BLUP	Kompetitive Allele Specific PCR (KASP)	Inbred lines and test cross progenies	*Striga* resistance		0.58	Badu-Apraku et al. (2019)
				Drought tolerance		0.42–0.65	
	GBLUP	GBS	Breeding lines	Drought tolerance		0.37–0.38	Beyene et al. (2015)
	RRBLUP and GBLUP	KASP	Inbred lines and half diallel population	Water-logging tolerance		0.53–0.84	Das et al. (2020)
	BLUP	KASP	Asian and African inbred lines	Drought tolerance	GY	0.71–0.75	Vivek et al. (2017)
					Anthesis–silking interval (ASI)	0.35–0.43	
	RRBLUP	Infinium Maize SNP50 Bead Chip	Subtropical maize lines	Drought tolerance	ASI	0.93	Shikha et al. (2017)
	Bayes A, Bayes B, and LASSO				100 kernel weight	0.92	
ii) Wheat	RRBLUP, RKHS, and Bayesian LASSO	Diversity Arrays Technology (DArT)	Advanced breeding and germplasm lines	GY		0.49–0.61	Crossa et al. (2010)
	Bayesian LASSO and RKHS	DArT	Breeding lines	GY		0.43–0.79	Crossa et al. (2011)
	Bayes A, Bayes B, Bayes C, and RRBLUP	DArT	Breeding lines	GY		0.48	Heffner et al. (2011)
	Bayesian LASSO	DArT	Breeding lines	GY		0.5–0.6	Burgueño et al. (2012)
	RRBLUP, Bayes A, Bayes B, Bayes C, LASSO, NN, and RKHS	DArT	Breeding lines	GY		0.6–0.7	Pérez-Rodríguez et al. (2012)
	GBLUP	DArT	Breeding lines	GY		0.2–0.4	Poland et al. (2012)
	RRBLUP, Bayes A, Bayes B, and Bayes C	9 K Illumina® Infinium	F1s	GY		0.3–0.6	Zhao et al. (2013)
	RRBLUP	9 K Illumina® and 90 K iSelect	Red winter wheat breeding lines	GY		0.14–0.43	Lozada et al. (2019)
	GBLUP	DArT and KASP	F4:6 population	GY		0.75	Michel et al. (2019)
	GBLUP	GBS	Breeding lines	GY		0.42–0.56	Juliana et al. (2019)
	GBLUP	GBS	Breeding population	GY		0.12–0.34	Sun et al. (2019)
	GBLUP and IBCF:MTME (item-based collaborative filtering: multi-trait multi-environment)	Illumina® 90 K	Winter wheat lines	GY		-0.21 to 0.42	Lozada and Carter (2020)
	GBLUP and BRR	Infinium iSelect 9 K	Germplasm	Leaf rust resistance (LRR)		0.35	Daetwyler et al. (2014)
				Stem rust resistance (SRR)		0.27	
				Yellow rust resistance (YRR)		0.44	
	RR RKHS RF Bayesian LASSO and multiple linear regression	DArT	Breeding lines	*Fusarium* head blight (FHB) resistance		0.006–0.463 0.118–0.575	Rutkoski et al. (2012)
				Deoxynivalenol (DON) resistance			
	RRBLUP Bayes Cπ and RKHS	Illumina Infinium 9 K and 90 K	Winter wheat breeding lines	FHB resistance		0.6	Mirdita et al. (2015)
				Septoria leaf blotch resistance		0.5	
	RRBLUP	GBS	Winter wheat breeding lines	Powdery mildew resistance		0.60	Sarinelli et al. (2019)
				GY		0.64	
				Test weight		0.71	
	RKHS and GBLUP	GBS	Lines from International Bread Wheat Screening Nursery	LRR	Seedling	0.31–0.74	Juliana et al. (2017)
					Adult	0.12–0.56	
				YRR	Seedling	0.70–0.78	
					Adult	0.34–0.71	
				SRR		0.31–0.65	

(Continued on following page)

TABLE 1 | (*Continued*) Genomic prediction for grain yield and related traits in different crops (i.e. Cereals, Pulses, Oilseeds and Horticultural crops).

Crop	Model	Genotyping Techniques	Population type	Trait	Prediction accuracy (PA)	Reference
iii) Rice	Bayesian LASSO	DArT	Inter-related synthetic population	GY	0.309	Grenier et al. (2015)
				Panicle weight	0.327	
	RRBLUP	GBS	Tropical rice breeding lines	GY	0.31	Spindel et al. (2015)
	GBLUP	Illumina HiSeq 2000	128 Japanese rice varieties	Field grain	0.30	Yabe et al. (2018)
				Field grain weight	0.28	
		Illumina HiSeq 4000 and HiSeqX		Variance of field grain	0.53	
	GBLUP, SVM, LASSO, and PLS	GBS	North Carolina design II population	GY	~0.5	Xu et al. (2018)
				Thousand grain weight (TGW)	~0.28	
	GBLUP	Illumina® HiSeq 2000	Hybrid population	GY	0.54	Cui et al. (2020)
				Grain length	0.92	
	GBLUP, RKHS, and Bayes B	GBS	Breeding lines	Panicle weight	0.30	Hassen et al. (2018)
				Nitrogen balance index	0.21	
	GBLUP	SNP	Breeding lines	GY	0.39	Wang et al. (2018)
				TGW	0.88	
	RRBLUP and GBLUP	GBS	Rice population	Blast resistance	0.17–0.73	Huang et al. (2019)
	GBLUP and RKHS	962 K Core SNP dataset	Germplasm	Drought tolerance	0.226–0.809	Bhandari et al. (2019)
iv) Barley	RRBLUP	Illumina GoldenGate	Breeding lines	GY	0.57	Sallam et al. (2015)
				DON	0.72	
				FHB	0.74	
	GBLUP and RKHS	GBS	Breeding lines	Thousand kernel weight (TKW)	0.67	Abed et al. (2018)
	GBLUP	Illumina	Breeding lines	GY	0.362	Tiede and Smith (2018)
				DON resistance	0.367	
B. Pulses						
i) Lentil	RRBLUP GBLUP Bayes A Bayes B Bayes Cπ Bayesian LASSO BRR and RKHS	Exome capture	Lentil diversity panel, RIL	Maturity duration	0.58–0.84	Haile et al. (2020)
ii) Common bean	GBLUP Bayes A Bayes B Bayes C Bayesian LASSO and BRR	GBS	RIL, multi-parent advanced generation inter-cross (MAGIC), germplasm	Cooking time	0.22–0.55	Diaz et al. (2021)
	RKHS	GBS	Breeding lines	Root rot resistance *Fusarium*	0.52	Diaz et al. (2021)
				Pythium	0.72–0.79	
iii) Chickpea	RRBLUP Bayesian LASSO and BRR	Whole-genome re-sequencing (WGRS)	Breeding lines	Drought tolerance	0.56–0.61	Li et al. (2018)
C. Oilseeds						
i) Groundnut	Bayesian generalized linear regression	Affymetrix GeneTitan®	Breeding lines	Yield	0.49–0.60	Pandey et al. (2020)
				Protein	0.41–0.46	
				Rust resistance	0.74–0.75	
				Late leaf spot resistance	0.57–0.65	
ii) *Brassica napus*	RRBLUP	Infinium Array 60 K	Test cross F1s	Seed yield	0.45	Jan et al., 2016
				Oil content	0.81	
				Lodging resistance	0.39	
	GBLUP	Transcriptome GBSt assay	Spring canola lines	Seed yield	0.69	Fikere et al. (2020)
				Oil content	0.64	
	GBLUP LASSO	Illumina Infinium 60 K	Double haploid population	Seed yield	0.27–0.55	Xiong et al. (2020)
iii) Sunflower	and multi-kernel BLUP	GBS	F1s from factorial mating design	Oil content	0.783	Mangin et al. (2017)
iv) Soybean	RRBLUP	iSelect Bead Chip	RILs from interspecific cross	Yield	0.68	Beche et al. (2021)
				Oil content	0.76	
	Bayes B and Bayesian LASSO	BARCSoySNP6K		Protein content	0.76	
	RRBLUP	iSelect Bead Chip BARCSoySNP6K	Breeding lines	Oil content Protein content	0.30 0.55	Stewart-Brown et al. (2019)

(Continued on following page)

TABLE 1 | (*Continued*) Genomic prediction for grain yield and related traits in different crops (i.e. Cereals, Pulses, Oilseeds and Horticultural crops).

Crop	Model	Genotyping Techniques	Population type	Trait	Prediction accuracy (PA)	Reference
D. Horticultural crops						
i) Apple	RRBLUP	HiScan Illumina and Infinium Array 20 K	Germplasm and biparental families	Firmness	0.81	Roth et al. (2020)
	RRBLUP and Bayesian LASSO	Infinium® II 8 K	F1 from factorial mating design	Fruit firmness	0.83	Kumar et al. (2012)
				Soluble solids	0.89	
ii) Citrus	GBLUP	GBS	F1 from different parental lines	Fruit weight	0.650	Imai et al. (2019)
				Sugar content	0.519	
				Acid content	0.666	
	GBLUP	Illumina HiSeq 2000	Varieties and their full sib families	Fruit weight distribution	0.89	Minamikawa et al. (2017)
iii) Apricot	GBLUP	GBS	F1 pseudo-testcross population	Glucose content	0.31	Nsibi et al. (2020)
	RRBLUP				0.78	
	Bayes A, Bayes B, Bayes C, Bayesian LASSO, and BRR			Ethylene content		
iv) Pear	GBLUP	GBS	Full sib families	Crispness	0.32	Kumar et al. (2019)
				Sweetness	0.62	
v) Capsicum	GBLUP, Bayesian LASSO, Bayes B, Bayes C, and RKHS	GBS	Core collection and RIL population	Fruit length	0.32	Hong et al. (2020)
				Fruit width	0.50	
				Fruit shape	0.34	
				Fruit weight	0.48	
vi) Tomato	RRBLUP	Infinium Assay	Germplasm	Fruit weight	0.814	Duangjit et al. (2016)
				Firmness	0.614	
				Soluble solids	0.714	
				Sugar content	0.649	
				Acidity	0.619	
				Biochemical profile	0.126–0.705	

there interactions. GS is superior to MAS, and the prediction efficiency is also higher for abiotic stress tolerance (Cerrudo et al., 2018). The usefulness of GS has been shown in wheat, maize, and rice for drought and heat tolerance.

Beyene et al. (2015) have reported a gain of 0.086 t/ha for grain yield, following the rapid cycling GS strategy in eight biparental populations of maize under drought conditions, and a final gain of 0.176 t/ha after three cycles of selection. This increased the genetic gain as the time required for selection was reduced significantly as compared to that of the conventional breeding scheme, where it was three times higher with phenotypic selection. Similarly, Das et al. (2020) reported a genetic gain of 0.110 and 0.135 t/ha/yr for grain yield under drought and 0.038 and 0.113 t/ha/yr under water logging in two maize populations, viz., Maize Yellow Synthetic 1 and Maize Yellow Synthetic 2, respectively, following rapid cycling genomic selection. Vivek et al. (2017) compared the performances of second cycle selection through phenotypic and rapid cycle genomic selection and found 10–20% superiority using the latter. Genomic prediction accuracies using multi-environment models for drought stress tolerance were higher than those using single-environment models in rice and wheat (Sukumaran et al., 2018; Bhandari et al., 2019). Prediction accuracies were higher for heat and drought stress in case of wheat when secondary traits contributing to yield were considered under stress rather than yield *per se* using genomic prediction (Rutkoski et al., 2016). Comparative analysis among different models leads to the conclusion that multi-trait models are superior when selection is carried out in severe drought conditions, while the random regression model was better than the repeatability model and multi-trait model under normal drought conditions and also use of secondary high-throughput traits in genomic prediction improved accuracies by ~70% (Sun et al., 2017).

Quality Improvement

Quality traits have varied genetic architectures, some being controlled oligogenically like grain color, while others are polygenic in nature, viz., grain size and protein content (Battenfield et al., 2016). GS has been carried out in wheat extensively for quality-related traits, viz., milling and flour quality, and when prediction accuracies were compared in biparental and multi-family populations, it was concluded that the prediction accuracies in multi-family populations were better (Heffner et al., 2011).

Protein content is known to be negatively correlated with yield due to physiological compensation (Lam et al., 1996). Michel et al. (2019) employed multi-trait genomic selection for grain yield, protein content, and dough rheological traits for efficient selection with optimized yield and protein content with better quality. The prediction accuracy for the quality traits depends on variability in the germplasm, the relationship among training and prediction populations, etc. (Crossa et al., 2014; Zhao et al., 2015). Joukhadar et al. (2021) used Bayesian regression and BRR for rapid improvement of grain yield as well as mineral content to biofortify wheat and reported Bayesian regression was better in predicting mineral content with an accuracy of 0.55. In rice, grain length and width are important quality parameters, and the

prediction accuracy for these traits ranged from 0.35 to 0.45 and 0.5 to 0.7, respectively, in 110 Japanese rice cultivars employing various GS models (Onogi et al., 2015). In barley, the prediction for quality traits like malting quality (prediction accuracy: 0.4–0.8) has shown the prospects of GS for screening large populations without the need of cost-intensive phenotyping (Schmidt et al., 2016).

GS in Oilseeds

Oilseeds are a source of livelihood to the smallholder farmers in developing countries of Asia and Africa. The yield potential is still to be realized by bridging the yield gap *via* inducing tolerance to biotic and abiotic stresses and improvement in quality (Janila et al., 2016). Different traits related to biotic and abiotic stresses have been mapped, but most of them are qualitative in nature, and the report of GS is limited in such potential crops. Oil quality and yield traits are influenced by the environment and GxE interactions (Patil et al., 2020). Hence, it is important to use the appropriate GS models to account for the GxE effects for accurate selection. Pandey et al. (2020) employed GS in groundnut with different models and validation schemes to account for GxE interaction effects. The model having genomic information generated from the SNP (G), genotypic effect of the line (L), environment effect (E), and their interactions (LxE and GxE) had better mean accuracy (0.58) for all the traits compared to other models. Jan et al. (2016) employed the RRBLUP model for GS in *Brassica* using 950 cross combinations derived from utilizing 475 lines and two testers, for the improvement of oil-specific traits, and the accuracy for oil content and oil yield was 0.81 and 0.75, respectively. Hence, they concluded that the GS model is helpful in pre-selecting superior cross combinations before extensive field evaluation over location and years saving resources. Fikere et al. (2020) employed GS for 22 traits related to yield, disease resistance, and quality in *B. napus* and reported prediction accuracy was highest for yield (0.69) followed by oil content (0.64) using GBLUP. They also evaluated genomic prediction for compositional fatty acid estimated under rainfed and irrigated conditions and concluded that the prediction accuracies for these traits were lower under non-irrigated conditions. Xiong et al. (2020) employed various prediction models, viz., LASSO, GBLUP, OLS, and OLS post-LASSO, for different traits in *B. napus* and reported the two-stage method OLS post-LASSO to be the most accurate (0.90 and 0.55 for oil content and single plant yield, respectively) with the provision of incorporating GxE interactions. For oil content in sunflower which is highly heritable and additive in nature, Mangin et al. (2017) reported that accuracy based on general combining ability (GCA) and GS were on par, and in case if there is no knowledge about one of the parents of hybrid combination, GS excels the GCA-based predictions. Similar inferences had been made by Reif et al. (2013) for the prediction of hybrid performance in sunflower.

From a cross between cultivated and wild progenitors of soybean (*G. max* X *G. sojae*), Beche et al. (2021) reported that the yield-related alleles were associated with the cultivated elite line, but the protein content alleles were from the wild progenitor. The difference in the distribution of trait-contributing alleles in such crosses has a greater impact on their predictive accuracy. When each allele is distributed equally in the population, the predictive accuracy for both the alleles is the same. In such cases, it is obvious that the less frequent allele's prediction is biased downward. Contiguous breeding programs are very common where new cross combinations are added each year. In such cases, using nested association mapping (NAM) population is better in terms of prediction accuracy (for yield 0.68 and oil and for protein content 0.76) than biparental population, showing the potential of NAM where connectedness is there among the population on the basis of the common parent (Beche et al., 2021). Similarly, Stewart-Brown et al. (2019) have reported that, for better predictions in soybean, it is important to have good relatedness among training and breeding populations. They have observed that the size of the training population has a larger effect on the prediction accuracy, compared to the marker density, but increasing the training population sizes beyond a limit had a diminishing return on the prediction accuracy. Hu et al. (2011) applied GS for biological process, i.e., embryogenesis capacity in soybean, and reported a good prediction accuracy (0.78).

GS in Pulses

In lentil, Haile et al. (2020) showed that if large-effect QTLs were present in the population, multi-trait–based Bayes B is the best GS model, while single-trait GS (STGS) is suitable in their absence. They also reported that, for low heritable traits with GxE interactions, MTGS improves predictability. Considering quality traits in *Phaseolus*, i.e., cooking time for screening of fast culinary genotypes, Diaz et al. (2021) evaluated GS using different populations (RIL, MAGIC, Andean, and Mesoamerican breeding lines). The trait was highly heritable (0.64–0.89), and genomic prediction accuracies for cooking time using MAGIC population were promising and high (0.55) compared to those of Mesoamerican genotypes (0.22).

Under the circumstance of less connectedness in the training and prediction populations, markers generated using the whole genome re-sequencing (WGRS) platform increase the prediction accuracy; however, Li et al. (2018) proposed first identifying causal variants and then utilizing them into the prediction. The prediction accuracy was 0.148–0.186 for yield under drought when using all the SNP from WGRS, but when filtered yield-related causal SNPs were employed, it was observed that prediction accuracy significantly improved (0.56–0.61). Diaz et al. (2021) employed GS for root rot resistance and reported high prediction accuracies (0.7–0.8) for both rots (*Pythium* and *Fusarium*) in *Phaseolus* and proposed it to be promising for improving quantitative tolerance.

GS in Horticultural Crops

Fruit and vegetables are indispensable in achieving nutritional security. However, the problem associated with their breeding, especially of fruits, has its own limitations, viz., long juvenile phase and highly heterozygous nature. Therefore, genetic gain is not much as per the Lush equation. In such crops, GS can be a perfect tool where prediction of performance for quality- and yield-related traits which are under a complex genetic system can

be utilized to improve selection accuracy and efficiency in developing varieties. The success of GS in annual crops has led the horticultural crop breeder to utilize its potential in perennial fruit as well as annual fruit and vegetable crops. Roth et al. (2020) evaluated 537 genotypes in apple for fruit texture traits and performed GS and reported the accuracy up to 0.81. It was suggested to have a large training population from which a tailored training population with a priori genetic relatedness information and ample variation can be formed and utilized to predict the performance of population under consideration. Kumar et al. (2012) have shown high prediction accuracy in apple for different quality traits utilizing a factorial mating design (0.70–0.90). Imai et al. (2019) reported that ssGBLUP predicts with higher accuracy (0.650, 0.519, and 0.666) than GBLUP (0.642, 0.432, and 0.655) for quality traits in citrus, viz., fruit weight, sugar content, and acid content from population where some individuals are not genotyped using information from genotyped related individuals, hence reducing the cost at hand.

As fruits are perishable produce and the post-harvest attribute of the fruits plays an important role in storability, attempts have been made to employ GS for such traits. In apricot, Nsibi et al. (2020) reported prediction accuracy ranging from 0.31 to 0.78 for glucose content and ethylene production. Minamikawa et al. (2017) compared different models of GS for fruit weight distribution among two groups of fruit sizes and reported that, among a large fruit size group, rrGBLUP (0.89) was superior to GBLUP (0.74) and the same was in the case of a small fruit size group, i.e., rrGBLUP (0.32) and GBLUP (0.30). Also, it was proposed to have breeding population or combined parental and breeding population as training population to have better accuracy than only having parental as training population which was consistent for all the quality-related traits. Kumar et al. (2019) employed GS in pear for various fruit quality traits ranging from texture to taste and observed the prediction accuracy ranged from 0.32 to 0.62 averaging to 0.42 and also suggested that training population should be multi-generational and evaluated rigorously over location and time, to have better prediction accuracy. Various GS models have been evaluated for different fruit-related traits in capsicum and reported that RKHS had better accuracy ranging from 0.75 to 0.82 and positively correlated with the number of markers (Hong et al., 2020). GS is also performed to evaluate the accuracy of prediction of different biochemical parameters important for fruit quality in tomato which ranged from 0.13 to 0.70 for aspartate content and also for other traits, viz., fruit weight (0.81), firmness (0.61), soluble solids (0.71), sugar content (0.65), and acidity (0.62) (Duangjit et al., 2016).

STATISTICAL TOOLS FOR IMPLEMENTING GENOMIC SELECTION

Several tools and packages have been developed for the evaluation of genomic prediction and implementation of GS, some of which are discussed below.

GMStool

It is a genome-wide association study (GWAS)-based tool for genomic prediction using genome-wide marker data. It searches for the optimum number of markers for prediction using appropriate statistical and machine learning/deep learning–based models and chooses the best prediction model (Jeong et al., 2020). Furthermore, it identifies SNP markers with the lowest *p*-values (e.g., top 100 markers) in the GWAS and then chooses the relevant markers set to be included in the final prediction model. GMStool is R-based and freely available through the GitHub repository at https://github.com/JaeYoonKim72/GMStool. The whole process or its algorithm is basically divided into three steps: data preparation, marker selection, and final prediction model. The detailed procedure of GMStool is discussed below.

Step 1: Input data are divided into training and test sets (user defined)

Step 2: The training set is further divided into small datasets for performing cross validation (i.e., k-folds, for example, five or ten folds) followed by marker selection in each group or fold. The process of marker selection is performed in each fold/group simultaneously.

Step 3: The selected marker from each fold is integrated into the final marker set for updating the model. Appropriate statistical and machine learning–based models are then used for genomic prediction.

solGS

It is an open-source tool based on the Linux operating system. The workflow of the tool is broadly divided into two steps, i.e., training of the prediction model and obtaining GEBV. However, there are three approaches available for training the prediction model, i.e., trait-based approach, trial approach, and custom lists approach. Here, model input and output could be visualized graphically and can be interactively explored or downloaded. It is designed to store a large amount of genotypic, phenotypic, and experimental data. In the background, it basically uses two R-based packages, i.e., nlme (Pinheiro et al., 2017) for data preprocessing and rrBLUP (Endelman, 2011) for statistical modeling. solGS was earlier used by the NEXTGEN Cassava project (http://nextgencassava.org) and implemented at the Cassavabase website (http://cassavabase.org/solgs).

rrBLUP

It is an R package based on BLUP, which is a mixed linear model framework (Endelman, 2011). It is one of the most widely used packages for genomic prediction in animal and plant breeding. This package estimates the marker effects from training datasets and ultimately estimates the GEBV for the selection candidates. The *mixed.solve* function, a linear mixed model equation which estimates marker effects and GEBV, is one of the most commonly used functions of this package. An additive relationship matrix of individuals can be calculated using genotypic data for the estimation of GEBV using GBLUP. rrBLUP is an open-source package and can be accessed at https://CRAN.R-project.org/package=rrBLUP.

BWGS

It is an integrated pipeline based on R and freely available at https://CRAN.R-project.org/package=BWGS. The BWGS (i.e., BreedWheat Genomic Selection) pipeline (Charmet et al., 2020) basically consists of three modules: i) missing data imputation, ii) dimension reduction, i.e., reducing the number of markers as it could enhance the speed of computation on large datasets, and iii) estimation of GEBV. It has a wide choice of totally 15 parametric and non-parametric statistical models for estimation of GEBV for selection candidates. It could be used for estimation of GEBV for a wide range of genetic architectures. This tool comprises mainly two functions: *bwgs.cv* and *bwgs.predict*. The former is used for missing value imputation, dimension reduction, and cross validation, while the later is used for model calibration and estimation of GEBV for selection candidates.

BGLR

This package is basically an extension of the BLR package (Perez and Campos, 2014). It can be used to implement several Bayesian models and also provides flexibility in terms of prior density distribution. Here, the response to be considered could be continuous or categorical (either binary or ordinal). It is freely available in the public domain through the CRAN mirror at https://CRAN.R-project.org/package=BGLR.

GenSel

The GenSel software program was developed and implemented under the BIGS (Bioinformatics to Implement Genomic Selection) project (Fernando and Garrick, 2009). It is used for estimation of molecular marker–based breeding values of animals for the trait of interest. This can serve the purpose through the command line (MAC or Linux) interface or as a user-friendly tool. The jobs are submitted and assigned in the queue for analysis. The software uses the Bayesian approach in the background for estimation of marker effects from the training data and further for estimation of GEBV for breeding candidates. This software program can be accessed at https://github.com/austin-putz/GenSel.

GSelection

This is an R-based package and is freely available at https://CRAN.R-project.org/package=GSelection. The package comprises of a set of functions to select the important markers and estimates the GEBV of selection candidates using an integrated model framework (Majumdar et al., 2019). The motivation behind this package is that not a single method performs best in case of all crop plants or animal breeding programs as they may have diverse genetic architectures, i.e., additive and non-additive genetic effects. This package has been developed by integrating the best performing model from each category of additive and non-additive genetic models.

lme4GS

lme4GS is an R-based package freely available and can be accessed through the GitHub repository at https://github.com/perpdgo/lme4GS. It is an extension of the lme4 R package, which is the standard package for fitting linear mixed models. lme4GS package is basically motivated from existing R packages pedigreemm (Vazquez et al., 2010) and lme4qtl (Ziyatdinov et al., 2018). lme4GS package can also be considered an extension of the rrBLUP (Endelman, 2011) package. Further, lme4GS package can be used for fitting mixed models with covariance structures defined by the user, bandwidth selection, and genomic prediction.

STGS

It is an R-based package developed for genomic predictions by estimating marker effects, and the same is further used for calculation of genotypic merit of individuals, i.e., GEBV. GS may be based on single-trait or multi-trait information. This package performs genomic selection only for a single trait, hence named STGS, i.e., single-trait genomic selection (Budhlakoti et al., 2019a). STGS is a comprehensive package which gives a single-step solution for genomic selection based on most commonly used statistical methods (i.e., RR, BLUP, LASSO, SVM, ANN, and RF). It is freely available through the CRAN server at https://CRAN.R-project.org/package=STGS.

MTGS

It is an R-based package developed for genomic predictions by estimating marker effects based on information available on multiple traits. Currently, STGS methods could not utilize additional information available when using multi-trait data. The package MTGS performs genomic selection using multi-trait information (Budhlakoti et al., 2019b). MTGS is a comprehensive package which gives a single-step solution for genomic selection using various MTGS-based methods (MRCE, MLASSO, i.e., multivariate LASSO, and KMLASSO, i.e., kernelized multivariate LASSO). It is freely available through the CRAN server at https://CRAN.R-project.org/package=MTGS.

FACTORS AFFECTING GENOMIC PREDICTION: EFFECTS OF MARKER DENSITY, POPULATION SIZE, TRAIT ARCHITECTURE, AND HERITABILITY

In general, increased marker density enhances the prediction accuracy using most of the GS models such as BLUP, LASSO, machine learning–based, or deep learning–based methods. However, there may be a chance of slow convergence in methods like Bayesian (Bayes A, Bayes B, Bayes Cπ, and Bayes Dπ), where convergence in terms of MCMC (i.e., Markov chain Monte Carlo) iteration is required (Arruda et al., 2016; Zhang et al., 2017; Norman et al., 2018; Zhang et al., 2019). Sometimes, low-density markers of a few hundreds to thousands also enable high prediction accuracies in breeding populations provided that there is a strong LD among the markers; however, it may be trait specific

and may vary with the architecture and heritability of studied traits (Lorenz et al., 2011; Werner et al., 2018). Also sometimes keeping a very high density of markers may have economic constraints as incorporation of such aspects into evaluation of GS strategies is also necessary for a profitable and efficient GS. Therefore, it is always difficult to give a benchmark for the number of markers to be used in such genomic studies; however, it is advisable to keep a moderate density, at least 2000 SNPs, so that prediction accuracy could not be significantly hampered (Abed et al., 2018). However, the cost of genotyping can also be significantly reduced by increasing the level of multiplexing without paying any penalty in terms of genomic prediction accuracy (e.g., genotyping a single line by GBS (96-plex) can cost 3.75 and 4.25 times less than using 9 K and 50 K arrays, respectively, in barley) (Abed et al., 2018). The position of SNPs and how they are placed in genomic arrangements over the chromosome may have a key role, for example, SNPs located in the intergenic space are slightly better at capturing the underlying haplotype diversity related to SNPs located in the genic space as the intergenic space is a playground of many important regulatory sequences, such as promoters and enhancers (Barrett et al., 2012; Abed et al., 2018). The use of high-quality SNP genotyping data (i.e., minor allele frequency (MAF)>0.1) could also be suggested to achieve a good prediction accuracy.

Population size has a significant role in the prediction accuracy whether it is conventional MAS or genomic selection, especially training population. If the population size or training population size is small, it is obvious that a decrease in accuracy is expected because the model will poorly estimate the marker effects and hence prediction accuracy. However, as an idea or estimate for the size of training population as 2*Ne*L (where N_e is the effective population size and L is the genome size in Morgan) and the number of markers as 10*Ne*L to achieve a prediction accuracy of 0.9 and reducing the size of the training population to $1*N_e*L$ results in a prediction accuracy of 0.7, provided that training population and breeding population are unrelated or both separated by many generations (Meuwissen, 2009). However, for most of the cases, training population and breeding population are related, so high genomic prediction accuracy could be achieved with a training population size much smaller than that referred above (Meuwissen, 2009).

Apart from these factors, prediction accuracy can also be affected by trait heritability especially for lower heritability ($h^2 < 0.4$) (Hayes et al., 2009). Numerous studies up-to-date showed that genomic selection accuracy is strongly influenced by trait heritability, i.e., the fraction of the phenotypic variance to the genetic variance of studied traits. Generally, it is assumed that the target trait with high heritability has good prediction accuracies and vice versa. However, as most of the agricultural traits have low to moderate heritability, it poses a challenge to genomic selection studies, especially in plants. However, low heritability traits would require a larger training population in order to attain the same prediction accuracy as in the case of traits with moderate to high heritability. However, to achieve this goal, sometimes cost may be a limiting factor, especially in developing countries. Moreover, it could be observed from the available literature that even for low heritable and complex traits, the performance of BLUP and its derivatives (e.g., GBLUP and RRBLUP), Bayesian methods (Bayes A, Bayes B, Bayes Cπ, and Bayes Dπ), and RKHS seems to be robust as compared to their counterparts (Crossa et al., 2010; Crossa et al., 2011; Heffner et al., 2011; Poland et al., 2012; Zhao et al., 2013; Spindel et al., 2015; Crossa et al., 2017; Wang et al., 2018; Xu et al., 2018; Juliana et al., 2019; Lozada et al., 2019; Michel et al., 2019), and at the same time, most of the models work fine with highly heritable traits, although the most suitable method is usually case-dependent. Sometimes missing observations also poses a challenge in estimating GEBV. However, the issue of low heritable trait and missing observation could be handled simultaneously, provided that data are available on multiple traits. In multiple traits, if we have few traits with low heritability and at the same time we have a good correlation with other highly heritable traits, i.e., by using the appropriate MTGS-based model, we can borrow information from other traits. In such scenarios, by using the MTGS model, we can estimate the GEBV more precisely and accurately.

CONCLUSION AND PROSPECTS

Genomic selection has shown its potential in plant and animal breeding research by increasing genetic gains in the last two decades. Revolution in terms of cheaper NGS technologies has made it possible to sequence the crop and animal genomes at a relatively low cost. It resulted in a number of completely sequenced crop and animal genomes with high-density SNP genotyping chips and their availability in the public domain, which may further boost the predictive ability of a GS model. Even after more than a decade in the field of genomic selection studies, still there is a lot of scope for improvement in this area. Methodological refinements (such as imputation of missing genotypic value, implementation of GxE interaction, information on epigenetic regulation, haplotypes, and including multi-trait information into prediction models) will be definitely helpful for a successful implementation of GS in plant and animal breeding programs. Consistent updation of the training set for GS is highly desirable by including the new markers in each generation. Evaluation of the training populations should be done in controlled and well-managed conditions as it significantly affects the performance of prediction models. There is a need for a structured program in the field of genomic selection including human resource development, advanced data recording methodologies, and trait phenotyping in order to come out with fruitful outcomes.

AUTHOR CONTRIBUTIONS

NB, AK, AR, and DM contributed to conceptualization. NB, AK, KC, AK, AP, RK, and UK reviewed and edited the paper. PJ, DM, and SK contributed to the final editing and correction. All authors contributed to the manuscript and approved the submitted version.

ACKNOWLEDGMENTS

We are grateful to the Government of India's DBT project "Germplasm Genomics and Trait Discovery in Wheat" and ICAR-CABin Scheme Network Project for the financial support to carry out this study.

REFERENCES

Abed, A., Pérez-Rodríguez, P., Crossa, J., and Belzile, F. (2018). When Less Can Be Better: How Can We Make Genomic Selection More Cost-Effective and Accurate in Barley? *Theor. Appl. Genet.* 131, 1873–1890. doi:10.1007/S00122-018-3120-8

Almeida, G. D., Makumbi, D., Magorokosho, C., Nair, S., Borém, A., Ribaut, J.-M., et al. (2013). QTL Mapping in Three Tropical maize Populations Reveals a Set of Constitutive and Adaptive Genomic Regions for Drought Tolerance. *Theor. Appl. Genet.* 126, 583–600. doi:10.1007/S00122-012-2003-7

Arelli, P. R., Young, L. D., and Concibido, V. C. (2009). Inheritance of Resistance in Soybean PI 567516C to LY1 Nematode Population Infecting Cv. Hartwig. *Euphytica* 165, 1–4. doi:10.1007/S10681-008-9760-Z

Arelli, P. R., Young, L. D., and Mengistu, A. (2006). Registration of High Yielding and Multiple Disease-Resistant Soybean Germplasm JTN-5503. *Crop Sci.* 46, 2723–2724. doi:10.2135/cropsci2005.12.0471crg

Arruda, M. P., Lipka, A. E., Brown, P. J., Krill, A. M., Thurber, C., Brown-Guedira, G., et al. (2016). Comparing Genomic Selection and Marker-Assisted Selection for Fusarium Head Blight Resistance in Wheat (*Triticum aestivum* L.). *Mol. Breed.* 36, 84. doi:10.1007/s11032-016-0508-5

Badu-Apraku, B., Talabi, A. O., Fakorede, M. A. B., Fasanmade, Y., Gedil, M., Magorokosho, C., et al. (2019). Yield Gains and Associated Changes in an Early Yellow Bi-parental maize Population Following Genomic Selection for Striga Resistance and Drought Tolerance. *BMC Plant Biol.* 19, 129. doi:10.1186/S12870-019-1740-Z

Baenziger, P. S. (2016). "Wheat Breeding and Genetics," in *Reference Module in Food Science*. Editor C. Beddows (Amsterdam, Netherlands: Elsevier). doi:10.1016/B978-0-08-100596-5.03001-8

Barrett, L. W., Fletcher, S., and Wilton, S. D. (2012). Regulation of Eukaryotic Gene Expression by the Untranslated Gene Regions and Other Non-coding Elements. *Cell. Mol. Life Sci.* 69, 3613–3634. doi:10.1007/S00018-012-0990-9

Battenfield, S. D., Guzmán, C., Gaynor, R. C., Singh, R. P., Peña, R. J., Dreisigacker, S., et al. (2016). Genomic Selection for Processing and End-Use Quality Traits in the CIMMYT Spring Bread Wheat Breeding Program. *Plant Genome* 9, 1–12. doi:10.3835/PLANTGENOME2016.01.0005

Beche, E., Gillman, J. D., Song, Q., Nelson, R., Beissinger, T., Decker, J., et al. (2021). Genomic Prediction Using Training Population Design in Interspecific Soybean Populations. *Mol. Breed.* 41, 1–15. doi:10.1007/S11032-021-01203-6

Ben Hassen, M., Bartholomé, J., Valè, G., Cao, T.-V., and Ahmadi, N. (2018). Genomic Prediction Accounting for Genotype by Environment Interaction Offers an Effective Framework for Breeding Simultaneously for Adaptation to an Abiotic Stress and Performance under normal Cropping Conditions in rice. *G3 Genes, Genomes, Genet.* 8, 2319–2332. doi:10.1534/g3.118.200098

Bernardo, R. (2008). Molecular Markers and Selection for Complex Traits in Plants: Learning from the Last 20 Years. *Crop Sci.* 48, 1649–1664. doi:10.2135/CROPSCI2008.03.0131

Bernardo, R., and Yu, J. (2007). Prospects for Genomewide Selection for Quantitative Traits in maize. *Crop Sci.* 47, 1082–1090. doi:10.2135/CROPSCI2006.11.0690

Beyene, Y., Semagn, K., Mugo, S., Tarekegne, A., Babu, R., Meisel, B., et al. (2015). Genetic Gains in Grain Yield through Genomic Selection in Eight Bi-parental Maize Populations under Drought Stress. *Crop Sci.* 55, 154–163. doi:10.2135/CROPSCI2014.07.0460

Bhandari, A., Bartholomé, J., Cao-Hamadoun, T.-V., Kumari, N., Frouin, J., Kumar, A., et al. (2019). Selection of Trait-specific Markers and Multi-Environment Models Improve Genomic Predictive Ability in rice. *PLOS ONE* 14, e0208871. doi:10.1371/JOURNAL.PONE.0208871

Budhlakoti, N., Mishra, D. C., Rai, A., and Chaturvedi, K. K. (2019a). Package 'STGS'. 1–11. Available at: https://cran.r-project.org/web/packages/STGS/STGS.pdf.

Budhlakoti, N., Mishra, D. C., and Rai, A. (2019b). Package 'MTGS'. 1–6.

Budhlakoti, N., Mishra, D. C., Rai, A., Lal, S. B., Chaturvedi, K. K., and Kumar, R. R. (2019c). A Comparative Study of Single-Trait and Multi-Trait Genomic Selection. *J. Comput. Biol.* 26, 1100–1112. doi:10.1089/CMB.2019.0032

Budhlakoti, N., Rai, A., and Mishra, D. C. (2020a). Effect of Influential Observation in Genomic Prediction Using LASSO Diagnostic. *Indian J. Agric. Sci.* 90, 1155–1159.

Budhlakoti, N., Rai, A., Mishra, D. C., Jaggi, S., Kumar, M., and Rao, A. R. (2020c). Comparative Study of Different Non-parametric Genomic Selection Methods under Diverse Genetic Architecture. *Ijgpb* 80, 395–401. doi:10.31742/IJGPB.80.4.4

Budhlakoti, N., Rai, A., and Mishra, D. C. (2020b). Statistical Approach for Improving Genomic Prediction Accuracy through Efficient Diagnostic Measure of Influential Observation. *Sci. Rep.* 10, 1–11. doi:10.1038/s41598-020-65323-3

Burgueño, J., de los Campos, G., Weigel, K., and Crossa, J. (2012). Genomic Prediction of Breeding Values when Modeling Genotype × Environment Interaction Using Pedigree and Dense Molecular Markers. *Crop Sci.* 52, 707–719. doi:10.2135/CROPSCI2011.06.0299

Cai, X., Huang, A., and Xu, S. (2011). Fast Empirical Bayesian LASSO for Multiple Quantitative Trait Locus Mapping. *BMC Bioinformatics* 12, 211–213. doi:10.1186/1471-2105-12-211/FIGURES/5

Cerrudo, D., Cao, S., Yuan, Y., Martinez, C., Suarez, E. A., Babu, R., et al. (2018). Genomic Selection Outperforms Marker Assisted Selection for Grain Yield and Physiological Traits in a maize Doubled Haploid Population across Water Treatments. *Front. Plant Sci.* 9, 366. doi:10.3389/FPLS.2018.00366/BIBTEX

Charmet, G., Tran, L.-G., Auzanneau, J., Rincent, R., and Bouchet, S. (2020). BWGS: A R Package for Genomic Selection and its Application to a Wheat Breeding Programme. *PLOS ONE* 15, e0222733. doi:10.1371/JOURNAL.PONE.0222733

Cheng, H., Kizilkaya, K., Zeng, J., Garrick, D., and Fernando, R. (2018). Genomic Prediction from Multiple-Trait Bayesian Regression Methods Using Mixture Priors. *Genetics* 209, 89–103. doi:10.1534/GENETICS.118.300650/-/DC1

Chiquet, J., Mary-Huard, T., Robin, S., and Robin, S. (2017). Structured Regularization for Conditional Gaussian Graphical Models. *Stat. Comput.* 27, 789–804. doi:10.1007/s11222-016-9654-1

Clark, S. A., Hickey, J. M., and Van Der Werf, J. H. (2011). Different Models of Genetic Variation and Their Effect on Genomic Evaluation. *Genet. Sel. Evol.* 43, 18. doi:10.1186/1297-9686-43-18

Cooper, M., Podlich, D. W., Micallef, K. P., Smith, O. S., Jensen, N. M., Chapman, S. C., et al. (2002). "Complexity, Quantitative Traits and Plant Breeding: a Role for Simulation Modelling in the Genetic Improvement of Crops," in *Quantitative Genetics, Genomics and Plant Breeding*. Editor M. S. Kang (Wallingford, UK: CAB International), 143–166. doi:10.1079/9780851996011.0143

Crossa, J., Campos, G. d. l., Pérez, P., Gianola, D., Burgueño, J., Araus, J. L., et al. (2010). Prediction of Genetic Values of Quantitative Traits in Plant Breeding Using Pedigree and Molecular Markers. *Genetics* 186, 713–724. doi:10.1534/GENETICS.110.118521

Crossa, J., Pérez, P., de los Campos, G., Mahuku, G., Dreisigacker, S., and Magorokosho, C. (2011). Genomic Selection and Prediction in Plant Breeding. *J. Crop Improvement* 25, 239–261. doi:10.1080/15427528.2011.558767

Crossa, J., Pérez, P., Hickey, J., Burgueño, J., Ornella, L., Cerón-Rojas, J., et al. (2014). Genomic Prediction in CIMMYT maize and Wheat Breeding Programs. *Heredity* 112, 48–60. doi:10.1038/HDY.2013.16

Crossa, J., Pérez-Rodríguez, P., Cuevas, J., Montesinos-López, O., Jarquín, D., de los Campos, G., et al. (2017). Genomic Selection in Plant Breeding: Methods, Models, and Perspectives. *Trends Plant Science* 22, 961–975. doi:10.1016/J.TPLANTS.2017.08.011

Cui, Z., Dong, H., Zhang, A., Ruan, Y., He, Y., and Zhang, Z. (2020). Assessment of the Potential for Genomic Selection to Improve Husk Traits in Maize. *G3: Genes, Genomes, Genet.* 10, 3741–3749. doi:10.1534/G3.120.401600

Daetwyler, H. D., Bansal, U. K., Bariana, H. S., Hayden, M. J., and Hayes, B. J. (2014). Genomic Prediction for Rust Resistance in Diverse Wheat Landraces. *Theor. Appl. Genet.* 127, 1795–1803. doi:10.1007/s00122-014-2341-8

Daetwyler, H. D., Hickey, J. M., Henshall, J. M., Dominik, S., Gredler, B., van der Werf, J. H. J., et al. (2010). Accuracy of Estimated Genomic Breeding Values for Wool and Meat Traits in a Multi-Breed Sheep Population. *Anim. Prod. Sci.* 50, 1004–1010. doi:10.1071/AN10096

Das, R. R., Vinayan, M. T., Patel, M. B., Phagna, R. K., Singh, S. B., Shahi, J. P., et al. (2020). Genetic Gains with Rapid-cycle Genomic Selection for Combined Drought and Waterlogging Tolerance in Tropical maize (Zea May S L.). *Plant Genome* 13, 1–15. doi:10.1002/tpg2.20035

de los Campos, G., Hickey, J. M., Pong-Wong, R., Daetwyler, H. D., and Calus, M. P. L. (2013). Whole-Genome Regression and Prediction Methods Applied to Plant and Animal Breeding. *Genetics* 193, 327–345. doi:10.1534/GENETICS.112.143313

De Roos, A. P. W., Hayes, B. J., and Goddard, M. E. (2009). Reliability of Genomic Predictions across Multiple Populations. *Genetics* 183, 1545–1553. doi:10.1534/GENETICS.109.104935

Diaz, S., Ariza-Suarez, D., Ramdeen, R., Aparicio, J., Arunachalam, N., Hernandez, C., et al. (2021). Genetic Architecture and Genomic Prediction of Cooking Time in Common Bean (Phaseolus vulgaris L.). *Front. Plant Sci.* 11, 2257. doi:10.3389/FPLS.2020.622213/BIBTEX

dos Santos, J. P. R., Pires, L. P. M., de Castro Vasconcellos, R. C., Pereira, G. S., von Pinho, R. G., and Balestre, M. (2016). Genomic Selection to Resistance to Stenocarpella Maydis in maize Lines Using DArTseq Markers. *BMC Genet.* 17, 86. doi:10.1186/S12863-016-0392-3

Duangjit, J., Causse, M., and Sauvage, C. (2016). Efficiency of Genomic Selection for Tomato Fruit Quality. *Mol. Breed.* 36, 29. doi:10.1007/S11032-016-0453-3

Endelman, J. B. (2011). Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. *The Plant Genome* 4, 250–255. doi:10.3835/PLANTGENOME2011.08.0024

Fernando, R., and Garrick, D. (2009). GenSel- User Manual for a Portfolio of Genomic Selection Related Analyses. Available at: http://taurus.ansci.iastate.edu/Site/Welcome_files/GenSel%20 Manual%20v2.pdf.

Fernando, R., and Grossman, M. (1989). Marker Assisted Selection Using Best Linear Unbiased Prediction. *Genet. Selection Evol.* 21 (421), 467–477. doi:10.1186/1297-9686-21-4-467

Fikere, M., Barbulescu, D. M., Malmberg, M. M., Maharjan, P., Salisbury, P. A., Kant, S., et al. (2020). Genomic Prediction and Genetic Correlation of Agronomic, Blackleg Disease, and Seed Quality Traits in Canola (Brassica Napus L.). *Plants* 9, 719–19. doi:10.3390/PLANTS9060719

Fones, H. N., Bebber, D. P., Chaloner, T. M., Kay, W. T., Steinberg, G., and Gurr, S. J. (2020). Threats to Global Food Security from Emerging Fungal and Oomycete Crop Pathogens. *Nat. Food* 1, 332–342. doi:10.1038/s43016-020-0075-0

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *J. Stat. Softw.* 33, 1–22. doi:10.18637/jss.v033.i01

Fristche-Neto, R., Akdemir, D., and Jannink, J.-L. (2018). Accuracy of Genomic Selection to Predict maize Single-Crosses Obtained through Different Mating Designs. *Theor. Appl. Genet.* 131, 1153–1162. doi:10.1007/s00122-018-3068-8

Furbank, R. T., and Tester, M. (2011). Phenomics - Technologies to Relieve the Phenotyping Bottleneck. *Trends Plant Sci.* 16, 635–644. doi:10.1016/J.TPLANTS.2011.09.005

Gianola, D., Fernando, R. L., and Stella, A. (2006). Genomic-Assisted Prediction of Genetic Value with Semiparametric Procedures. *Genetics* 173, 1761–1776. doi:10.1534/GENETICS.105.049510

Gianola, D., Okut, H., Weigel, K. A., and Rosa, G. J. (2011). Predicting Complex Quantitative Traits with Bayesian Neural Networks: a Case Study with Jersey Cows and Wheat. *BMC Genet.* 12, 87. doi:10.1186/1471-2156-12-87

Gopalakrishnan, S., Sharma, R. K., Anand Rajkumar, K., Joseph, M., Singh, V. P., Singh, A. K., et al. (2008). Integrating Marker Assisted Background Analysis with Foreground Selection for Identification of superior Bacterial Blight Resistant Recombinants in Basmati rice. *Plant Breed.* 127, 131–139. doi:10.1111/J.1439-0523.2007.01458.X

Grenier, C., Cao, T.-V., Ospina, Y., Quintero, C., Châtel, M. H., Tohme, J., et al. (2015). Accuracy of Genomic Selection in a Rice Synthetic Population Developed for Recurrent Selection Breeding. *PloS one* 10, e0136594. doi:10.1371/JOURNAL.PONE.0136594

Habier, D., Fernando, R. L., Kizilkaya, K., and Garrick, D. J. (2011). Extension of the Bayesian Alphabet for Genomic Selection. *BMC Bioinformatics* 12, 186–197. doi:10.1186/1471-2105-12-186/FIGURES/2

Habier, D., Fernando, R. L., and Dekkers, J. C. M. (2009). Genomic Selection Using Low-Density Marker Panels. *Genetics* 182, 343–353. doi:10.1534/GENETICS.108.100289

Habier, D., Fernando, R. L., and Garrick, D. J. (2013). Genomic BLUP Decoded: a Look into the Black Box of Genomic Prediction. *Genetics* 194, 597–607. doi:10.1534/GENETICS.113.152207

Haile, T. A., Heidecker, T., Wright, D., Neupane, S., Ramsay, L., Vandenberg, A., et al. (2020). Genomic Selection for Lentil Breeding: Empirical Evidence. *Plant Genome* 13, 1–30. doi:10.1002/tpg2.20002

Hayes, B. J., Bowman, P. J., Chamberlain, A. J., and Goddard, M. E. (2009). Invited Review: Genomic Selection in Dairy Cattle: Progress and Challenges. *J. Dairy Sci.* 92, 433–443. doi:10.3168/JDS.2008-1646

Heffner, E. L., Jannink, J.-L., Sorrells, M. E., Heffner, E. L., Sorrells, M. E., Univ, C., et al. (2011). Genomic Selection Accuracy Using Multifamily Prediction Models in a Wheat Breeding Program. *The Plant Genome* 4, 65–75. doi:10.3835/plantgenome.2010.12.0029

Heffner, E. L., Lorenz, A. J., Jannink, J. L., and Sorrells, M. E. (2010). Plant Breeding with Genomic Selection: Gain Per Unit Time and Cost. *Crop Sci.* 50, 1681–1690. doi:10.2135/cropsci2009.11.0662

Heffner, E. L., Sorrells, M. E., and Jannink, J.-L. (2009). Genomic Selection for Crop Improvement. *Crop Sci.* 49, 1–12. doi:10.2135/CROPSCI2008.08.0512

Henderson, C. R., Kempthorne, O., Searle, S. R., and von Krosigk, C. M. (1959). The Estimation of Environmental and Genetic Trends from Records Subject to Culling. *Biometrics* 15, 192. doi:10.2307/2527669

Holliday, J. A., Wang, T., and Aitken, S. (2012). Predicting Adaptive Phenotypes from Multilocus Genotypes in Sitka Spruce (Picea Sitchensis) Using Random Forest. *Using Random For.* 2, 1085–1093. doi:10.1534/g3.112.002733

Hong, J. P., Ro, N., Lee, H. Y., Kim, G. W., Kwon, J. K., Yamamoto, E., et al. (2020). Genomic Selection for Prediction of Fruit-Related Traits in Pepper (Capsicum spp.). *Front. Plant Sci.* 11, 570871. doi:10.3389/FPLS.2020.570871/BIBTEX

Howard, R., Carriquiry, A. L., and Beavis, W. D. (2014). Parametric and Nonparametric Statistical Methods for Genomic Selection of Traits with Additive and Epistatic Genetic Architectures. *G3: Genes, Genomes, Genet.* 4, 1027–1046. doi:10.1534/G3.114.010298/-/DC1

Hu, Z., Li, Y., Song, X., Han, Y., Cai, X., Xu, S., et al. (2011). Genomic Value Prediction for Quantitative Traits under the Epistatic Model. *BMC Genet.* 12, 15. doi:10.1186/1471-2156-12-15

Huang, M., Balimponya, E. G., Mgonja, E. M., McHale, L. K., Luzi-Kihupi, A., Wang, G.-L., et al. (2019). Use of Genomic Selection in Breeding rice (Oryza Sativa L.) for Resistance to rice Blast (Magnaporthe Oryzae). *Mol. Breed.* 39, 1–16. doi:10.1007/S11032-019-1023-2

Imai, A., Kuniga, T., Yoshioka, T., Nonaka, K., Mitani, N., Fukamachi, H., et al. (2019). Single-step Genomic Prediction of Fruit-Quality Traits Using Phenotypic Records of Non-genotyped Relatives in Citrus. *PLOS ONE* 14, e0221880. doi:10.1371/JOURNAL.PONE.0221880

Jan, H. U., Abbadi, A., Lücke, S., Nichols, R. A., and Snowdon, R. J. (2016). Genomic Prediction of Testcross Performance in Canola (Brassica Napus). *PLOS ONE* 11, e0147769. doi:10.1371/JOURNAL.PONE.0147769

Janila, P., Variath, M. T., Pandey, M. K., Desmae, H., Motagi, B. N., Okori, P., et al. (2016). Genomic Tools in Groundnut Breeding Program: Status and Perspectives. *Front. Plant Sci.* 7, 289. doi:10.3389/FPLS.2016.00289/BIBTEX

Jeong, S., Kim, J.-Y., and Kim, N. (2020). GMStool: GWAS-Based Marker Selection Tool for Genomic Prediction from Genomic Data. *Sci. Rep.* 10, 1–12. doi:10.1038/s41598-020-76759-y

Jia, Y., and Jannink, J.-L. (2012). Multiple-Trait Genomic Selection Methods Increase Genetic Value Prediction Accuracy. *Genetics* 192, 1513–1522. doi:10.1534/GENETICS.112.144246

Joukhadar, R., Thistlethwaite, R., Trethowan, R. M., Hayden, M. J., Stangoulis, J., Cu, S., et al. (2021). Genomic Selection Can Accelerate the Biofortification of spring Wheat. *Theor. Appl. Genet.* 134, 3339–3350. doi:10.1007/S00122-021-03900-4

Juarez, M., Legua, P., Mengual, C. M., Kassem, M. A., Sempere, R. N., Gómez, P., et al. (2013). Relative Incidence, Spatial Distribution and Genetic Diversity of Cucurbit Viruses in Eastern Spain. *Ann. Appl. Biol.* 162, 362–370. doi:10.1111/AAB.12029

Juliana, P., Poland, J., Huerta-Espino, J., Shrestha, S., Crossa, J., Crespo-Herrera, L., et al. (2019). Improving Grain Yield, Stress Resilience and Quality of Bread Wheat Using Large-Scale Genomics. *Nat. Genet.* 51, 1530–1539. doi:10.1038/s41588-019-0496-6

Juliana, P., Singh, R. P., Singh, P. K., Crossa, J., Rutkoski, J. E., Poland, J. A., et al. (2017). Comparison of Models and Whole-Genome Profiling Approaches for Genomic-Enabled Prediction of Septoria Tritici Blotch, Stagonospora Nodorum Blotch, and Tan Spot Resistance in Wheat. *Plant Genome* 10, 1–16. doi:10.3835/PLANTGENOME2016.08.0082

Klápště, J., Dungey, H. S., Telfer, E. J., Suontama, M., Graham, N. J., Li, Y., et al. (2020). Marker Selection in Multivariate Genomic Prediction Improves Accuracy of Low Heritability Traits. *Front. Genet.* 11, 499094. doi:10.3389/FGENE.2020.499094/FULL

Krishnan, S. G., Singh, A. K., Rathour, R., Nagarajan, M., Bhowmick, P. K., Ellur, R. K., et al. (2019). Rice Variety Pusa Samba 1850. *Indian J. Genet.* 79, 109–110.

Kumar, S., Chagné, D., Bink, M. C. A. M., Volz, R. K., Whitworth, C., and Carlisle, C. (2012). Genomic Selection for Fruit Quality Traits in Apple (Malus×domestica Borkh.). *PLoS One* 7, e36674. doi:10.1371/JOURNAL.PONE.0036674

Kumar, S., Kirk, C., Deng, C. H., Shirtliff, A., Wiedow, C., Qin, M., et al. (2019). Marker-trait Associations and Genomic Predictions of Interspecific Pear (Pyrus) Fruit Characteristics. *Sci. Rep.* 9, 1–10. doi:10.1038/s41598-019-45618-w

Lam, H.-M., Coschigano, K. T., Oliveira, I. C., Melo-Oliveira, R., and Coruzzi, G. M. (1996). The Molecular-Genetics of Nitrogen Assimilation into Amino Acids in Higher Plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 47, 569–593. doi:10.1146/annurev.arplant.47.1.569

Legarra, A., and Reverter, A. (2018). Semi-parametric Estimates of Population Accuracy and Bias of Predictions of Breeding Values and Future Phenotypes Using the LR Method. *Genet. Sel Evol.* 50, 53–18. doi:10.1186/S12711-018-0426-6/FIGURES/3

Li, Y., Ruperao, P., Batley, J., Edwards, D., Khan, T., Colmer, T. D., et al. (2018). Investigating Drought Tolerance in Chickpea Using Genome-wide Association Mapping and Genomic Selection Based on Whole-Genome Resequencing Data. *Front. Plant Sci.* 9, 190. doi:10.3389/FPLS.2018.00190/BIBTEX

Liu, B., Asseng, S., Müller, C., Ewert, F., Elliott, J., Lobell, D. B., et al. (2016). Similar Estimates of Temperature Impacts on Global Wheat Yield by Three Independent Methods. *Nat. Clim Change* 6, 1130–1136. doi:10.1038/NCLIMATE3115

Liu, X., Wang, H., Wang, H., Guo, Z., Xu, X., Liu, J., et al. (2018). Factors Affecting Genomic Selection Revealed by Empirical Evidence in maize. *Crop J.* 6, 341–352. doi:10.1016/J.CJ.2018.03.005

Long, N., Gianola, D., Rosa, G. J. M., and Weigel, K. A. (2011). Application of Support Vector Regression to Genome-Assisted Prediction of Quantitative Traits. *Theor. Appl. Genet.* 123, 1065–1074. doi:10.1007/S00122-011-1648-Y

Longin, C. F. H., Reif, J. C., and Würschum, T. (2014). Long-term Perspective of Hybrid versus Line Breeding in Wheat Based on Quantitative Genetic Theory. *Theor. Appl. Genet.* 127, 1635–1641. doi:10.1007/S00122-014-2325-8

Lorenz, A. J., Chao, S., Asoro, F. G., Heffner, E. L., Hayashi, T., Iwata, H., et al. (2011). Genomic Selection in Plant Breeding. *Adv. Agron.* 110, 77–123. doi:10.1016/B978-0-12-385531-2.00002-5

Lorenz, A. J., Smith, K. P., and Jannink, J. L. (2012). Potential and Optimization of Genomic Selection for Fusarium Head Blight Resistance in Six-Row Barley. *Crop Sci.* 52, 1609–1621. doi:10.2135/cropsci2011.09.0503

Lorenzana, R. E., and Bernardo, R. (2009). Accuracy of Genotypic Value Predictions for Marker-Based Selection in Biparental Plant Populations. *Theor. Appl. Genet.* 120, 151–161. doi:10.1007/s00122-009-1166-3

Lozada, D. N., and Carter, A. H. (2020). Genomic Selection in Winter Wheat Breeding Using a Recommender Approach. *Genes* 11, 1–14. doi:10.3390/GENES11070779

Lozada, D. N., Mason, R. E., Sarinelli, J. M., and Brown-Guedira, G. (2019). Accuracy of Genomic Selection for Grain Yield and Agronomic Traits in Soft Red winter Wheat. *BMC Genet.* 20, 1–12. doi:10.1186/s12863-019-0785-1

Luan, T., Woolliams, J. A., Lien, S., Kent, M., Svendsen, M., and Meuwissen, T. H. E. (2009). The Accuracy of Genomic Selection in Norwegian Red Cattle Assessed by Cross-Validation. *Genetics* 183, 1119–1126. doi:10.1534/GENETICS.109.107391

Lush, J. L. (1943). *Animal Breeding Plans*. Edn 2. Charleston, South Carolina: Bibliolife DBA of Bibilio Bazaar.

Mackay, T. F. C. (2001). The Genetic Architecture of Quantitative Traits. *Annu. Rev. Genet.* 35, 303–339. doi:10.1146/annurev.genet.35.102401.090633

Madhavi, K. R., Rambabu, R., Abhilash Kumar, V., Vijay Kumar, S., Aruna, J., Ramesh, S., et al. (2016). Marker Assisted Introgression of Blast (Pi-2 and Pi-54) Genes in to the Genetic Background of Elite, Bacterial Blight Resistant Indica rice Variety, Improved Samba Mahsuri. *Euphytica* 212, 331–342. doi:10.1007/S10681-016-1784-1

Maenhout, S., De Baets, B., Haesaert, G., and Van Bockstaele, E. (2007). Support Vector Machine Regression for the Prediction of maize Hybrid Performance. *Theor. Appl. Genet.* 115, 1003–1013. doi:10.1007/s00122-007-0627-9

Majumdar, S. G., Rai, A., and Mishra, D. C. (2020). Integrated Framework for Selection of Additive and Nonadditive Genetic Markers for Genomic Selection. *J. Comput. Biol.* 27, 845–855. doi:10.1089/CMB.2019.0223

Majumdar, S. G., Rai, A., and Mishra, D. C. (2019). Package 'GSelection', 1–14. Available at: https://rdrr.io/cran/GSelection/man/GSelection-package.html.

Mangin, B., Bonnafous, F., Blanchet, N., Boniface, M.-C., Bret-Mestries, E., Carrère, S., et al. (2017). Genomic Prediction of sunflower Hybrids Oil Content. *Front. Plant Sci.* 8, 1633. doi:10.3389/FPLS.2017.01633/BIBTEX

Marulanda, J. J., Mi, X., Melchinger, A. E., Xu, J.-L., Würschum, T., and Longin, C. F. H. (2016). Optimum Breeding Strategies Using Genomic Selection for Hybrid Breeding in Wheat, maize, rye, Barley, rice and Triticale. *Theor. Appl. Genet.* 129, 1901–1913. doi:10.1007/s00122-016-2748-5

Meuwissen, T. H. (2009). Accuracy of Breeding Values of 'unrelated' Individuals Predicted by Dense SNP Genotyping. *Genet. Sel Evol.* 41, 35–39. doi:10.1186/1297-9686-41-35/TABLES/3

Meuwissen, T. H. E., Hayes, B. J., and Goddard, M. E. (2001). Prediction of Total Genetic Value Using Genome-wide Dense Marker Maps. *Genetics* 157, 1819–1829. doi:10.1093/GENETICS/157.4.1819

Michel, S., Löschenberger, F., Ametz, C., Pachler, B., Sparry, E., and Bürstmayr, H. (2019). Combining Grain Yield, Protein Content and Protein Quality by Multi-Trait Genomic Selection in Bread Wheat. *Theor. Appl. Genet.* 132, 2767–2780. doi:10.1007/S00122-019-03386-1

Minamikawa, M. F., Nonaka, K., Kaminuma, E., Kajiya-Kanegae, H., Onogi, A., Goto, S., et al. (2017). Genome-wide Association Study and Genomic Prediction in Citrus: Potential of Genomics-Assisted Breeding for Fruit Quality Traits. *Sci. Rep.* 7, 1–13. doi:10.1038/s41598-017-05100-x

Mirdita, V., He, S., Zhao, Y., Korzun, V., Bothe, R., Ebmeyer, E., et al. (2015). Potential and Limits of Whole Genome Prediction of Resistance to Fusarium Head Blight and Septoria Tritici Blotch in a Vast Central European Elite winter Wheat Population. *Theor. Appl. Genet.* 128, 2471–2481. doi:10.1007/S00122-015-2602-1

Mishra, D. C., Budhlakoti, N., Majumdar, S. G., and Rai, A. (2021). Innovations in Genomic Selection : Statistical Perspective. 101–111. Available at: https://ssca.org.in/media/9_Spl_Proceedings_2021_006072021_Dwijesh_Mishra_Final.pdf.

Mohan, M., Nair, S., Bhagwat, A., Krishna, T. G., Yano, M., Bhatia, C. R., et al. (1997). Genome Mapping, Molecular Markers and Marker-Assisted Selection in Crop Plants. *Mol. Breed.* 3, 87–103. doi:10.1023/A:1009651919792

Moore, J. H., and Williams, S. M. (2009). Epistasis and its Implications for Personal Genetics. *Am. J. Hum. Genet.* 85, 309–320. doi:10.1016/J.AJHG.2009.08.006

Nakaya, A., and Isobe, S. N. (2012). Will Genomic Selection Be a Practical Method for Plant Breeding? *Ann. Bot.* 110, 1303–1316. doi:10.1093/AOB/MCS109

Neeraja, C. N., Maghirang-Rodriguez, R., Pamplona, A., Heuer, S., Collard, B. C. Y., Septiningsih, E. M., et al. (2007). A Marker-Assisted Backcross Approach for Developing Submergence-Tolerant rice Cultivars. *Theor. Appl. Genet.* 115 (6), 767–776. doi:10.1007/s00122-007-0607-0

Norman, A., Taylor, J., Edwards, J., and Kuchel, H. (2018). Optimising Genomic Selection in Wheat: Effect of Marker Density, Population Size and Population

Structure on Prediction Accuracy. *G3 Genes|Genomes|Genetics* 8, 2889–2899. doi:10.1534/G3.118.200311

Nsibi, M., Gouble, B., Bureau, S., Flutre, T., Sauvage, C., Audergon, J.-M., et al. (2020). Adoption and Optimization of Genomic Selection to Sustain Breeding for Apricot Fruit Quality. *G3 Genes|Genomes|Genetics* 10, 4513–4529. doi:10.1534/G3.120.401452

Ogutu, J. O., Schulz-Streeck, T., and Piepho, H.-P. (2012). Genomic Selection Using Regularized Linear Regression Models: ridge Regression, Lasso, Elastic Net and Their Extensions. *BMC Proc.* 6, S10. doi:10.1186/1753-6561-6-S2-S10

Onogi, A., Ideta, O., Inoshita, Y., Ebana, K., Yoshioka, T., Yamasaki, M., et al. (2015). Exploring the Areas of Applicability of Whole-Genome Prediction Methods for Asian rice (Oryza Sativa L.). *Theor. Appl. Genet.* 128, 41–53. doi:10.1007/S00122-014-2411-Y

Pandey, M. K., Chaudhari, S., Jarquin, D., Janila, P., Crossa, J., Patil, S. C., et al. (2020). Genome-based Trait Prediction in Multi- Environment Breeding Trials in Groundnut. *Theor. Appl. Genet.* 133, 3101–3117. doi:10.1007/S00122-020-03658-1/TABLES/5

Pérez, P., and de los Campos, G. (2014). Genome-Wide Regression and Prediction with the BGLR Statistical Package. *Genetics* 198, 483–495. doi:10.1534/genetics.114.164442

Pérez-Rodríguez, P., Gianola, D., González-Camacho, J. M., Crossa, J., Manès, Y., and Dreisigacker, S. (2012). Comparison between Linear and Non-parametric Regression Models for Genome-Enabled Prediction in Wheat. *G3: Genes, Genomes, Genet.* 2, 1595–1605. doi:10.1534/G3.112.003665/-/DC1

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., Heisterkamp, S., and Van Willigen, B. (2017). Package 'nlme'. Available at: https://cran.r-project.org/web/packages/nlme/nlme.pdf.

Poland, J., Endelman, J., Dawson, J., Rutkoski, J., Wu, S., Manes, Y., et al. (2012). Genomic Selection in Wheat Breeding Using Genotyping-by-Sequencing. *Plant Genome* 5, 1–11. doi:10.3835/PLANTGENOME2012.06.0006

Qin, F., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2011). Achievements and Challenges in Understanding Plant Abiotic Stress Responses and Tolerance. *Plant Cel Physiol.* 52, 1569–1582. doi:10.1093/PCP/PCR106

Rai, K. N., Hash, C. T., Singh, A. K., and Velu, G. (2008). Adaptation and Quality Traits of a Germplasm-Derived Commercial Seed Parent of Pearl Millet. *Plant Genet. Resour. Newsl.* 154, 20–24.

Reif, J. C., Zhao, Y., Würschum, T., Gowda, M., and Hahn, V. (2013). Genomic Prediction of sunflower Hybrid Performance. *Plant Breed* 132, 107–114. doi:10.1111/pbr.12007

Reynolds, M. P., and Ortiz, R. (2010). "Adapting Crops to Climate Change: a Summary," in *Climate Change and Crop Production*. Editor M. P. Reynolds (Wallingford,UK: CABI), 1–8. doi:10.1079/9781845936334.0001

Ribaut, J.-M., and Ragot, M. (2007). Marker-assisted Selection to Improve Drought Adaptation in maize: the Backcross Approach, Perspectives, Limitations, and Alternatives. *J. Exp. Bot.* 58, 351–360. doi:10.1093/JXB/ERL214

Rio, S., Mary-Huard, T., Moreau, L., and Charcosset, A. (2019). Genomic Selection Efficiency and A Priori Estimation of Accuracy in a Structured Dent maize Panel. *Theor. Appl. Genet.* 132, 81–96. doi:10.1007/s00122-018-3196-1

Roth, M., Muranty, H., Di Guardo, M., Guerra, W., Patocchi, A., and Costa, F. (2020). Genomic Prediction of Fruit Texture and Training Population Optimization towards the Application of Genomic Selection in Apple. *Hortic. Res.* 7, 1–14. doi:10.1038/s41438-020-00370-5

Rothman, A. J., Levina, E., and Zhu, J. (2010). Sparse Multivariate Regression with Covariance Estimation. *J. Comput. Graphical Stat.* 19, 947–962. doi:10.1198/JCGS.2010.09188

Rutkoski, J., Benson, J., Jia, Y., Brown-Guedira, G., Jannink, J.-L., and Sorrells, M. (2012). Evaluation of Genomic Prediction Methods for Fusarium Head Blight Resistance in Wheat. *The Plant Genome* 5, 51–61. doi:10.3835/PLANTGENOME2012.02.0001

Rutkoski, J., Poland, J., Mondal, S., Autrique, E., Pérez, L. G., Crossa, J., et al. (2016). Canopy Temperature and Vegetation Indices from High-Throughput Phenotyping Improve Accuracy of Pedigree and Genomic Selection for Grain Yield in Wheat. *G3: Genes, Genomes, Genet.* 6, 2799–2808. doi:10.1534/G3.116.032888

Sallam, A. H., Endelman, J. B., Jannink, J. L., and Smith, K. P. (2015). Assessing Genomic Selection Prediction Accuracy in a Dynamic Barley Breeding Population. *Plant Genome* 8, 20. doi:10.3835/PLANTGENOME2014.05.0020

Sallam, A. H., and Smith, K. P. (2016). Genomic Selection Performs Similarly to Phenotypic Selection in Barley. *Crop Sci.* 56, 2871–2881. doi:10.2135/CROPSCI2015.09.0557

Sarinelli, J. M., Murphy, J. P., Tyagi, P., Holland, J. B., Johnson, J. W., Mergoum, M., et al. (2019). Training Population Selection and Use of Fixed Effects to Optimize Genomic Predictions in a Historical USA winter Wheat Panel. *Theor. Appl. Genet.* 132, 1247–1261. doi:10.1007/S00122-019-03276-6

Schmidt, M., Kollers, S., Maasberg-Prelle, A., Großer, J., Schinkel, B., Tomerius, A., et al. (2016). Prediction of Malting Quality Traits in Barley Based on Genome-wide Marker Data to Assess the Potential of Genomic Selection. *Theor. Appl. Genet.* 129, 203–213. doi:10.1007/s00122-015-2639-1

Sehgal, D., Rosyara, U., Mondal, S., Singh, R., Poland, J., and Dreisigacker, S. (2020). Incorporating Genome-wide Association Mapping Results into Genomic Prediction Models for Grain Yield and Yield Stability in CIMMYT Spring Bread Wheat. *Front. Plant Sci.* 11, 197. doi:10.3389/FPLS.2020.00197

Semagn, K., Beyene, Y., Warburton, M. L., Tarekegne, A., Mugo, S., Meisel, B., et al. (2013). Meta-analyses of QTL for Grain Yield and Anthesis Silking Interval in 18 maize Populations Evaluated under Water-Stressed and Well-Watered Environments. *BMC Genomics* 14, 313–316. doi:10.1186/1471-2164-14-313/TABLES/4

Shikha, M., Kanika, A., Rao, A. R., Mallikarjuna, M. G., Gupta, H. S., and Nepolean, T. (2017). Genomic Selection for Drought Tolerance Using Genome-wide SNPs in Maize. *Front. Plant Sci.* 8, 1–12. doi:10.3389/fpls.2017.00550

Singh, A. K., Gopala Krishnan, S., Ellur, R. K., Bhowmick, P. K., Nagarajan, M., Vinod, K. K., et al. (2017a). Notification of Basmati rice Variety, Pusa Basmati 1728. *Indian J. Genet.* 77, 584.

Singh, A. K., Gopala Krishnan, S., Nagarajan, M., Bhowmick, P. K., Ellur, R. K., Haritha, B., et al. (2017b). Notification of Basmati rice Variety Pusa Basmati 1637. *Indian J. Genet.* 77, 583–584.

Spindel, J., Begum, H., Akdemir, D., Virk, P., Collard, B., Redoña, E., et al. (2015). Correction: Genomic Selection and Association Mapping in Rice (Oryza Sativa): Effect of Trait Genetic Architecture, Training Population Composition, Marker Number and Statistical Model on Accuracy of Rice Genomic Selection in Elite, Tropical Rice Breeding Lines. *Plos Genet.* 11, e1005350. doi:10.1371/JOURNAL.PGEN.1005350

Stewart-Brown, B. B., Song, Q., Vaughn, J. N., and Li, Z. (2019). Genomic Selection for Yield and Seed Composition Traits within an Applied Soybean Breeding Program. *G3 Genes, Genomes, Genet.* 9, 2253–2265. doi:10.1534/g3.118.200917

Sukumaran, S., Jarquin, D., Crossa, J., and Reynolds, M. (2018). Genomic-enabled Prediction Accuracies Increased by Modeling Genotype × Environment Interaction in Durum Wheat. *Plant Genome* 11, 170112. doi:10.3835/PLANTGENOME2017.12.0112

Sun, J., Poland, J. A., Mondal, S., Crossa, J., Juliana, P., Singh, R. P., et al. (2019). High-throughput Phenotyping Platforms Enhance Genomic Selection for Wheat Grain Yield across Populations and Cycles in Early Stage. *Theor. Appl. Genet.* 132, 1705–1720. doi:10.1007/S00122-019-03309-0

Sun, J., Rutkoski, J. E., Poland, J. A., Crossa, J., Jannink, J. L., and Sorrells, M. E. (2017). Multitrait, Random Regression, or Simple Repeatability Model in High-Throughput Phenotyping Data Improve Genomic Prediction for Wheat Grain Yield. *Plant Genome* 10, 1–12. doi:10.3835/PLANTGENOME2016.11.0111

Tanaka, E. (2020). Simple Outlier Detection for a Multi-environmental Field Trial. *Biometrics* 76, 1374–1382. doi:10.1111/BIOM.13216

Tanaka, E. (2018). Simple Robust Genomic Prediction and Outlier Detection for a Multi-Environmental Field Trial. arXiv preprint arXiv:1807.07268, 1–25.

Tester, M., and Langridge, P. (2010). Breeding Technologies to Increase Crop Production in a Changing World. *Science* 327, 818–822. doi:10.1126/SCIENCE.1183700

Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *J. R. Stat. Soc. Ser. B (Methodological)* 58, 267–288. doi:10.1111/J.2517-6161.1996.TB02080.X

Tiede, T., and Smith, K. P. (2018). Evaluation and Retrospective Optimization of Genomic Selection for Yield and Disease Resistance in spring Barley. *Mol. Breed.* 38, 55. doi:10.1007/S11032-018-0820-3

Usai, M. G., Goddard, M. E., and Hayes, B. J. (2009). LASSO with Cross-Validation for Genomic Selection. *Genet. Res.* 91, 427–436. doi:10.1017/S0016672309990334

VanRaden, P. M. (2008). Efficient Methods to Compute Genomic Predictions. *J. Dairy Sci.* 91, 4414–4423. doi:10.3168/JDS.2007-0980

Varshney, R. K., Mohan, S. M., Gaur, P. M., Chamarthi, S. K., Singh, V. K., Srinivasan, S., et al. (2014a). Marker-Assisted Backcrossing to Introgress Resistance to Fusarium Wilt Race 1 and Ascochyta Blight in C 214, an Elite Cultivar of Chickpea. *Plant Genome* 7, 35. doi:10.3835/plantgenome2013.10.0035

Varshney, R. K., Pandey, M. K., Janila, P., Nigam, S. N., Sudini, H., Gowda, M. V. C., et al. (2014b). Marker-assisted Introgression of a QTL Region to Improve Rust Resistance in Three Elite and Popular Varieties of Peanut (*Arachis hypogaea* L.). *Theor. Appl. Genet.* 127, 1771–1781. doi:10.1007/S00122-014-2338-3

Varshney, R. K., Singh, V. K., Kumar, A., Powell, W., and Sorrells, M. E. (2018). Can Genomics Deliver Climate-Change Ready Crops? *Curr. Opin. Plant Biol.* 45, 205–211. doi:10.1016/J.PBI.2018.03.007

Váry, Z., Mullins, E., Mcelwain, J. C., and Doohan, F. M. (2015). The Severity of Wheat Diseases Increases when Plants and Pathogens Are Acclimatized to Elevated Carbon Dioxide. *Glob. Change Biol.* 21, 2661–2669. doi:10.1111/GCB.12899

Vasistha, N. K., Balasubramaniam, A., Mishra, V. K., Srinivasa, J., Chand, R., and Joshi, A. K. (2017). Molecular Introgression of Leaf Rust Resistance Gene Lr34 Validates Enhanced Effect on Resistance to Spot Blotch in spring Wheat. *Euphytica* 213, 1–10. doi:10.1007/s10681-017-2051-9

Vazquez, A. I., Bates, D. M., Rosa, G. J. M., Gianola, D., and Weigel, K. A. (2010). Technical Note: An R Package for Fitting Generalized Linear Mixed Models in Animal Breeding1. *J. Anim. Sci.* 88, 497–504. doi:10.2527/JAS.2009-1952

Viswanatha, K. P., Patil, R., Upadhyaya, H. D., Khan, H., Gururaj, S., ., S., et al. (2020). Genetic Diversity, Association and Principle Component Analyses for Agronomical and Quality Traits in Genomic Selection Training Population of Groundnut (*Arachis hypogaea* L.). *Ijgpb* 80, 282–290. doi:10.31742/IJGPB.80.3.7

Vivek, B. S., Krishna, G. K., Vengadessan, V., Babu, R., Zaidi, P. H., Kha, L. Q., et al. (2017). Use of Genomic Estimated Breeding Values Results in Rapid Genetic Gains for Drought Tolerance in Maize. *Plant Genome* 10. doi:10.3835/PLANTGENOME2016.07.0070/FORMAT/PDF

Wang, X., Xu, Y., Hu, Z., and Xu, C. (2018). Genomic Selection Methods for Crop Improvement: Current Status and Prospects. *Crop J.* 6, 330–340. doi:10.1016/j.cj.2018.03.001

Werner, C. R., Voss-Fels, K. P., Miller, C. N., Qian, W., Hua, W., Guan, C. Y., et al. (2018). Effective Genomic Selection in a Narrow-Genepool Crop with Low-Density Markers: Asian Rapeseed as an Example. *Plant Genome* 11, 170084. doi:10.3835/plantgenome2017.09.0084

WHO/FAO (2003). *Diet, Nutrition and the Prevention of Chronic Diseases: Recommendations for Preventing Excess Weight Gains and Obesity*. Geneva, Switzerland: WHO, 1–148.

Xiong, S., Wang, M., Zou, J., Meng, J., and Liu, Y. (2020). A Two-Stage Method for Improving the Prediction Accuracy of Complex Traits by Incorporating Genotype by Environment Interactions inBrassica Napus. *Discrete Dyn. Nat. Soc.* 2020, 1–12. doi:10.1155/2020/7959508

Xu, S. (2007). An Empirical Bayes Method for Estimating Epistatic Effects of Quantitative Trait Loci. *Biometrics* 63, 513–521. doi:10.1111/J.1541-0420.2006.00711.X

Xu, Y., Wang, X., Ding, X., Zheng, X., Yang, Z., Xu, C., et al. (2018). Genomic Selection of Agronomic Traits in Hybrid rice Using an NCII Population. *Rice (N Y)* 11, 32–10. doi:10.1186/S12284-018-0223-4/FIGURES/5

Yabe, S., Yoshida, H., Kajiya-Kanegae, H., Yamasaki, M., Iwata, H., Ebana, K., et al. (2018). Description of Grain Weight Distribution Leading to Genomic Selection for Grain-Filling Characteristics in rice. *PLOS ONE* 13, e0207627. doi:10.1371/JOURNAL.PONE.0207627

Yuan, Y., Cairns, J. E., Babu, R., Gowda, M., Makumbi, D., Magorokosho, C., et al. (2019). Genome-wide Association Mapping and Genomic Prediction Analyses Reveal the Genetic Architecture of Grain Yield and Flowering Time under Drought and Heat Stress Conditions in maize. *Front. Plant Sci.* 9, 1919. doi:10.3389/FPLS.2018.01919/FULL

Zhang, H., Yin, L., Wang, M., Yuan, X., and Liu, X. (2019). Factors Affecting the Accuracy of Genomic Selection for Agricultural Economic Traits in maize, Cattle, and Pig Populations. *Front. Genet.* 10, 189. doi:10.3389/FGENE.2019.00189/BIBTEX

Zhang, X., Pérez-Rodríguez, P., Burgueño, J., Olsen, M., Buckler, E., Atlin, G., et al. (2017). Rapid Cycling Genomic Selection in a Multiparental Tropical maize Population. *G3: Genes, Genomes, Genet.* 7, 2315–2326. doi:10.1534/G3.117.043141

Zhao, Y., Gowda, M., Liu, W., Würschum, T., Maurer, H. P., Longin, F. H., et al. (2012). Accuracy of Genomic Selection in European maize Elite Breeding Populations. *Theor. Appl. Genet.* 124, 769–776. doi:10.1007/S00122-011-1745-Y

Zhao, Y., Mette, M. F., and Reif, J. C. (2015). Genomic Selection in Hybrid Breeding. *Plant Breed* 134, 1–10. doi:10.1111/PBR.12231

Zhao, Y., Zeng, J., Fernando, R., and Reif, J. C. (2013). Genomic Prediction of Hybrid Wheat Performance. *Crop Sci.* 53, 802–810. doi:10.2135/CROPSCI2012.08.0463

Ziyatdinov, A., Ve1;zquez-Santiago, M., Brunel, H., Martinez-Perez, A., Aschard, H., and Soria, J. M. (2018). lme4qtl: Linear Mixed Models With Flexible Covariance Structure For Genetic Studies Of Related Individuals. *BMC Bioinformat.* 19, 68. doi:10.1186/s12859-018-2057-x

Zou, H., and Hastie, T. (2005). Regularization and Variable Selection via the Elastic Net. *J. R. Stat. Soc B* 67, 301–320. doi:10.1111/j.1467-9868.2005.00503.x

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.

Copyright © 2022 Budhlakoti, Kushwaha, Rai, Chaturvedi, Kumar, Pradhan, Kumar, Kumar, Juliana, Mishra and Kumar. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genetics

ORIGINAL RESEARCH
published: 09 February 2022
doi: 10.3389/fgene.2021.782366

Transcriptome Analysis of Bread Wheat Genotype KRL3-4 Provides a New Insight Into Regulatory Mechanisms Associated With Sodicity (High pH) Tolerance

Geeta Prasad[1†], Shikha Mittal[1†], Arvind Kumar[2], Divya Chauhan[1], Tanmaya Kumar Sahu[1], Sundeep Kumar[1], Rakesh Singh[1], Mahesh C. Yadav[1] and Amit Kumar Singh[1]**

[1]Division of Genomic Resources, ICAR-NBPGR, New Delhi, India, [2]ICAR-Central Soil Salinity Research Institute, Karnal, India

OPEN ACCESS

Edited by:
Abhishek Bohra, Indian Institute of Pulses Research (ICAR), India

Reviewed by:
Dawei Xue, Hangzhou Normal University, China
Dharmendra Singh, Indian Agricultural Research Institute (ICAR), India

***Correspondence:**
Amit Kumar Singh
amit.singh5@icar.gov.in
Arvind Kumar
Arvind.kumar@icar.gov.in

[†]These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Plant Genomics, a section of the journal Frontiers in Genetics

Received: *24 September 2021*
Accepted: *14 December 2021*
Published: *09 February 2022*

Citation:
Prasad G, Mittal S, Kumar A, Chauhan D, Sahu TK, Kumar S, Singh R, Yadav MC and Singh AK (2022) Transcriptome Analysis of Bread Wheat Genotype KRL3-4 Provides a New Insight Into Regulatory Mechanisms Associated With Sodicity (High pH) Tolerance. Front. Genet. 12:782366. doi: 10.3389/fgene.2021.782366

Globally, sodicity is one of the major abiotic stresses limiting the wheat productivity in arid and semi-arid regions. With due consideration, an investigation of the complex gene network associated with sodicity stress tolerance is required to identify transcriptional changes in plants during abiotic stress conditions. For this purpose, we sequenced the flag leaf transcriptome of a highly tolerant bread wheat germplasm (KRL 3–4) in order to extend our knowledge and better understanding of the molecular basis of sodicity tolerance. A total of 1,980 genes were differentially expressed in the flag leaf due to sodicity stress. Among these genes, 872 DEGs were upregulated and 1,108 were downregulated. Furthermore, annotation of DEGs revealed that a total of 1,384 genes were assigned to 2,267 GO terms corresponding to 502 (biological process), 638 (cellular component), and 1,127 (molecular function). GO annotation also revealed the involvement of genes related to several transcription factors; the important ones are *expansins*, *peroxidase*, *glutathione-S-transferase*, and metal ion transporters in response to sodicity. Additionally, from 127 KEGG pathways, only 40 were confidently enriched at a p-value <0.05 covering the five main KEGG categories of metabolism, i.e., environmental information processing, genetic information processing, organismal systems, and cellular processes. Most enriched pathways were prioritized using MapMan software and revealed that lipid metabolism, nutrient uptake, and protein homeostasis were paramount. We have also found 39 SNPs that mapped to the important sodicity stress-responsive genes associated with various pathways such as ROS scavenging, serine/threonine protein kinase, calcium signaling, and metal ion transporters. In a nutshell, only 19 important candidate genes contributing to sodicity tolerance in bread wheat were identified, and these genes might be

Abbreviations: TF, transcription factor; ROS, Reactive Oxygen species; GO, gene ontology; QC, quality control; BP, biological process; MF, molecular function; CC, cellular component; DEGs, differentially expressed genes; KEGG, Kyoto Encyclopedia of genes and genomes; AS, alternative splicing; SNP, single nucleotide polymorphism; PC-PLA1, phospholipases A1; pPLA2-II, phospholipase A2; KCR, 3-ketoacyl-CoA reductase; KCS, 3-ketoacyl-CoA synthase; GPAT, Glycerol-3-phosphate acyl-transferase; DGAT, acyl-CoA:diacylglycerol acyltransferase; SPX, phosphate signaling regulatory protein; PHT1, phosphate transporter 1; NLA, Nitrogen limitation adaptation, GRX, Glutaredoxins; GUX, glucuronosyltransferase; CARE, cis-acting regulatory elements; STPK, Serine-threonine protein kinases; TAG, Triacylglycerol; BBI, Bowman-Birk protease inhibitor; AGPs, Arabinogalactan proteins; DRE, dehydration-responsive element.

helpful for better understanding and further improvement of sodicity tolerance in bread wheat.

Keywords: KRL 3-4, transcriptome, sodicity, DEGs, ROS

INTRODUCTION

Achieving global food security in the 21st century seemed arduous due to acute environmental challenges including extreme climatic vulnerability (Fujimori et al., 2019) and persistent land degradation (Subramaniam and Masron, 2021). Salts specifically excess sodium ions affected land degradation distorted nearly 1,128 million ha (Mha) land and causing US$ 27.3 billion economic losses per year (Qadir et al., 2014). Neoteric projections indicated that ~12 Mha of productive land salinized every year by the dint of natural and anthropogenic processes, and more than half of the total croplands is to be expectedly salinized by 2050 (UNCCD, 2017). Bread wheat (*Triticum aestivum* L.), is the staple food for nearly 2.5 billion of the world population, and constitutes a major share in carbohydrates (55%) and food calorie (20%) consumption in dietary intake (Ramadas et al., 2020). It concedes as a moderately tolerant crop to sodicity (exchangeable sodium percentage, ESP: 35%–50%) stress (Maas, 2019), although with the tolerance, transmogrify with plant species, stress intensity, and developmental stage (Sheoran et al., 2021). Furthermore, an extensive assessment of production loss in wheat cultivated in salt-affected soils revealed 20%–43% yield penalty, only due to adverse soil constraints (Qadir et al., 2014; Chaurasia et al., 2021).

The sodic soil (pH>8.5) is detrimental for mineral elements absorption and impedes with ionic balances in the soils. Globally, more than 30% of the surface land of the earth is covered with high pH soils (Chen and Barak, 1982). Sodic soils having a high level of exchangeable sodium on clay particles tend to develop low porosity with dense structure and high soil strength resulting in restricted root penetration, nutrient uptake, plant water availability, and metabolic activities, thereby adversely affecting crop productivity (Anzooman et al., 2019; Minhas et al., 2019). Previous studies have reported that plants exhibit poor seed germination, seedling survival, stunted shoot and root growth, and decrease in nutrient solubility under sodicity (pH >8.5), consequently, very poor grain yield or fails to reach maturity. The high pH directly affects the mineral element absorption and impedes with ionic balances. Therefore, the elevated soil pH >8.5 is much more detrimental to plants than neutral salts (Yaduvanshi et al., 2012).

Plants have inherent mechanisms to survive under different abiotic and biotic stresses. It is important to unveil the molecular mechanism involved in the regulation of abiotic stress in order to improve the tolerance level of plants. Earlier studies have shown that in adverse conditions, plants cope with the situation through generic stress signal perception and transduction that alters genes responsible for stress and protects plants in the stress condition (Shinozaki and Yamaguchi-Shinozaki, 2007). Past studies revealed that a large number of stress-responsive genes are involved in the stress signaling network. These can be divided into two groups on the basis of their functions and products. The first group include genes involved in protection against cell damage during stress, cell viability such as osmolyte biosynthetic enzymes, antioxidant proteins, and chaperones. In the second group, various transcription factors (TFs) and other parts composed of protein phosphatase and protein kinases are involved (Shinozaki and Yamaguchi-Shinozaki, 2007). Osmotic homeostasis and reactive oxygen species (ROS) scavenging have a critical role in adapting high salt conditions. In addition, iron acquisition, nitrate assimilation, and calcium metabolism could effectively promote a mechanism of tolerance to environment changes in plants (An et al., 2016). Moreover, higher organic acids, such as citric acid accumulation, facilitates tolerance to salinity in soybean (Li et al., 2017), whereas H^+ secretion in *Arabidopsis* under salt stress might be involved in the regulation of salt tolerance (Shi and Zhu, 2002). Thus, the modulation of ROS scavenging, osmotic homeostasis, and H^+ secretion jointly helps plants to survive in adverse condition and induce salt tolerance in plants compared with sensitive plants that lack such mechanism of tolerance (Fu et al., 2018).

Until now, research has been primarily focused on saline stress condition, and there are limited studies conducted for understanding of the molecular basis of sodicity stress tolerance (high pH stress) in crop plants (Meng et al., 2017). In our study, a wheat genotype KRL 3–4 that is considered to be one of the most sodicity stress tolerant line was selected to study transcriptional level changes in response to higher pH (Singh et al., 2010). RNA samples collected from wheat flag leaves of plants grown under sodic and normal soils were processed for high-throughput RNA sequencing. GO and KEGG functional enrichment analyses revealed the role of specific DEGs under sodicity stress. Furthermore, a major hub of genes that might have a role in tolerance against sodicity stress were identified. Moreover, bioinformatics analysis unveiled *cis*-acting regions of putative candidate genes that are involved in sodicity stress tolerance.

MATERIALS AND METHODS

Background of Wheat Genotype KRL3-4

The genotype was developed at ICAR-Central Soil Salinity Research Institute, Karnal, India, through conventional breeding method (bulk selection) involving the variety HD 1982 (released in the year 1975 for north eastern plain zone, pedigree: E5557/HD H45) as a female and Kharchia 65 (released in the year 1970 for salt affected lands in all the wheat growing zones, pedigree Kharchia local/EG 953) as a male parent. The selected genotype is tall (140–155 cm) in nature and exceptionally unique in salt tolerance, having pale green foliage and higher early seedling vigor similar to their parent Kharchia 65. It possesses high yielding ability along with higher levels of tolerance to salinity, sodicity, and water-logging compared with their

parent Kharchia 65. This conclusion was substantiated by 4 years of observations of the All India Salinity/Alkalinity Tolerance Screening Nursery (SATSN) conducted in different salt-affected agro-ecosystems during 2004–2005 to 2008–2009 (**Supplementary Table S1**). In the previous literature, Dr. Abdul Mujeeb-Kazi (distinguished scientist and ex program leader of CIMMYT) had already mentioned that Kharchia 65 is the most salt-tolerant bread wheat in the world (Wang et al., 2003).

This germplasm already has been registered (IC408331; INGR09087), in ICAR-National Bureau of Plant Genetic Resources (NBPGR), New Delhi, salt and waterlogging tolerance. The salt tolerance of KRL 3–4 is characterized by better regulation over lower uptake of sodium and highest uptake of magnesium salt-affected soils (Singh et al., 2010). Additionally, it possesses higher tolerance to boron, iron, and aluminum concentrations in the soil solution, and their full potential reflected on highly sodic condition (pH >9.3), where it performs as a halophyte, and it competes with the parent Kharchia 65 in yield as well as in other physiological traits. KRL 3–4 also had the highest grain Ca, Mg, Fe, Zn, and S concentrations compared with most of the Indian cultivars, as it substantiates in an experiment conducted in hostile soil conditions (pH: 4.5–9.5) across the six different agro-ecosystems for 2 years (Khokhar et al., 2018).

Plant Material and Experimental Design

In our study, a wheat genotype KRL 3–4 was selected to understand the transcriptional changes in response to sodicity stress. The genotype was grown in two concrete lysimeters (dimension: 2 m × 2 m × 1 m) situated at ICAR-Central Soil Salinity Research Institute, experimental farm (29°42′31″N, 76°57′6″E) Karnal, India. The first lysimeter, filled homogeneously with sandy loam soil (Inceptisols, pH: 7.4, ESP:6.3, ECe: 0.9 dS m^{-1}, organic carbon: 2.3 g kg^{-1} soil) was considered as control. The other lysimeter was filled homogeneously with incubated synthetic sodic soils formulated by mixing the carbonate and bicarbonate salts in the soils (pH: 9.5, ESP: 74.6, ECe: 1.6 dS m^{-1}, organic carbon: 2.1 g kg^{-1} soil). The sowing seeds were surface sterilized in 1% sodium hypochlorite ($NaClO^{-}$) for 10 min and then washed with distilled water several times and planted in the lysimeters. Plants that were grown in normal soil were considered control plants, and other plants grown in the sodic lysimeters (stress intensity: pH-9.5) were considered treated plants. At the anthesis stage, fully matured flag leaf samples were taken from both control and treated plants with two biological replicates. The leaf samples were immediately frozen in liquid nitrogen and stored at −80°C until analysis.

RNA Extraction and Illumina Deep Sequencing

Total RNA was isolated from the two biological replicates of flag leaves from normal and treated plants (after 12-h exposure to sodicity stress) using RNeasy Plant Mini Kit (Qiagen) according to the instructions of the manufacturer. Equal amounts of total RNA of each two biological replicates were used for the RNA sequencing. Initially, the concentration of RNA was checked by nanodrop and later purity through agarose gel electrophoresis. Furthermore, RNA integrity was estimated by Agilent Bioanalyzer 2100 system (Agilent Technologies Co. Ltd., Beijing, China). Accordingly, the samples with RIN value higher than 9 were used for transcriptome sequencing. After quality control (QC) of the RNA samples, poly (A) enrichment, RNA fragmentation, random hexamer-primed cDNA synthesis, linker ligation, size selection, and PCR amplification reactions were performed to prepare cDNA libraries for each sample. Finally, the qualified libraries fed into HiSeq sequencer according to its effective concentration and expected data volume. Illumina HiSeq 2500 platform was used to sequence libraries, and 150-bp paired end reads were generated.

Quality Control, Alignment, and Differential Expression Gene Analysis

Raw reads of control and treated samples were evaluated for their quality using FastQC (version 0.11.8) (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/). The adaptor sequences and low-quality sequence reads were removed from the data sets using trim galore (version 0.6.2) considering the threshold parameters, including phred score 33, quality score 20, and length 20 (https://github.com/FelixKrueger/TrimGalore/releases). Raw sequences were transformed into clean reads after data processing. The quality of cleaned reads was also checked through base quality score distribution, sequence quality score distribution, average base content per read, and GC distribution in the reads. The obtained high-quality reads were mapped to the reference genome of *Triticum aestivum*, i.e., *IWGSC* v.1.0 (ftp://ftp.ensemblgenomes.org/pub/plants/release-42) using bwa (*version 0.7.5*, http://bio-bwa.sourceforge.net/) with default parameters. These mapped reads were spliced using Cufflinks software based on the reference genome sequence. Quantification of the gene expression levels was estimated as fragments per kilo base of transcript per million fragments mapped (FKPM). Differential expression analysis between control and treated samples was performed using the DESeq R package (1.10.1) (Anders and Huber, 2010). DESeq provides reliability for determining differential expression in digital gene expression data using a model based on the negative binomial distribution. The resulting *p*-values were adjusted using Benjamin and Hochberg's approach for controlling the false discovery rate. Genes with an adjusted *p*-value <0.05 were found using DESeq and were assigned as differentially expressed. A fold change ≥2 and FDR <0.01 were considered as the thresholds for determining the differential expression of a gene.

Gene Ontology and Gene Pathway Enrichment Analysis

Gene ontology (GO) enrichment analysis of the DEGs was implemented using topGO R package (https://bioconductor.org/packages/release/bioc/html/topGO.html). The GO terms, i.e., biological process (BP), cellular component (CC), or molecular function (MF) were assigned to the DEGs. Kyoto

Encyclopedia of genes and Genomes (KEGG) is a database resource for understanding high-level functions and utilities of the biological systems, such as the cell, organism and ecosystem from molecular-level information, especially large-scale molecular datasets generated by genome sequencing and other high-throughput experimental technologies (http://www.genome.jp/kegg/). KOBAS software (http://kobas.cbi.pku.edu.cn/) was used to analyze the statistical enrichment of differential expression of genes in KEGG pathways. Furthermore, MapMan (version 3.5.1; http://mapman.gabipd.org/web/guest) was used for pathway analysis of DEGs as p-value <0.05.

Alternative Splicing Analysis

Alternative splicing (AS) events for identified DEGs under sodicity stress were determined using an AStalavista web tool (version 3; http://genome.crg.es/astalavista/) with default parameters. The results obtained for all AS events were further analyzed to unravel the molecular mechanisms of AS during exposure of wheat crop to the sodicity stress.

Identification of Single Nucleotide Polymorphism

The freebayes tool (Garrison and Marth, 2012) and SNP effect predictor (Cingolani et al., 2012) were used to identify and annotate the SNPs, respectively. SNPs and INDEL variants were filtered through bcftools. Based on genomic locations, the variants were annotated as intronic, exonic, 5′UTR, and 3′UTR.

Candidate Gene Identification

The candidate genes for sodicity stress tolerance were identified on the basis of similarity (blastx with similarity >80% and e-value <0.001) with published or validated genes for their role in sodicity stress tolerance (Singh et al., 2018). The heatmap was made using ggplot2 R package (https://cran.r-project.org/web/packages/ggplot2/index.html). Furthermore, the chromosomal localization of the identified candidate genes was performed using MapChart (Voorrips, 2002), and the 2-kb upstream region was used for the identification of *cis*-acting regulatory elements using PlantCARE database (Rombauts et al., 1999).

RESULTS

Sequencing Statistics

In order to understand the mechanism involved at the transcriptional level for sodicity stress tolerance, transcriptome analysis was carried out for control and treated samples of wheat flag leaves. The libraries were sequenced to obtain 150-bp paired-end reads using the Illumina high-throughput sequencing platform. In total, approximately 37.24 Gb of raw data was generated from the two biological replicates of the wheat flag leaf. After trimming, the volume of the total clean reads for all the samples was 36.54 Gb. The clean reads of each sample was more than 8.75 Gb, and the percentage of Q20 bases was ≥97.40%. Sequencing quality distribution was examined over the complete sequence length to detect the sites (base positions) with an unusually low sequencing quality, where incorrect bases may be incorporated at abnormally high levels. Approximately, 97.60%–99.06% of reads are mapped to the wheat reference genome (*IWGSC* RefSeq v1.0, http://www.wheatgenome.org/), whereas GC content varied from 51% to 53% for the control and sodicity stress treated samples (**Table 1**).

Identification of Differentially Expressed Genes and Gene Ontology Enrichment

In total, we identified 97,514 differentially expressed genes (DEGs) between control and treated flag leaf wheat samples. The genes with expression level greater than two fold (adjusted p-value < 0.05) were considered as significant DEGs under sodicity stress condition with reference to the controlled conditions. At this threshold, there were only 1,980 significant DEGs between control and sodicity stress conditions (**Supplementary Table S2**). Among these genes, 872 DEGs showed upregulation, and 1,108 showed downregulation under sodicity stress condition (**Figure 1A**).

Furthermore, GO enrichment analysis was carried out to characterize the main biological functions of DEGs in wheat flag leaves under sodicity stress condition. All the DEGs were annotated into three classes, i.e., BP, MF, and CC. Annotation of DEGs revealed that a total of 1,384 genes were assigned 2,267 GO terms corresponding to 502 BP, 638 CC, and 1,127 MF (**Supplementary Table S3**). The dominant terms in CC category were "membrane," "membrane part," "cell," "cell part," and "organelle," whereas in MF, "catalytic activity," "binding," "transporter activity," "molecular function regulator," and "molecular transducer activity" were the most dominant GO terms. However, in case of BP, most of the DEGs were classified in "response to stimulus," "cellular process," "metabolic process" followed by "biogenesis," and "localization" (**Figure 1B**).

As expected, the most enriched BP terms for over-presented DEGs were associated with metal ion transport, response to auxin, water and oxidative stress, transmembrane transport, oligopeptide transport, lipid transport, carbohydrate metabolic process, plant-type cell wall organization, which acted as indicators of significant biological processes underlying the specific sodicity stress responses of plants. The most enriched MF terms for over-presented DEGs were metal ion binding, ATP binding, calcium ion binding, catalytic activity, channel activity, iron and magnesium ion binding, oxidoreductase activity, transferase activity, voltage-gated anion channel activity, etc. Meanwhile, with regard to the over-representation of CC terms among the DEGs, we found extracellular region, membrane, intrinsic to membrane, integral to the membrane, and intrinsic to plasma membrane to be the most enriched.

Furthermore, a large number of TF genes were found to be differentially expressed under sodicity stress condition. In total, 985 DEGs were encoded for 53 TF families

TABLE 1 | Summary of sequencing and mapping statistics of the bread wheat germplasm (KRL3-4) genotype used in the present study.

	Replicate	Raw data (Gb)	Clean data (Gb)	Mapping %	GC %
Control	Replicate 1	9.28	9.11	98.43	53
	Replicate 2	9.02	8.87	97.61	52
Sodicity stress	Replicate 1	9.98	9.77	99.06	52
	Replicate 2	8.95	8.78	98.99	51

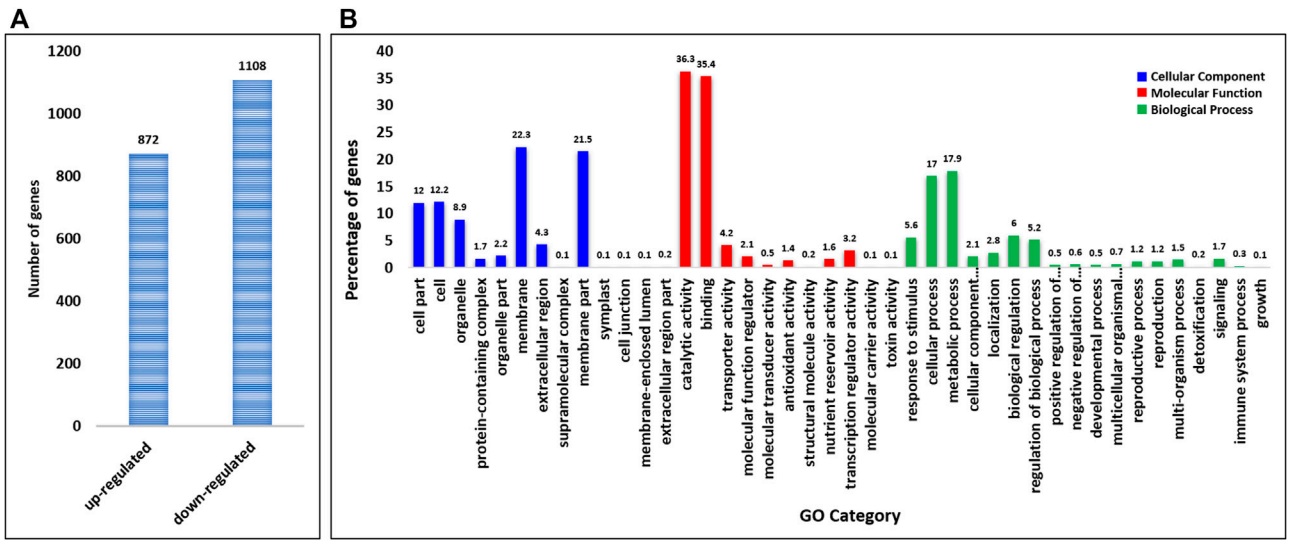


FIGURE 1 | (A) Significant up and downregulated differentially expressed genes (DEGs) identified under sodicity stress. **(B)** Gene Ontology (GO) annotation of DEGs in sodicity treated wheat flag leaves summarized in three main categories: biological process (BP), molecular function (MF), and cellular component (CC).

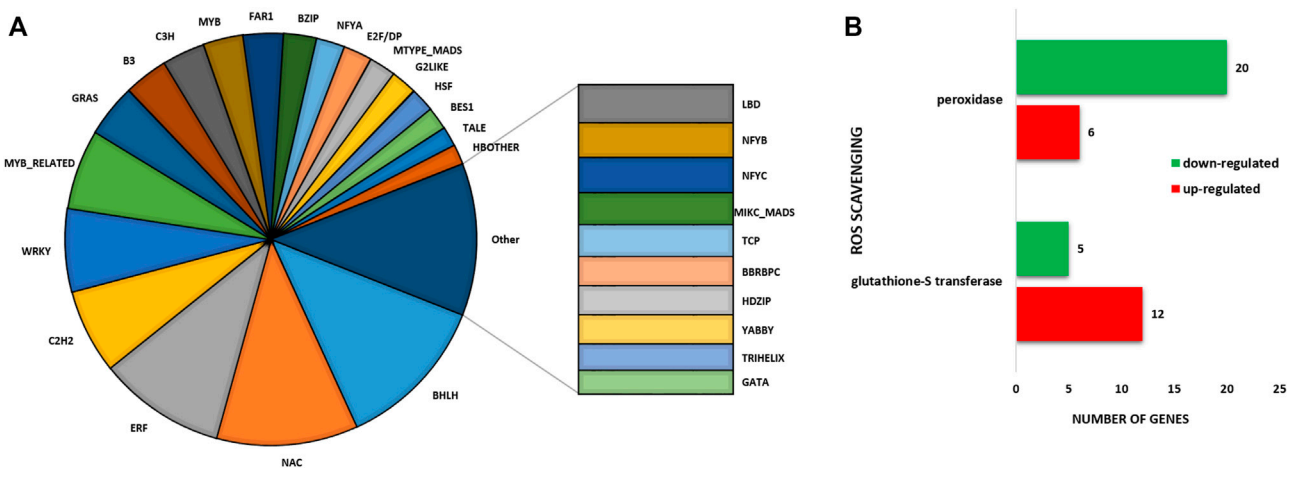


FIGURE 2 | (A) Pie chart depicting the involvement of different transcription factors (TFs) in sodicity treated wheat flag leaf samples. **(B)** Bar diagram representing number of the identified reactive oxygen species (ROS) scavenging genes.

(**Figure 2A**). The important TF families that were represented by DEGs included *bHLH*, *ERF*, *bZIP*, *NAC*, *MYB-RELATED*, *MYB*, and *WRKY*. Among the TF families, the highest number of genes belonged to the *bHLH* family (109) followed by *NAC* (99), *ERF* (89), *C2H2* (59), *WRKY* (58), *MYB_RELATED* (56), and *GRAS* (37), and so on. We also checked the expression of

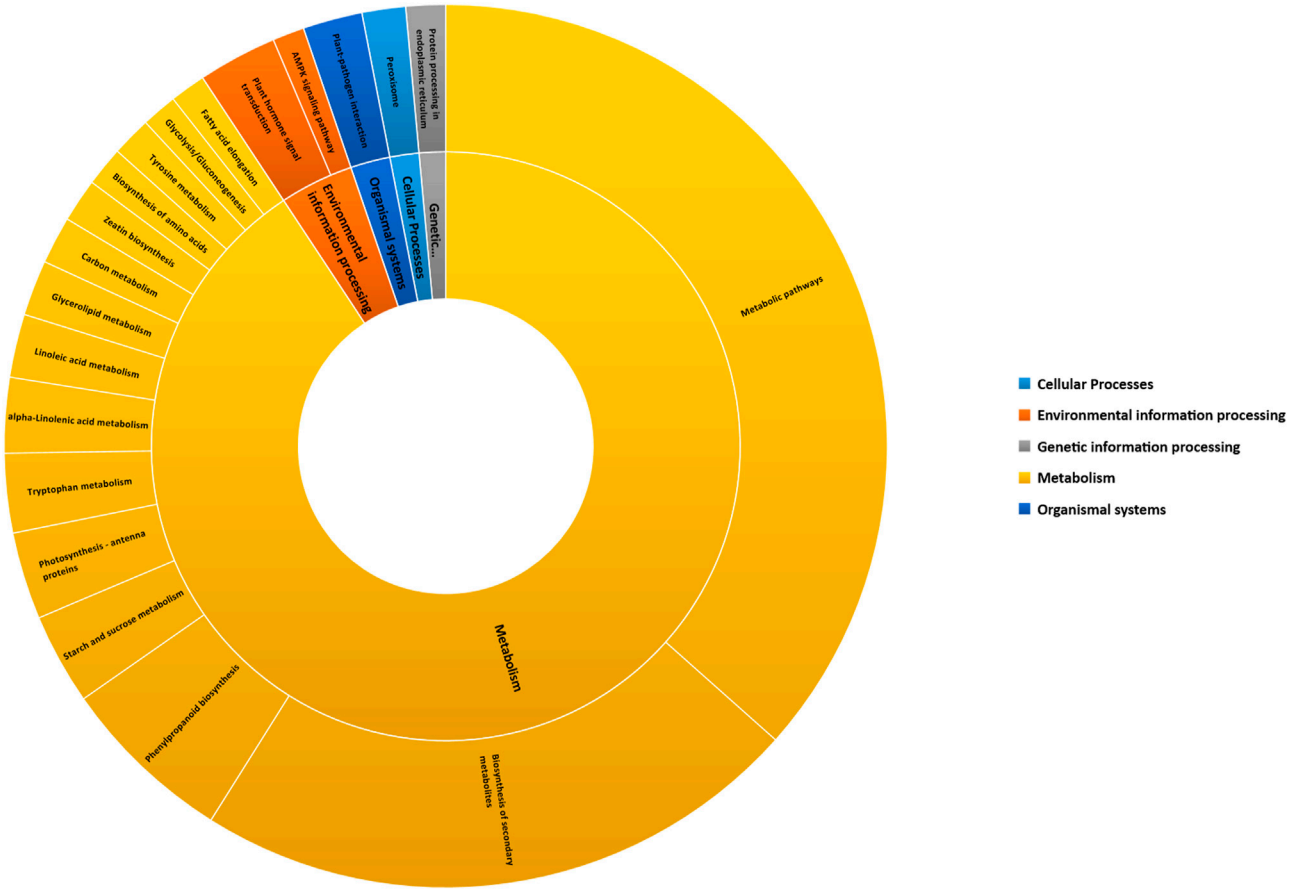


FIGURE 3 | Enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways categorized in top five KEGG pathway categories, i.e., cellular processes, environmental information processing, genetic information processing, metabolism, and organismal systems.

these TF in control as well as in sodicity treated samples and visualized using heat map. Moreover, we have also identified 37 DEGs encoding for serine/threonine kinase family proteins, out of which 22 showed upregulation, and the remaining 15 showed downregulation under sodicity stress. Similarly, genes encoding for other stress tolerance and signaling-associated proteins were also identified, including 22 genes of calmodulin ion binding and calcium signaling, 11 genes for metal ion transport, 4 genes of ABC transporter, 4 genes of SWEET sugar transporter, and 1 gene encoding potassium transporter. In this study, 43 DEGs appear to protect plant cells from oxidative damage by *ROS* scavenging, such as peroxidase genes (26) and *glutathione S-transferase* genes (17). For peroxidase genes, 6 of the 26 DEGs were upregulated, and the other 20 genes were observed to be downregulated. Besides, the peroxidase gene family, a number of *glutathione S-transferase* genes were found differentially expressed. Most of these genes (12/17) were noticed to be upregulated during sodicity stress (**Figure 2B**; **Supplementary Table S4**).

Kyoto Encyclopedia of Genes and Genomes and MapMan Analysis of Differentially Expressed Genes Under Sodicity Stress

The KEGG pathway analysis was performed to unveil the active biological pathways involved in DEGs under sodicity stress in KRL3-4 flag leaf. KEGG is a database resource for interpretation of large-scale molecular datasets generated by genome sequencing for understanding high-level functions and utilities of the biological system (Kanehisa et al., 2017). The pathway analysis showed that out of 127 KEGG pathways, 40 pathways were observed to be significantly enriched at a p-value <0.05 covering the five main KEGG categories of metabolism, environmental information processing, genetic information

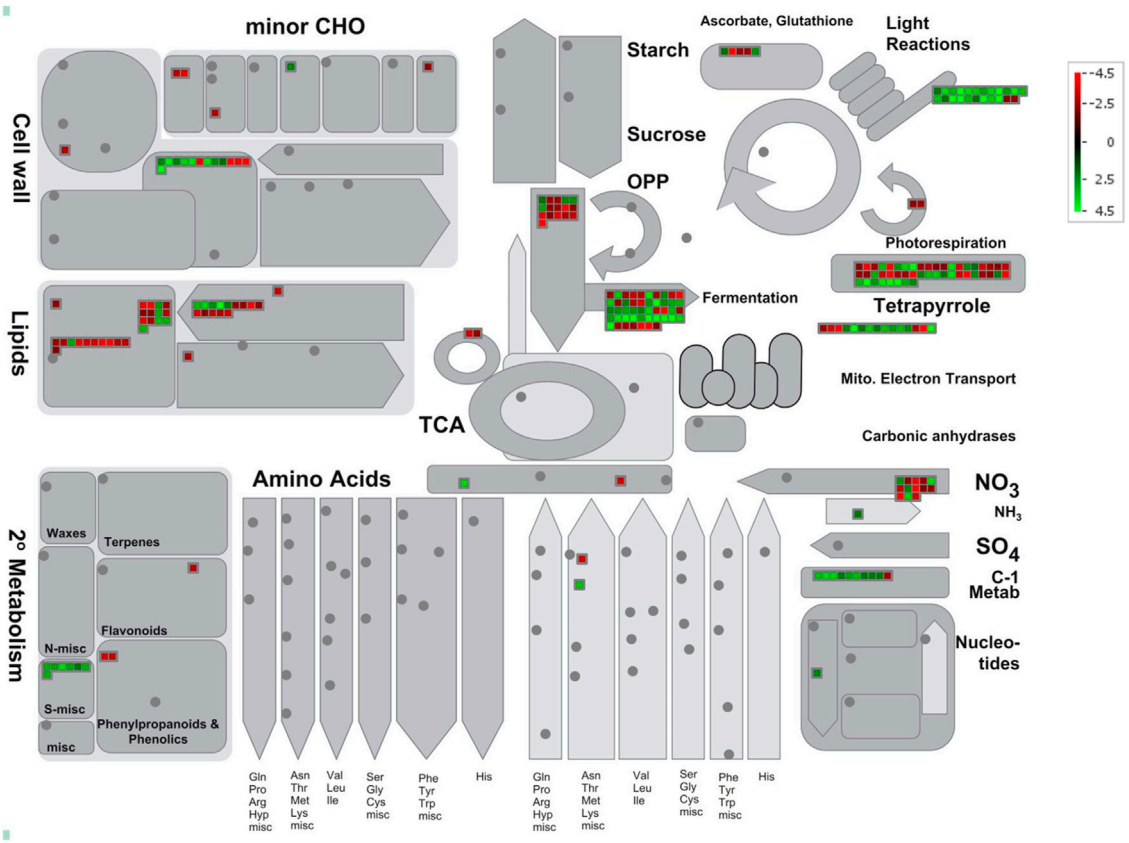


FIGURE 4 | MapMan representing the differential expression of genes involved in metabolism pathway under sodicity stress condition.

processing, organismal systems, and cellular processes (**Figure 3**). The KEGG enrichment pathways indicated that the specific DEGs, found under sodicity stress in flag leaf, were widely enriched in the pathways of metabolic, biosynthesis of secondary metabolites, phenylpropanoid biosynthesis, starch and sucrose metabolism, photosynthesis, and plant hormone signal transduction. According to previous studies, these pathways were found related to abiotic stress tolerance. The list of DEGs involved in 40 KEGG enriched pathway is given in **Supplementary Table S5**.

The differentially expressed sodicity stress-responsive genes were also visualized using MapMan in order to find out alterations in metabolic pathways during sodicity stress (**Supplementary Table S6**). The MapMan analysis revealed that lipid metabolism and protein homeostasis, nutrient uptake, etc., were the most enriched pathways under sodicity stress. In lipid metabolism, particularly, the expression of 12 genes involved in lipid degradation, i.e., 9 *phospholipases A1* (*PC-PLA1*) and 3 *phospholipase A2* (*pPLA2-II*) were highly upregulated in response to sodicity stress, whereas the genes related to fatty acid biosynthesis and elongation (11 genes), i.e., *3-ketoacyl-CoA reductase* (*KCR*) and *3-ketoacyl-CoA synthase* (*KCS*) showed mixed expression. *Glycerol-3-phosphate acyltransferase* (*GPAT1-3*) and *acyl-CoA:diacylglycerol acyltransferase* (*DGAT2*), involved in glycerolipid biosynthesis, were found to be up and downregulated, respectively. Similarly, in the case of protein homeostasis, the genes encoding Bowman–Birk protease inhibitor, pepsin-type protease, amino acid hydrolase was found to be upregulated under sodicity stress. On the other hand, we have also observed the upregulation of *phosphate signaling regulatory protein* (*SPX*), *phosphate transporter 1* (*PHT1*), and *nitrogen limitation adaptation* (*NLA*) genes in nutrient uptake pathway. We identified one gene for *PHT1* (*Traescs2a02g047300*), one for *NLA* (*Traescs5b02g373000*) and seven for *SPX* (*Traescs2a02g169600*, *Traescs2b02g195900*, *Traescs2d02g177100*, *Traescs7a02g376200*, *Traescs7a02g554100*, *Traescs7b02g478000*, and *Traescs7d02g372600*) **(Figure 4)**. The cellular overview pathway exposed that the genes encoding for cell wall organization showed mixed expression under sodicity stress. For example, the genes encoding *glutaredoxins* (GRXs) and *arabinogalactan* protein (xylogen) were found highly upregulated while alpha class *expansins* and *glucuronosyltransferase* (GUX) showed downregulation. GRXs are small redox proteins, which use glutathione to catalyze the reduction of disulfide bonds of substrate proteins to maintain cellular redox homeostasis. Furthermore, diverse functions such as transcriptional regulation of defense responses, flower development, oxidative stress response,

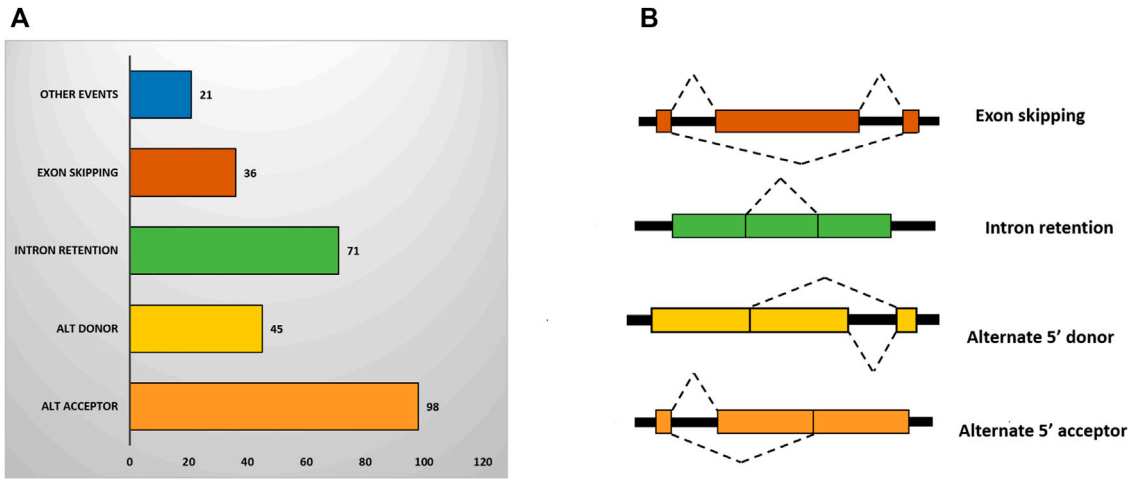


FIGURE 5 | (A) Bar chart showing different types of alternative splicing events in DEGs. **(B)** Structure of different types of alternative splicing (AS) events where black box and colored box represent introns and exon, respectively.

redox signaling, hormonal regulation, iron homeostasis, and environmental adaptation have been reported for various plant GRXs.

Alternative Splicing Analysis

Alternative splicing (AS) is an important post-transcriptional mechanism that increases the complexity and diversity of transcriptome and proteome in higher eukaryotes. In plants, AS is found to be involved in a range of functions including plant growth and development, and responses to biotic and abiotic stresses. We have categorized the DEGs into four primary classes of AS events that are retained intron (RI), skipping exon (SE), alternative 5′ splice site (A5SS), and alternative 3′ splice site (A3SS) (**Figure 5B**). Our study revealed, 271 AS events in the DEGs under sodicity stress. In total, 237 (11.96%) genes out of 1,980 have shown alternative splicing of which the highest number belonged to alternate 3′ acceptor (36.16%) followed by intron retention (26.19%), alternate 5′ donor (16.60%) and exon skipping (13.28%) (**Figure 5A**). Functional categorization of these alternatively spliced DEGs revealed their involvement in different stress-responsive pathways, including the plant hormone signal transduction, calcium signaling pathway, biosynthesis of secondary metabolites, glutathione metabolism, AMPK signaling pathway, etc. In terms of GO biological processes, these were found to be involved in metal ion transport, response to water, transmembrane transport, lipid and carbohydrate metabolic process, cell wall organization, and abscisic acid-activated signaling pathway, while, the major molecular functions represented by these alternatively spliced DEGs involve abscisic acid binding, metal ion binding, iron ion binding, phospholipase A1 activity, transmembrane transporter activity, etc.

Polymorphism in the Expressed Sequences

Trait variations are primarily due to sequence level changes in the coding region. The wheat genotype KRL3-4 included in our study represents one of the most sodicity tolerant genotypes of the world. Therefore, one might expect that alignment of the flag leaf transcripts from the control and susceptible conditions against a reference genome of the *cv.* Chinese spring (a relatively susceptible variety) may unravel novel SNPs and INDELs that may have a role in sodicity stress tolerance. A total of 36.64 Gb clean reads was obtained after RNA sequencing of both control and sodicity treated wheat flag leaf samples to discover SNP and INDELs. On an average, one variant was called within every 17 kb in the treating sample and 32 kb in the control sample. In total, 2,35,985 and 1,63,420 polymorphisms (SNPs and INDELs) were identified in the control and treated samples, respectively. In the case of control samples, 39,295,381, and 39,200 SNPs were found as missense, nonsense, and silent mutations, respectively, whereas in sodicity treated wheat flag leaf samples, 26,575, 234, and 26,377 were observed as missense, nonsense, and silent mutations, respectively.

SNPs and INDELs were mapped to exonic regions that included 8.02% of the total reads (63,239 read counts) in treated and 12.95% (94,262 read counts) in control samples. Moreover, 4.10% of the total reads (32,339) were mapped on 5′UTR and 3′UTR regions in treated samples (**Figure 6**). In addition, polymorphism due to transitions (Ts) and transversions (Tv) was also determined in each sample. In control, the Ts and Tv were estimated at 7,17,799 and 3,68,641, while in the sodicity treated sample, 4,94,965 and 2,54,817 transition and transversion were detected, respectively. In addition, amino acid substitutions were also identified for SNPs mapped to *IWGSC* wheat reference genome in control and sodicity treated samples. In control sample, the major

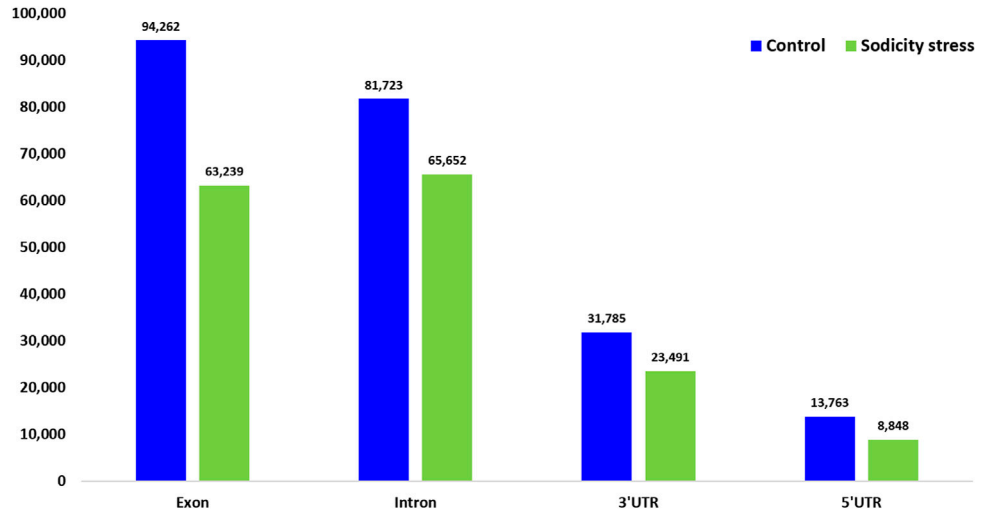


FIGURE 6 | Number of single-nucleotide polymorphism (SNPs) detected in different genomic regions in control and sodicity treated wheat flag leaf using freebayes program.

TABLE 2 | List of identified 18 candidate genes for sodicity stress using BLAST search.

Gene ID	Chr	Gene start (bp)	Gene end (bp)	Description	Expression (control)	Expression (sodicity)
TraesCS2D02G061500	2D	25,338,876	25,340,723	NAC domain-containing protein	1.45	14.24
TraesCS2D02G290400	2D	372,695,288	372,696,521	Salt-responsive protein	1.14	12.12
TraesCS3B02G255600	3B	412,595,778	412,597,321	Heat shock factor C1e	1.94	18.68
TraesCS3D02G276600	3D	383,545,476	383,548.591	Auxin-responsive protein	166.37	766.49
TraesCS4D02G102000	4D	80,220,933	80,221,864	HMA domain-containing protein	3.39	38.49
TraesCS4D02G118500	4D	99,161,360	99,162,709	Serine/threonine protein kinase	2.92	69.49
TraesCS5A02G503900	5A	669,584,344	669,585,225	Cold-responsive LEA/RAB-related COR	225.31	990.82
TraesCS5B02G516900	5B	680,671,587	680,672,746	LEA_2 domain-containing protein	4.83	52.68
TraesCS5B02G054800	5B	59,926,012	59,928,225	bHLH domain-containing protein	53.63	228.76
TraesCS5B02G260200	5B	442,821,771	442,822,446	SRC1-clade calcium sensor	1.94	26.13
TraesCS5D02G026300	5D	23,863,473	23,864,186	C2H2-ZF	6.29	246.94
TraesCS5D02G145800	5D	232,453,697	232,455,753	WRKY domain-containing protein	3.87	64.41
TraesCS5D02G245300	5D	353,795,287	353,796,446	AP2/ERF domain-containing protein	1.94	28.48
TraesCS6A02G047600	6A	24,048,539	24,050,177	Peroxidase	1.45	29.12
TraesCS6A02G098500	6A	65,681,325	65,684,847	Chloride channel protein	48.02	308.12
TraesCS6D02G144500	6D	115,043,730	115,046,100	Potassium/sodium cation transporter	40.54	213.00
TraesCS7A02G215100	7A	180,647,622	180,649,835	Fe2OG dioxygenase domain containing	0.49	16.12
TraesCS7A02G516900	7A	701,461,957	701,462,696	Calcium-dependent signaling calcium sensor	0.97	24.82

amino acid substitutions observed were alanine to valine (1,584), alanine to threonine (1,386), and valine to isoleucine (1,004), whereas in treating samples, alanine to valine (1,031), alanine to threonine (902), and valine to isoleucine (709) were noticed to be the major amino acid substitutions.

Candidate Gene Identification for Sodicity Stress

We have identified 19 sodicity stress responsive genes in our dataset. These candidate genes were identified on the basis of similarity with published validated salt-responsive genes (**Table 2**). The identified candidate genes encoded for a diverse group of proteins including *LEA*, salt-responsive protein, potassium/sodium cation transporter (HKT), calcium-dependent signaling calcium sensor, *glutathione-S-transferase*, serine/threonine protein kinase, etc. These candidate genes were found to be upregulated under sodicity stress. In terms of GO annotation, these genes were observed to be involved in metal ion binding, hydrolase activity, protein dimerization activity, diacylglycerol O-acyltrasferase activity, DNA binding, polygalacturonase activity, etc. Some candidate genes corresponded to different TF families including *C2H2* (*TraesCS5D02G026300*), *NAC* (*TraesCS2D02G061500*), *WRKY* (*TraesCS5D02G145800*), *AP2/ERF* (*TraesCS5D02G245300*),

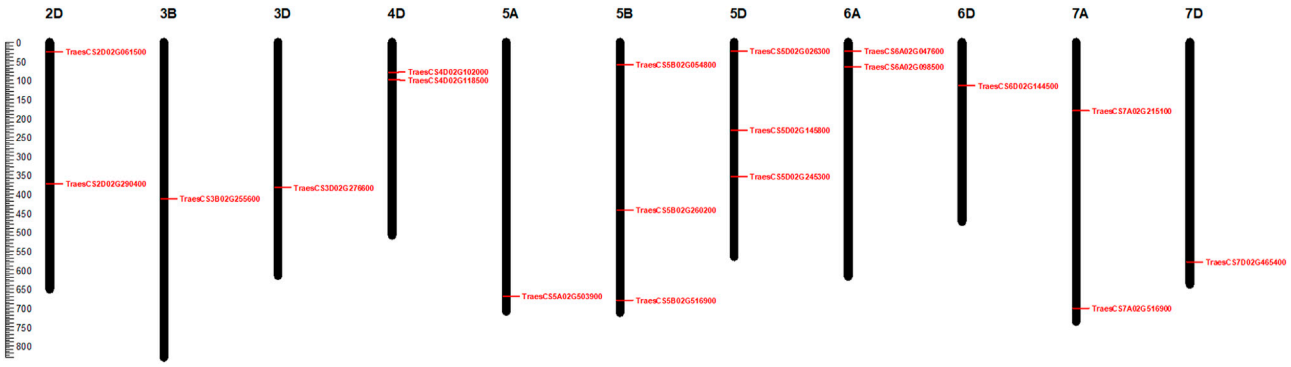


FIGURE 7 | Chromosomal localization of 19 candidate sodicity responsive genes using MapChart.

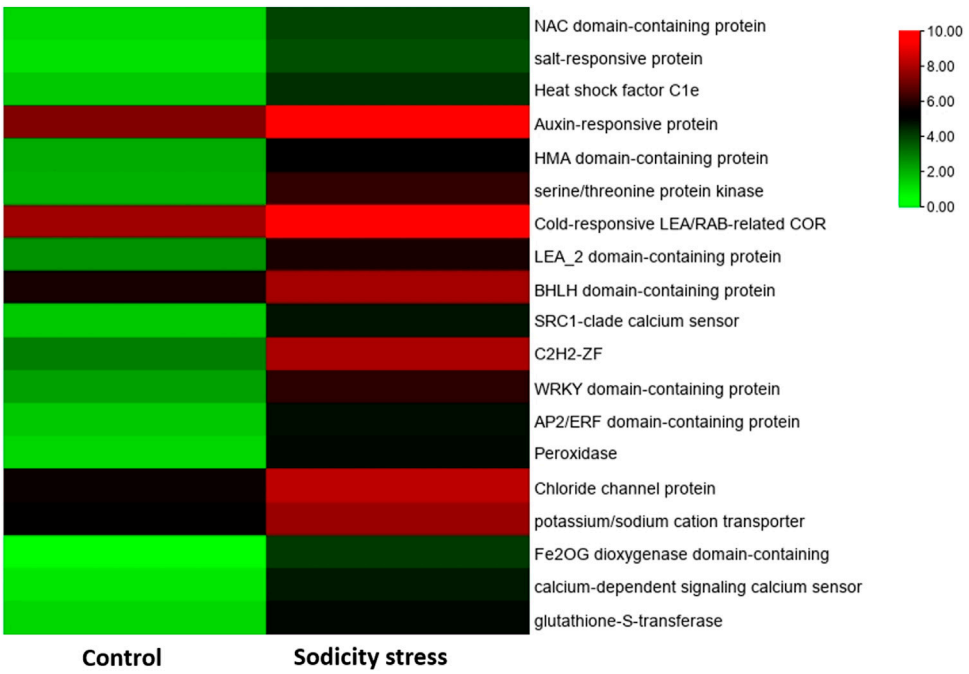


FIGURE 8 | Heat map showing differential expression of candidate sodicity-responsive genes.

ARF (*TraesCS3D02G276600*), and *bHLH* (*TraesCS5B02G054800*) and showed upregulation in case of sodicity stress compared with control.

Chromosomal distribution analysis showed that 19 identified candidate genes were unevenly mapped in 10 out of 21 wheat chromosomes. On chromosome 5, a total of seven candidate genes were located that represent 36.84% of the entire candidate genes identified. Besides, four genes were located on chromosome 3, whereas on 3B, 3D, and 6D chromosomes, one gene each was located. On the other hand, 2D, 4D, 6A and 7A chromosomes were found to contain two genes each. In contrast, no genes were found in the remaining 11 chromosomes. The sodicity responsive candidate genes were observed to be unevenly distributed among the sub-genomes A, B, and D, with 1, 4, and 10 members representing 5.26%, 21.05%, and 52.63% of the total candidate genes, respectively (**Figure 7**). The chromosomal distribution of the 19 candidate genes identified in our study is similar to a previous study of bread wheat where the maximum candidate salt-responsive genes were found on D sub-genome than on A and B sub-genomes (Singh et al., 2018) suggesting that sodicity stress controlling genes is preferentially located on the D sub-genome of wheat. The difference in expression patterns of the 19 candidate genes between control and sodicity stress was represented using heatmap (**Figure 8**). Heat map analysis indicated that nine genes showed contrasting expression between control and sodicity stress, while the rest of the

TABLE 3 | List of *cis*-acting elements found in candidate sodicity-responsive genes retrieved from PlantCARE with its consensus sequence and transcription factors and their functional description.

S. No	Motif	Consensus/Core-sequence	Transcription factor	Function	References
1	A-Box	CCGTCC	bZIPs	Conserved sequence regulates õ-amylase activity	Sheshadri et al. (2016); Liu et al. (2018); Ijaz et al. (2020)
2	ABRE3a	TACGTG	NAC	*Cis*-acting element involved in abiotic stress and signaling pathway	Zhang et al. (2020)
3	CAAT-box	CCAAT/CAAT	Growth regulating factors (GRFs)	*Cis*-element regulate plant growth and abiotic stress response	Huang et al. (2021)
4	CCGTCC-motif	CCGTCC	ERFs	Involve in ABA response and has a role in meristem specific regulation	Li et al. (2020)
5	CCGTCC-box	CCGTCC	Iron-responsive element (IRE)	Is related to meristem specific activation	Wang et al. (2016)
6	CGTCA-motif	CGTCA	MeJA	*Cis*-regulatory element involved in the MeJa response	Xing et al. (2020)
7	DRE-Core	GCCGAC	DREB/ERFs	It mediates tolerance against biotic and abiotic stress and has a role in signaling	Zhou et al. (2010)
8	G-Box	CACGTG	bZIPs, bHLH and NAC	Ubiquitous *cis*-acting element responses to various stress, i.e., drought, high salt, cold/freezing	Qian et al. (2021); Shah et al. (2021)
9	MYB	CAACAG	MYB-TF family	Plays a role in developmental processes and stress response	Chen et al. (2018); Hao et al. (2021)
10	STRE	AGGGG	ZnF	Regulates pathways initiated by abiotic and biotic stresses	Estruch (2000)
11	Sp1	GGGCGG	HD-ZIP	Putative role in drought and salinity stress	Pandey et al. (2016)
12	TGACG-motif	TGACG	MeJA	JA-responsive element	Liao et al. (2017)
13	as-1	TGACG	bZIPs, MeJA	Involvement in stress and salicylic acid-responsive element	Banerjee et al. (2015)

genes were having a mixed expression level. Therefore, this analysis suggests that these genes might have a more certain functional role in sodicity stress tolerance.

Analysis of *cis*-Regulatory Elements of Identified Candidate Genes Under Sodicity Stress

The *cis*-acting regulatory elements (CARE) have been shown to play a major role in the regulation of gene expression at the transcriptional level in plants. Indeed, these elements are the central component in plant abiotic stress responses since transcription regulation involves association between TFs and, in particular, *cis*-acting regulatory elements (CAREs) of a specific gene (Kaur and Pati, 2016). From the 19 candidate genes identified in this study, the majority belonged to TF categories, which have *cis*-binding motifs that binds to *cis*-elements of the downstream stress-responsive genes and modulate their expression. Identified 19 putative candidate genes were analyzed for their *cis*-regulatory elements using PlantCARE with default parameters (Lescot et al., 2002; Doi et al., 2008). A total of 17 CARE were A-box, ABRE3a, CAAT-box, CCGTCC-motif, CCGTCC-box, CGTCA-motif, DRE-core, G-Box, MTB, STRE, Sp1, TGACG-motif, as-1, AAGAA-motif, MBS, MYC, and TATA-box were observed among identified sodicity responsive candidate genes that have a significant role in stress tolerance. The cis-regulatory elements located in promoter regions of putative 19 sodicity responsive genes with their functional description are given in **Table 3**.

DISCUSSION

Sodicity stress is one of the serious abiotic stresses that causes severe threat to global food security, leading to desertification of land affecting crop yield and quality (Singh et al., 2015; Li et al., 2018). Most plants, including wheat and other cereal species can barely persist in sodicity soils due to high-pH toxicity, low osmotic potential of soil, and imbalanced organic composition which lead to inhibition of plant growth resulting into reduced crop productivity (Msimbira and Smith, 2020). Therefore, it is important to understand detailed molecular mechanisms underlying sodicity stress tolerance in wheat. The comprehensive understanding of cellular pathways and genes involved in salinity stress response can contribute to the development of high yielding sodicity stress tolerant wheat cultivars.

Transcriptomic approach has emerged as a very effective tool to unravel key pathways/genes responsible for abiotic stress tolerance in plants (Elamin et al., 2019). So far, there is only one report of transcriptomic analysis of wheat in response to sodicity stress. Hence, we conducted a high-throughput RNA sequencing of wheat flag leaves RNA samples under sodicity stress conditions. The statistical analysis revealed 1,891 DEGs in response to sodicity stress. The GO analysis revealed that the DEGs were related mainly to the TFs, serine-threonine kinase protein, metal ion transporters and ROS scavenging under sodicity stress. Similarly, a previous study of wheat revealed higher ROS scavenging ability as a common feature in defending salt-alkali stress (Meng et al., 2017). In our study, many TF families were differentially expressed under sodicity

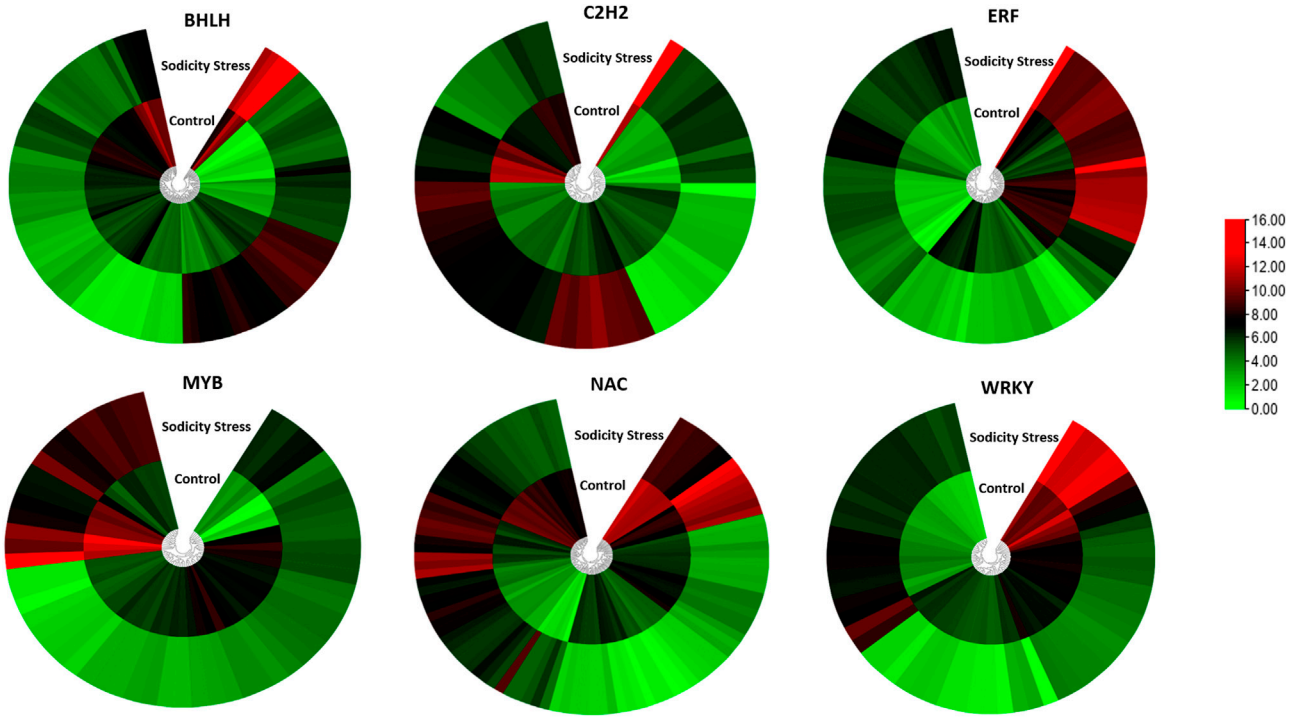


FIGURE 9 | Heatmap representing differential expression of sodicity-responsive TF families in wheat leaf samples.

stress, for instance, *bHLH*, *NAC*, *ERF*, *WRKY*, *C2H2*, *MYB*, etc., were highly enriched in the sodicity stress treatments (**Figure 9**). Past studies have reported involvement of these TFs in response to environmental stresses, such as cold, salinity, drought, iron deficiency, and low nitrogen (Wang et al., 2017; Mittal et al., 2018; Guo et al., 2021b; Wang et al., 2017; Guo et al., 2021a). The TF families such as *bHLH*, *NAC*, *WRKY*, *ERF*, and *MYB* were highly enriched in *Tamarix hispida*, *Ziziphus acidojujuba*, and *Glycine soja* roots under sodicity stress condition (Ge et al., 2011; Wang et al., 2014; Guo et al., 2017). Conversely, studies on different plants such as, *Arabidopsis*, rice, chickpea, and tomato demonstrated that high expression of *NAC* transcription factors support in attaining sodicity stress tolerance by regulating stress-responsive genes and enhanced physiological activities (Huang and Dai, 2015). Additionally, a zinc finger protein, is related to the osmotic stress tolerance in *Arabidopsis* (Muthamilarasan et al., 2014; Ma et al., 2016). Similarly, the overexpression of *OsWRKY8* increased tolerance to osmotic stress in *Arabidopsis* by modulating ABA-independent responsive gene (Yu et al., 2010). The high enrichment of these TF families in sodicity environment suggests their involvement in the regulation of sodicity tolerance.

In addition, we also identified 37 DEGs encoding *serine–threonine protein kinases* (*STPK*). Phosphorylation by *STPK* is one of the primary signaling events that occur in response to environmental stresses (Ichimura et al., 2000). Various studies have depicted the role of *STPK* in abiotic stresses (salinity, drought, and cold), in many crops including *Arachis hypogea* (Rudrabhatla and Rajasekharan, 2002), alfalfa (Chinchilla et al., 2003), *Arabidopsis* (Umezawa et al., 2004), *Oryza sativa* (Diédhiou et al., 2008), and *Triticum aestivum* (Mao et al., 2010). For example, in *Arabidopsis*, upregulation of *AtSnRK2.8* enhanced drought tolerance (Umezawa et al., 2004) and high expression of *OsSAPK4* significantly increased the salinity tolerance of transgenic plants (Diédhiou et al., 2008). In wheat, ectopic expression of *TaSnRK2.4* leads to extended primary roots, late seeding development, and improved tolerance to abiotic stresses in *Arabidopsis* (Mao et al., 2010). The fact that a large number of *STPK* genes were highly induced under sodicity condition suggested their involvement in signal transduction under salinity stress.

In plants, metal ion transporters are important proteins involved in ion homeostasis under sodicity stress conditions. Our study revealed differential expression of transporter encoding genes such as potassium transporter, *ABC* transporter, *SWEET* sugar transporter, and so on, in sodicity treated wheat flag leaf. The sugar metabolism, synthesis of amino acids, and enhanced activity of transmembrane kinase activity are important players in conferring enhanced level of salt stress tolerance in plants and sugar accumulation that provides energy for *ABC* transport regulation where amino acids are the key precursors for the synthesis of some important secondary metabolites (Jia et al., 2020). Moreover, our study also revealed that transmembrane transport protein was enriched

TABLE 4 | Polymorphisms identified in the DEGs by comparing with IWGSC Chinese spring (RefSeq v1.0) reference genome.

Annotation	Gene ID	Chr	Position	Reference genome	KRL3-4
ROS scavenging	TraesCS1A02G186600	1A	337,677,633	C	T
			337,677,955	A	C
	TraesCS1B02G113700	1B	133,076,406	G	–
	TraesCS2B02G126200	2B	94,223,974	T	C
			94,224,203	G	A
	TraesCS3A02G297100	3A	531,701,836	GTCGGAAGTGAGT CCCTT	A
			531,701,361	C	GAAA
			531,701,367	AAAG	C
			531,701,386	G	A
			531,701,395	C	TACCGGGCG
			531,701,415	TACCAGGCA	G
			531,701,429	A	AA
			531,701,436	TG	T
			531,702,306	G	T
	TraesCS3B02G471500	3B	720,177,490	C	G
			720,177,653	A	C
	TraesCS3B02G471900	3B	720,263,094	CGTC	TGTG
	TraesCS3D02G305300	3D	419,464,759	C	–
			419,464,777	C	–
	TraesCS3D02G305400	3D	419,515,860	TTAT	TT
			582,055,996	ATAGT	AT
			582,056,035	C	A
			582,056,048	GTATATAACTAGCG C	ATATC
			582,056,107	C	T
			582,056,130	C	G
			582,056,161	C	T
	TraesCS5A02G424000	5A	609,790,147	GGTCTACGACTGCG	GG
	TraesCS5D02G432600	5D	488,783,912	C	–
			488,784,312	C	–
Serine/threonine protein kinase	TraesCS1A02G080600	1A	63,476,951	CAAAAAACCCT	CAAAAAAAACCCT
	TraesCS1B02G098600	1B	105,108,534	C	A
	TraesCS1D02G082500	1D	64,420,937	ATTTTTTCCA	ATTTTTTTCCA
	TraesCS2B02G124100	2B	92,490,345	A	G
			92,490,898	G	C
	TraesCS2D02G107100	2D	59,450,560	ATG	AG
			59,450,572	A	G
Calcium signaling	TraesCS5D02G269400	5D	372,454,779	GTTTTTTTTATTTT CAAA	–
Metal ion transport	TraesCS3A02G258400	3A	480,147,175	GTTTTTTTGTTCTA	F
	TraesCS5A02G043000	5A	38,797,696	C	TCAT

Note. DEGs, differentially expressed genes; KRL 3–4, bread wheat germplasm; ROS, reactive oxygen species.

under sodicity stress, such as *ABC* transporters, phosphate transporters, transmembrane amino acid transporters, and transmembrane anion transporters. It is evident that the tolerance of wheat to sodicity stress was achieved by increasing the activities of transmembrane proteins and ion channels to exchange ions and small molecules.

KEGG pathway analysis was performed to identify related pathways for 1,891 DEGs that were enriched in this study. Our study revealed the involvement of phenylpropanoid biosynthesis, biosynthesis of secondary metabolites, and plant hormone signal transduction pathways in sodicity condition. In plants, multiple studies have reported that these pathways are significant in abiotic and biotic stress response (Goyal et al., 2016). Phenylpropanoid pathway with the third highest number of genes is a metabolic pathway accountable for the production of different plant secondary metabolites having roles in development and stress-related processes (Baxter et al., 2014).

MapMan Pathways Expressed in Wheat Flag Leaf Under Sodicity Stress

In our study, 604 DEGs were mapped using the MapMan tool to analyze the impact of sodicity stress on wheat flag leaves. Our results suggested that in response to sodicity stress, a large number of genes were found to be associated with lipid metabolism, protein homeostasis, cell wall synthesis, and modification. In addition, transcriptional regulation of defense responses, flower development, oxidative stress response, redox

signaling, hormonal regulation, iron homeostasis, and environmental adaptation were also reported for various plant *GRXs*.

Lipids are structural, storage, and signaling molecule components in plant cells. They are the principal constituents of plant membranes and the cuticle (Harwood, 1980). In this study, Mapman analysis showed upregulation of various lipid metabolism genes including 12 genes encoding for *phospholipase*, 9 genes encoding for *phosphocholine phosphatase*, and 1 gene encoding for *glycerol-3-phosphate acyltransferase* under sodicity stress. All three together are involved in the synthesis of *triacylglycerol* (*TAG*) that is an important component of lipids in cell membranes. Prior studies have shown that different stress conditions induce TAG production in vegetative tissues leading to upregulation of several key genes involved during TAG biosynthesis. Abscisic acid (ABA), a plant hormone, accumulates under stress and functions as a signaling molecule to regulate plant development and metabolic pathways. A transcription factor, ABSCISIC ACID INSENSITIVE 4 (ABI4), that is a key component in the ABA signaling pathway was upregulated under stress conditions and was observed to bind the DGAT1 promoter, which induces its expression in *Arabidopsis* (Lu et al., 2020). Therefore, this important observation supports the fact that salt stress generally disturbs the integrity of the plasma membrane and the high level of TAG might be contributing to salt stress tolerance in KRL3-4 by maintaining integrity of plant membranes. Another component involved in lipid metabolism is 3-ketoacyl-CoA synthase (11), which was found to be downregulated under sodicity condition. Earlier research showed that cuticular wax was found to be associated with plant resistance in response to abiotic stress. *3-Ketoacyl-CoA synthase* (*KCS*) catalyzes the biosynthesis of very-long-chain fatty acid (VLCFA) wax precursors. *3-Ketoacyl-CoA* (*KCS*) synthase encoded by three salt-responsive genes (*LOC_Os02g11070*, *LOC_Os05g49900*, and *LOC_Os02g56860*) have a role in wax biosynthesis (Yuenyong et al., 2018). Furthermore, a recent study in navel orange showed that a novel KCS family gene named *CsKCS6* regulate the production of fatty acid precursors involved in wax synthesis and delivers tolerance of transgenic *Arabidopsis* plants in abiotic stress. Their experiment demonstrates that the expression of *CsKCS6* significantly increased the amount of VLCFAs in the cuticular wax in transgenic *Arabidopsis* plants. The mechanism of *KCS6*-like genes to influence cuticle permeability in transgenic plants involves ectopic expression of *CsKCS6*, which decreased the amount of primary alcohols and aldehydes suggesting that wax biosynthetic genes might be suppressed by ectopic expression of *CsKCS6* in transgenic plants. Hence, a significant decrease in leaf chlorophyll permeability, water loss rate, ion leakage, increased root length, and survival rate were observed in the transgenic plants under drought and high salt stress indicating that ectopic expression of *CsKCS6* can reduce the cuticle permeability and enhance plant tolerance to drought and high salt (Guo et al., 2020). Furthermore, two *peroxygenases/caleosin* encoded by *CLO/PXG*-like genes were also found, showing downregulation during sodicity-induced condition in our analysis. *Caleosins* are a calcium sensor carrying one EF-hand calcium-binding motif. A transcriptomic study on fungus has shown that the *CLO/PXG*-like genes have complex patterns of developmental and tissue-specific expression and exhibited response toward biotic and abiotic stresses and play key role in lipid metabolism, signaling, reproduction, and pathogenesis (Rahman et al., 2018). In another study on rice, *caleosin family protein 5 (OsClo5)* or *peroxygenase (PXG)* modulates salt stress tolerance *via* calcium-dependent membrane associated hemoprotein that catalyzes the *PXG* reaction (Blée et al., 2012) indicating its potential roles in enzymatic activities and Ca^{2+} signaling in plants. *OsClo5* interacts with *OsDi19-5* and inhibits transcription of *OsUSP* and *OsMST* and that activity leads to the enhanced sensitivity of rice seedlings in response to salt stress compared with the wild-type, where *OsClo5* with T-DNA knockdown mutation distorted heterodimer formation with *OsDi19-5* and promotes the weak transcriptional repression effect on the two target genes, i.e., *OsUSP* and *OsMST*. This confers the reduced sensitivity to salinity or the enhanced tolerance to salt stress in rice seedlings in comparison with wild-type. Thus, *OsClo5* negatively affected salt stress tolerance in rice in cooperation with the transaction repression control of *OsDi19-5* to two target genes *OsUSP* and *OsMST* in rice (Jing et al., 2021).

In addition to lipid metabolism, many DEGs encoding the Bowman–Birk protease inhibitor (BBI), such as *TraesCS3D02G034200*, *TraesCS3A02G046000*, *TraesCS3A02G046100*, *TraesCS1D02G022000*, *TraesCS5A02G484800*, *TraesCS5B02G498100*, and *TraesCS5D02G498300*, showed upregulation. The Bowman–Birk inhibitor (BBI) is one of the subfamilies of serine protease inhibitors. It has been reported to function in controlling abiotic stresses, such as salinity and drought stresses (Prakash et al., 1996). In a previous study on wheat, a salt-responsive gene *WRSI5* that contains a Bowman–Birk domain was found responsible for tolerance to salt stress. It was characterized from the salt-tolerant cultivar Shangrong No.3 (SR3) and salt-susceptible cultivar JN177. The *WRIS5* gene functions as trypsin inhibitor as it exhibits a high level of trypsin-inhibiting activity, which signifies that BBI suppressed the trypsin activity, thus, promoting elevated *WRSI5* expression level in SR3 roots in salt, drought, or oxidative stress. Furthermore, the SR3 allowed limited transport of K^+ than Na^+ from root to shoot leading to low Na^+ concentration in SR3 leaves resulting in increased leaf growth rate. In addition, overexpression of *WRSI5* in *Arabidopsis* improves seedling growth at 150 mM NaCl condition, suggesting that *WRSI5* controls plant growth rate or long-distance Na^+ transport in SR3 plants under salt stress (Shan et al., 2008). Similarly, BBI family genes expressed in our study might function in sodicity tolerance.

Furthermore, our study showed upregulation of 13 *glutaredoxin* (*GRX*) genes in response to sodicity stress. The *GRX* family proteins are involved in various cellular

functions, such as redox regulation and protection under abiotic and biotic stresses (Lillig et al., 2008). A recent report showed overexpression of *glutaredoxin* (*OsGrx_C7*) ubiquitously expressed in rice including root and shoot, thus indicating positive response in salt induced stress. However, silencing of this gene exhibited increased sensitivity of rice plants to salt stress. Furthermore, *OsGrx_C7* was suggested to be a positive regulator of salt tolerance by reinforcing the expression of transporters employed in Na^+ homeostasis in overexpressing rice plants (Verma et al., 2021). Therefore, *GRX* genes found in our study can be considered to play a role in sodicity tolerance through Na^+ homeostasis. Additionally, eight downregulated DEGs encoding for α-*expansins* were also identified in our study. *Expansins* are key regulators of cell wall formation and required for cell enlargement under various environmental stresses (Wu and Cosgrove, 2000). Several studies have shown that the expression of *expansin*-encoding genes are induced under high salt and tends to promote salt tolerance in plants (Qiu et al., 2014). *Expansin* is a superfamily of plant proteins and comprised of four families, such as α-*expansin*, β-*expansin*, *expansin*-like A, and *expansin*-like B. The α and β-*expansin* proteins entail cell wall loosening activity and are involved in cell expansion and plant developmental processes. Recently, transgenic tobacco plants, including α-*expansin 4* (*EXPA4*) overexpression and RNA interference (RNAi) mutants were assayed for their tolerance to abiotic stress. The study revealed that transgenic plants with altered *EXPA4* demonstrated that RNAi mutants have increased hypersensitivity to salt and drought stress and vice-versa in overexpressed tobacco lines, consequently, leading to lower cell damage, higher fresh weight, soluble sugar, proline accumulation, and induces several stress responsive gene expression. Therefore, *EXPA4* indicated its involvement in salt and drought stress response *via* changes in the expression levels of some stress-responsive genes (Chen et al., 2018). Moreover, downregulation of arabinogalactan proteins (*AGPs*) was also seen in our study. *AGPs* were seen involved in abiotic stress response such as salinity regulating cell wall expansion. Several studies have reported downregulation of *AGPs* in response to salt treatments. For example, in tomato, different *AGPs* were strongly repressed in salt-treated tomato roots (Ouyang et al., 2007). Similarly, two rice *OsFLA* (*OsFLA10/18*) and two wheat *TaFLA3/4* genes were significantly downregulated under salinity (Faik et al., 2006; Ma and Zhao, 2010). *AGPs* also act as a sodium carrier through the mechanism of vesicle trafficking (Garcia de la Garma et al., 2015). Moreover, the tobacco (*Nicotiana tabacum* L.) cells cultured in saline medium, which were deficient in *AGPs* on the plasma membrane, demonstrates a reduced rate of cell enlargement and a decreased cell wall extensibility. However, non-saline tobacco cells showed higher upregulation of *AGPs* in salt-stress indicating an important increase rate in *AGP* diffusion through a much more highly porous pectic network leading to improved cell wall formation (Zhu et al., 1993; Lamport et al., 2006). This mechanism may contribute to sodium homeostasis during salt adaptation to high saline concentrations (Olmos et al., 2017). Overall, our study revealed diverse classes of genes as key mediator of salinity stress response and could be further validated using functional genomics tools to uncover their exact role in salt tolerance.

Nutrient Uptake Regulation in Bread Wheat Germplasm Under Sodicity Stress

Sodic soil is known to decrease the availability of nutrients such as phosphorus required for plant growth and development (Dotaniya and Meena, 2015). In plants, the primary source of phosphorus is inorganic orthophosphate (Pi), but in sodic soil, it forms an insoluble compound with calcium ions and becomes inaccessible (Péret et al., 2011). As a result, a deficiency in Pi stimulates a set of plant adaptive responses with the purpose of decreasing Pi usage and increasing Pi uptake and recycling *via* upregulation of regulators and transporters involved in P-homeostasis (Ticconi and Abel, 2004; Desnos, 2008; Lin et al., 2009; Rouached et al., 2010). Various genes related to low phosphorus stress have been identified in different plant species (Rausch and Bucher, 2002; Jain et al., 2007; Wang et al., 2011). For example, *phosphate transporter 1* (*PHT1*) gene family participates in the uptake of Pi from the soil (Muchhal et al., 1996; Mudge et al., 2002; Shin et al., 2004; Jia et al., 2011). In our study, we have observed one *PHT1* (*Traescs2a02g047300*) gene, which showed upregulation under sodicity stress. *PHT1* transporters are found to be involved in the uptake of phosphate from the soil and transport it to the shoot and mycorrhizal symbiotic interface. In *Arabidopsis*, rice, and maize, *PHT1* expression is strongly upregulated during phosphorus starvation (Smith and Read, 2008; Huang et al., 2011). Similarly, in wheat also, two *PHT1* members, i.e., *TAPHT1.1* and *TaPHT2.1*, overexpressed in Pi-deficient condition (Shukla et al., 2016). Our study also revealed the upregulation of seven *phosphate signaling regulatory protein* (*SPX*), mapped to the 2A, 2B, 2D, 7A, 7B, and 7D chromosomes (*Traescs2a02g169600*, *Traescs2b02g195900*, *Traescs2d02g177100*, *Traescs7a02g376200*, *Traescs7a02g554100*, *Traescs7b02g478000*, *Traescs7d02g372600*). These genes perform multiple functions in plant tolerance to phosphorus starvation and plays a vital role in sensing the concentration of phosphorus in cytosol (Duan et al., 2008; Wild et al., 2016). During Pi starvation, *SPX1* dissociates with the *phosphate response 1* (*PHR1*) TF. *PHR1* controls many Pi related including miRNA399, which targets *phosphate 2* (*PHO2*). Furthermore, *PHO2* reduction leads to *PHT1* and *PHO2* accumulation and, last, an increase in the plant capacity to uptake Pi and translocate it to shoots. Various genes participate in the proper functioning of this signaling pathway, such as *SUMO E3 ligase SIZ1*, *phosphate transporter traffic facilitator* (*PHF1*) and *nitrogen limitation adaptation* (*NLA*) (Lin et al., 2013). According to our MapMan results, we have also found one gene encoding *NLA* (*Traescs5b02g373000*), highly upregulated under sodicity stress. *NLA* functions of the plasma membrane to direct the degradation of PHT1s is a process necessary for Pi uptake ability of plants (Lin et al., 2013). Hence, we can conclude that although we have observed the expression of many transporter proteins and their

respective genes under sodicity stress, the physiological role of Pi transporters in response to sodicity stress is still limited. Understanding the molecular mechanism involved in the regulation of phosphate transporters is very important, due to the constant lowering availability of phosphorus reserves in nature and will facilitate the development of more efficient Pi-utilizing plants.

Cis-Regulatory Elements Identified in Candidate Sodicity Stress-Responsive Genes

Plants undergo adaptation during stress conditions by triggering a network of signaling cascade that leads to activation of a series of target gene expression. Studies have shown that several TFs, when binding to specific *cis*-elements, act as molecular regulators of gene expression during stress responses (Gujjar et al., 2014). PlantCARE analysis revealed recognition of many *cis*-elements in genes expressed in response to sodicity stress (**Table 3**). For instance, A-box, CAAT-box, CGTCA-motif, G-Box, STRE, TGACG-motif, as-1, TATA-box were present in all 19 genes followed by MYB and MYC observed in 18 genes, whereas CCGTCC motif, CCGTCC-box, DRE core, and Sp1 were seen in 17 genes, and ABRE3a, AAGAA-motif, MBS were located in 13 genes. Previous studies have reported that these *cis*-elements in many genes are expressed in response to abiotic stresses. For example, "CGTCA-motif" was reported in drought stress-responsive genes of trifoliate orange and suggested to play crucial role in drought response *via* jasmonic acid-mediated signaling (Xiong et al., 2020). Similarly, "G-box" *cis*-acting DNA regulatory element are recognized by the *bZIP* and *bHLHs* family TFs to mediate responses toward abiotic and biotic stresses (Qian et al., 2021). Prior studies have shown that *bZIP* and *bHLHs* family TFs bind to "G-box" in the promoter of cold, drought, and salinity-responsive genes in *Arabidopsis* and other plant species (Liu et al., 2008; Shah et al., 2021). In wheat, a *TabHLH13* gene that responds *via* "G-box" elements were upregulated under salt stress (Fu et al., 2014). Another important *cis*-element, i.e., "STRE" is recognized by *zinc-finger DNA binding* TF (Estruch, 2000). "STRE" is functional in both orientations ("CCCCT" or "AGGGG"). A study conducted on yeast showed that the "STRE" *cis*-element located in the promoter of *CTT1* and *DDR2*, a DNA damage-responsive gene mediates transcriptional induction in response to different stress conditions (Kobayashi and McEntee, 1993). In contrast, "TGACG-motif," with consensus sequence "TGACG" is targeted by MeJA family TFs as it acts as JA-responsive element (Liao et al., 2017). In rice, two "TGACG-motif" (JA-responsive element) *cis*-elements were found in the promoter region of *OsTMP14* that is related to stress tolerance. Subsequently, activation sequence-1 ("as-1") with consensus sequence "TGACG" situated in all identified genes showed the role in abiotic stress and salicylic acid responsiveness (Banerjee et al., 2015). In the current study, two *cis*-acting elements ("*MYB*" and "*MYC*") were located in 18 candidate genes that might be involved in sodicity stress tolerance. The *MYB* cis-element with consensus sequence "CAACAG" mostly interacts with MYB TFs. In past reports, overexpression of the wheat MYB TFs, such as *TaMYB30-B*, *TaMYB33*, *TaMYB56-B*, and *TaSIM*, import tolerance to salt and drought stresses in transgenic *Arabidopsis* (Yu et al., 2017), whereas *MYC cis*-element with consensus sequence "CATGTG" functions in salt and drought stress response through interaction with *bHLH* and *NAC* family TFs. The role of "*MYC*" has also been seen in drought tolerance through the dehydration-inducible expression of *EARLY RESPONSIVE TO DEHYDRATION STRESS 1 (ERD1)* gene in *Arabidopsis*. Since *ERD1* promoter contains the "CATGTG" motif that interacts with NAC, which specifically binds to the MYC-like sequence "CATGTG," the expression of NAC TFs genes, such as *ANAC019*, *ANAC055*, and *ANAC072*, was induced by drought, high salinity, and abscisic acid (Tran et al., 2004).

Dehydration-responsive element (*DRE*) core having "GCCGAC" consensus sequence is found in promoters of various environmental stress-responsive genes. DREB TFs binds at "DRE-core" (Liu et al., 1998). In *Arabidopsis*, overexpression of *DREB1A* promotes enhanced expression of downstream drought stress-responsive genes (Liu et al., 1998). In addition, DREB proteins also have a role in phytohormone signaling pathway such as abscisic acid, salicylic acid, jasmonate acid, ethylene, and gibberellic acid (Choi et al., 2002). "Sp1" is another *cis*-element present in 17 identified candidate genes. *Homeodomain leucine zipper* (*HD-ZIP*) TFs targets the consensus sequence "GGGCGG" of "Sp1" for abiotic stress response. The presence of "Sp1" in the promoter region of banana *HD-ZIP* genes have shown its participation in various functions and were found to be upregulated by drought and salinity stress (Pandey et al., 2016).

Importance of Polymorphism in the Expressed Sequences

SNP identification may have a critical role in conferring sodicity stress tolerance in wheat flag leaf. We noticed 39 SNPs mapped to the important sodicity stress-responsive genes associated with various pathways, such as ROS scavenging, serine/threonine protein kinase, calcium signaling, and metal ion transporters) that are known to have a role in abiotic stress responses. From these 39, a total of 29 SNPs were mapped to ROS-scavenging genes that were differentially expressed during sodicity stress condition, thus, stimulating stress tolerance through glutathione metabolism and the biosynthesis of secondary metabolite pathways in wheat. In addition, SNPs were also mapped to highly upregulated metal ion transport, AMPK signaling, and ABC transport activity (**Table 4**). These results suggest that besides variation in expression pattern, sequence level polymorphism in key abiotic stress genes may also have a role in conferring enhanced level of sodicity stress tolerance in KRL 3–4.

CONCLUSION

In summary, our study provided detailed insights into genes, regulatory elements, and various metabolic pathways that could be important for enhanced sodicity stress tolerance in bread wheat genotype KRL3–4. Although many metabolic pathways were altered, lipid metabolism and protein homeostasis-related pathways appear to have a most important role in sodicity tolerance. The genomic information, particularly 19 candidate genes identified in our study, could be further validated using functional genomics approaches for understanding their detailed role in sodicity stress tolerance in wheat.

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/, PRJNA763033.

AUTHOR CONTRIBUTIONS

GP handled the investigation, formal analysis, and writing of the original draft. SM contributed in the data curation, investigation, formal analysis, writing, reviewing, and editing, AK was involved in resource acquisition, writing, reviewing, and editing of the draft. DC performed data analysis, writing, reviewing, and editing, TS, SK, and RS participated in the writing, reviewing, and editing. AS contributed in the conceptualization, supervision, resources, writing, reviewing, and editing of the manuscript.

FUNDING

Financial support received from ICAR-National Innovations in Climate Resilient Agriculture (NICRA) Project code 1006607.

ACKNOWLEDGMENTS

This work was supported by ICAR-NBPGR under PROJECT PGR/DGR-BUR-DEL-01.02 and No. BT/Ag/Network/Wheat/2019-20. The authors are also thankful to the Head of Division, Genomic Resources, and Director, ICAR-NBPGR, New Delhi, for extending lab facilities for this work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.782366/full#supplementary-material

REFERENCES

An, J.-P., Li, H.-H., Song, L.-Q., Su, L., Liu, X., You, C.-X., et al. (2016). The Molecular Cloning and Functional Characterization of MdMYC2, a bHLH Transcription Factor in Apple. *Plant Physiol. Biochem.* 108, 24–31. doi:10.1016/j.plaphy.2016.06.032

Anders, S., and Huber, W. (2010). Differential Expression Analysis for Sequence Count Data. *Genome Biol.* 11 (11), 1–12. doi:10.1186/GB-2010-11-10-R106

Anzooman, M., Christopher, J., Dang, Y. P., Taylor, J., Menzies, N. W., and Kopittke, P. M. (2019). Chemical and Physical Influence of Sodic Soils on the Coleoptile Length and Root Growth Angle of Wheat Genotypes. *Ann. Bot.* 124, 1043–1052. doi:10.1093/AOB/MCZ094

Banerjee, J., Sahoo, D. K., Raha, S., Sarkar, S., Dey, N., and Maiti, I. B. (2015). A Region Containing an As-1 Element of Dahlia Mosaic Virus (DaMV) Subgenomic Transcript Promoter Plays a Key Role in Green Tissue- and Root-specific Expression in Plants. *Plant Mol. Biol. Rep.* 33, 532–556. doi:10.1007/s11105-014-0766-5

Baxter, A., Mittler, R., and Suzuki, N. (2014). ROS as Key Players in Plant Stress Signalling. *J. Exp. Bot.* 65, 1229–1240. doi:10.1093/jxb/ert375

Blée, E., Flenet, M. F., Boachon, B., and Fauconnier, M.-L. (2012). A Non-canonical Caleosin fromArabidopsisefficiently Epoxidizes Physiological Unsaturated Fatty Acids with Complete Stereoselectivity. *FEBS J.* 279, 3981–3995. doi:10.1111/J.1742-4658.2012.08757.X

Chaurasia, S., Singh, A. K., Kumar, A., Songachan, L. S., Yadav, M. C., Kumar, S., et al. (2021). Genome-wide Association Mapping Reveals Key Genomic Regions for Physiological and Yield-Related Traits under Salinity Stress in Wheat (*Triticum aestivum* L.). *Genomics* 113, 3198–3215. doi:10.1016/j.ygeno.2021.07.014

Chen, L.-J., Zou, W.-S., Fei, C.-Y., Wu, G., Li, X.-Y., Lin, H.-H., et al. (2018). α-Expansin EXPA4 Positively Regulates Abiotic Stress Tolerance but Negatively Regulates Pathogen Resistance in Nicotiana Tabacum. *Plant Cel Physiol* 59, 2317–2330. doi:10.1093/PCP/PCY155

Chen, Y., and Barak, P. (1982). Iron Nutrition of Plants in Calcareous Soils. *Adv. Agron.* 35, 217–240. doi:10.1016/S0065-2113(08)60326-0

Chinchilla, D., Merchan, F., Megias, M., Kondorosi, A., Sousa, C., and Crespi, M. (2003). Ankyrin Protein Kinases: A Novel Type of Plant Kinase Gene Whose Expression Is Induced by Osmotic Stress in Alfalfa. *Plant Mol. Biol.* 51, 555–566. doi:10.1023/A:1022337221225

Choi, D.-W., Rodriguez, E. M., and Close, T. J. (2002). Barley Cbf3 Gene Identification, Expression Pattern, and Map Location. *Plant Physiol.* 129, 1781–1787. doi:10.1104/PP.003046

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., et al. (2012). A Program for Annotating and Predicting the Effects of Single Nucleotide Polymorphisms, SnpEff. *Fly* 6, 80–92. doi:10.4161/fly.19695

Desnos, T. (2008). Root Branching Responses to Phosphate and Nitrate. *Curr. Opin. Plant Biol.* 11, 82–87. doi:10.1016/J.PBI.2007.10.003

Diédhiou, C. J., Popova, O. V., Dietz, K.-J., and Golldack, D. (2008). The SNF1-type Serine-Threonine Protein Kinase SAPK4regulates Stress-Responsive Gene Expression in rice. *BMC Plant Biol.* 8, 1. doi:10.1186/1471-2229-8-49

Doi, K., Hosaka, A., Nagata, T., Satoh, K., Suzuki, K., Mauleon, R., et al. (2008). Development of a Novel Data Mining Tool to Find Cis-Elements in rice Gene Promoter Regions. *BMC Plant Biol.* 8, 1. doi:10.1186/1471-2229-8-20

Dotaniya, M. L., and Meena, V. D. (2015). Rhizosphere Effect on Nutrient Availability in Soil and its Uptake by Plants: A Review. *Proc. Natl. Acad. Sci. India, Sect. B Biol. Sci.* 85, 1–12. doi:10.1007/S40011-013-0297-0

Duan, K., Yi, K., Dang, L., Huang, H., Wu, W., and Wu, P. (2008). Characterization of a Sub-family of Arabidopsis Genes with the SPX Domain Reveals Their Diverse Functions in Plant Tolerance to Phosphorus Starvation. *Plant J.* 54, 965–975. doi:10.1111/J.1365-313X.2008.03460.X

Elamin, F., Abdelazeem, N., Salah, I., Mirghani, Y., and Wong, F. (2019). A Randomized Clinical Trial Comparing Hall vs Conventional Technique in Placing Preformed Metal Crowns from Sudan. *PLoS One* 14, e0217740. doi:10.1371/journal.pone.0217740

Estruch, F. (2000). Stress-controlled Transcription Factors, Stress-Induced Genes and Stress Tolerance in Budding Yeast. *FEMS Microbiol. Rev.* 24, 469–486. doi:10.1111/j.1574-6976.2000.tb00551.x

Faik, A., Abouzouhair, J., and Sarhan, F. (2006). Putative Fasciclin-like Arabinogalactan-Proteins (FLA) in Wheat (*Triticum aestivum*) and rice (Oryza Sativa): Identification and Bioinformatic Analyses. *Mol. Genet. Genomics* 276, 478–494. doi:10.1007/S00438-006-0159-Z

Fu, H., Liu, L., Dong, Z., Guo, S., and Gao, H. (2018). Dissociation between Iron and Heme Biosyntheses Is Largely Accountable for Respiration Defects of Shewanella Oneidensis Fur Mutants. *Appl. Environ. Microbiol.* 84, e00039. doi:10.1128/AEM.00039-18

Fu, S., Liu, G., Zhang, L., Xia, C., Zhao, G., and Jia, J. (2014). Cloning and Analysis of a Salt Stress Related Gene TabHLH13 in Wheat. *Plant Genet. Resour.* 15, 1006–1011. Available at: https://scholar.google.com/scholar_lookup?title=Cloning+and+Analysis+of+a+Salt+Stress+Related+Gene+TabHLH13+in+Wheat.&journal=J.+Plant+Genet.+Resour.&author=Fu+S.+L.&author=Liu+G.+X.&author=Zhang+L.+C.&author=Xia+C.&author=Zhao+G.+Y.&author=Jia+J.+Z.&publication_year=2014&volume=15&pages=1006–1011 (Accessed September 15, 2021). doi:10.1016/S1875-2780(08)60090-5

Fujimori, S., Hasegawa, T., Krey, V., Riahi, K., Bertram, C., Bodirsky, B. L., et al. (2019). A Multi-Model Assessment of Food Security Implications of Climate Change Mitigation. *Nat. Sustain.* 2 (2), 386–396. doi:10.1038/s41893-019-0286-2

Garcia de la Garma, J., Fernandez-Garcia, N., Bardisi, E., Pallol, B., Asensio-Rubio, J. S., Bru, R., et al. (2015). New Insights into Plant Salt Acclimation: the Roles of Vesicle Trafficking and Reactive Oxygen Species Signalling in Mitochondria and the Endomembrane System. *New Phytol.* 205, 216–239. doi:10.1111/NPH.12997

Garrison, E., and Marth, G. (2012). Haplotype-based Variant Detection from Short-Read Sequencing. Available at: https://arxiv.org/abs/1207.3907v2 (Accessed September 15, 2021).

Ge, Y., Li, Y., Lv, D.-K., Bai, X., Ji, W., Cai, H., et al. (2011). Alkaline-stress Response in Glycine Soja Leaf Identifies Specific Transcription Factors and ABA-Mediated Signaling Factors. *Funct. Integr. Genomics* 11, 369–379. doi:10.1007/s10142-010-0191-2

Goyal, E., Amit, S. K., Singh, R. S., Mahato, A. K., Chand, S., and Kanika, K. (2016). Transcriptome Profiling of the Salt-Stress Response in *Triticum aestivum* Cv. Kharchia Local. *Sci. Rep.* 6, 1–14. doi:10.1038/srep27752

Gujjar, R. S., Akhtar, M., and Singh, M. (2014). Transcription Factors in Abiotic Stress Tolerance. *Ind. J. Plant Physiol.* 19, 306–316. doi:10.1007/s40502-014-0121-8

Guo, M., Li, S., Tian, S., Wang, B., and Zhao, X. (2017). Transcriptome Analysis of Genes Involved in Defense against Alkaline Stress in Roots of Wild Jujube (Ziziphus Acidojujuba). *PLoS One* 12, e0185732. doi:10.1371/journal.pone.0185732

Guo, W., Wu, Q., Yang, L., Hu, W., Liu, D., and Liu, Y. (2020). Ectopic Expression of CsKCS6 from Navel Orange Promotes the Production of Very-Long-Chain Fatty Acids (VLCFAs) and Increases the Abiotic Stress Tolerance of *Arabidopsis thaliana*. *Front. Plant Sci.* 11, 1503. doi:10.3389/FPLS.2020.564656

Guo, Y., Ren, G., Zhang, K., Li, Z., Miao, Y., and Guo, H. (2021a). Leaf Senescence: Progression, Regulation, and Application. *Mol. Hortic.* 1, 1–25. doi:10.1186/s43897-021-00006-9

Guo, Y., Yu, Z.-Y., Wu, J., Gong, H., Kesteven, S., Iismaa, S. E., et al. (2021b). The Ca2+-Activated Cation Channel TRPM4 Is a Positive Regulator of Pressure Overload-Induced Cardiac Hypertrophy. *Elife* 10, e66582. doi:10.7554/ELIFE.66582

Hao, L., Shi, S., Guo, H., Zhang, J., Li, P., and Feng, Y. (2021). Transcriptome Analysis Reveals Differentially Expressed MYB Transcription Factors Associated With Silicon Response in Wheat. *Sci. Rep.* 111 (11), 1–9. doi:10.1038/s41598-021-83912-8

Harwood, J. L. (1980). Plant Acyl Lipids: Structure, Distribution, and Analysis. *Lipids Struct. Funct.*, 1–55. doi:10.1016/B978-0-12-675404-9.50007-2

Huang, C. Y., Shirley, N., Genc, Y., Shi, B., and Langridge, P. (2011). Phosphate Utilization Efficiency Correlates with Expression of Low-Affinity Phosphate Transporters and Noncoding RNA, IPS1, in Barley. *Plant Physiol.* 156, 1217–1229. doi:10.1104/PP.111.178459

Huang, D., and Dai, W. (2015). Molecular Characterization of the Basic helix-loop-helix (bHLH) Genes that Are Differentially Expressed and Induced by Iron Deficiency in Populus. *Plant Cel Rep* 34, 1211–1224. doi:10.1007/s00299-015-1779-8

Huang, W., He, Y., Yang, L., Lu, C., Zhu, Y., Sun, C., et al. (2021). Genome-Wide Analysis of Growth-Regulating Factors (GRFs) in Triticum aestivum. *PeerJ* 9, 1–24. doi:10.7717/peerj.10701

Ichimura, K., Mizoguchi, T., Yoshida, R., Yuasa, T., and Shinozaki, K. (2000). Various Abiotic Stresses Rapidly Activate Arabidopsis MAP Kinases ATMPK4 and ATMPK6. *Plant J.* 24, 655–665. doi:10.1046/j.1365-313X.2000.00913.x

Ijaz, U., Pervaiz, T., Ahmed, T., Seemab, R., Shahid, M., Noman, M., et al. (2020). Plant Cis-Regulatory Elements: Methods of Identification and Applications. *Asian J. Agric. Biol.* 8, 207–222. doi:10.35495/AJAB.2019.08.352

Jain, A., Vasconcelos, M. J., Raghothama, K. G., and Sahi, S. V. (2007). Molecular Mechanisms of Plant Adaptation to Phosphate Deficiency. *Plant Breed. Rev.* 29, 359–419. doi:10.1002/9780470168035.CH7

Jia, H., Ren, H., Gu, M., Zhao, J., Sun, S., Zhang, X., et al. (2011). The Phosphate Transporter Gene OsPht1;8 Is Involved in Phosphate Homeostasis in Rice. *Plant Physiol.* 156, 1164–1175. doi:10.1104/PP.111.175240

Jia, Y., Li, X., Liu, Q., Hu, X., Li, J., Dong, R., et al. (2020). Physiological and Transcriptomic Analyses Reveal the Roles of Secondary Metabolism in the Adaptive Responses of Stylosanthes to Manganese Toxicity. *BMC Genomics* 21, 1–17. doi:10.1186/s12864-020-07279-2

Jing, P., Kong, D., Ji, L., Kong, L., Wang, Y., Peng, L., et al. (2021). OsClo5 Functions as a Transcriptional Co-repressor by Interacting with OsDi19-5 to Negatively Affect Salt Stress Tolerance in rice Seedlings. *Plant J.* 105, 800–815. doi:10.1111/TPJ.15074

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017). KEGG: New Perspectives on Genomes, Pathways, Diseases and Drugs. *Nucleic Acids Res.* 45, D353–D361. doi:10.1093/NAR/GKW1092

Kaur, G., and Pati, P. K. (2016). Analysis of Cis-Acting Regulatory Elements of Respiratory Burst Oxidase Homolog (Rboh) Gene Families in Arabidopsis and rice Provides Clues for Their Diverse Functions. *Comput. Biol. Chem.* 62, 104–118. doi:10.1016/j.compbiolchem.2016.04.002

Khokhar, J. S., Sareen, S., Tyagi, B. S., Singh, G., Wilson, L., King, I. P., et al. (2018). Variation in Grain Zn Concentration, and the Grain Ionome, in Field-Grown Indian Wheat. *PLoS One* 13, e0192026. doi:10.1371/JOURNAL.PONE.0192026

Kobayashi, N., and McEntee, K. (1993). Identification of Cis and Trans Components of a Novel Heat Shock Stress Regulatory Pathway in *Saccharomyces cerevisiae*. *Mol. Cel. Biol.* 13, 248–256. doi:10.1128/MCB.13.1.248-256.1993

Lamport, D. T. A., Kieliszewski, M. J., and Showalter, A. M. (2006). Salt Stress Upregulates Periplasmic Arabinogalactan Proteins: Using Salt Stress to Analyse AGP Function*. *New Phytol.* 169, 479–492. doi:10.1111/J.1469-8137.2005.01591.X

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, M., Guo, R., Jiao, Y., Jin, X., Zhang, H., and Shi, L. (2017). Comparison of Salt Tolerance in Soja Based on Metabolomics of Seedling Roots. *Front. Plant Sci.* 8, 1101. doi:10.3389/FPLS.2017.01101/BIBTEX

Li, S., Li, X., Wei, Z., and Liu, F. (2020). ABA-mediated Modulation of Elevated CO_2 on Stomatal Response to Drought. *Curr. Opin. Plant Biol.* 56, 174–180. doi:10.1016/j.pbi.2019.12.002

Li, Z., Tian, Y., Xu, J., Fu, X., Gao, J., Wang, B., et al. (2018). A Tomato ERF Transcription Factor, SlERF84, Confers Enhanced Tolerance to Drought and Salt Stress but Negatively Regulates Immunity against *Pseudomonas syringae* Pv. Tomato DC3000. *Plant Physiol. Biochem.* 132, 683–695. doi:10.1016/j.plaphy.2018.08.022

Liao, Y., Jiang, Y., Xu, J., Hu, C., Quan, C., Zhou, J., et al. (2017). Overexpression of a Thylakoid Membrane Protein geneOsTMP14improves Indica rice Cold Tolerance. *Biotechnol. Biotechnological Equipment* 31(4), 1–8. Available at: http://mc.manuscriptcentral.com/tbeq31. doi:10.1080/13102818.2017.1334590

Lillig, C. H., Berndt, C., and Holmgren, A. (2008). Glutaredoxin Systems. *Biochim. Biophys. Acta (Bba) - Gen. Subjects* 1780, 1304–1317. doi:10.1016/J.BBAGEN.2008.06.003

Lin, W.-Y., Huang, T.-K., and Chiou, T.-J. (2013). NITROGEN LIMITATION ADAPTATION, a Target of MicroRNA827, Mediates Degradation of Plasma

Membrane-Localized Phosphate Transporters to Maintain Phosphate Homeostasis in Arabidopsis. *Plant Cell* 25, 4061–4074. doi:10.1105/TPC.113.116012

Lin, W.-Y., Lin, S.-I., and Chiou, T.-J. (2009). Molecular Regulators of Phosphate Homeostasis in Plants. *J. Exp. Bot.* 60, 1427–1438. doi:10.1093/JXB/ERN303

Liu, J. X., Srivastava, R., and Howell, S. H. (2008). Stress-induced Expression of an Activated Form of AtbZIP17 Provides Protection From Salt Stress in Arabidopsis. *Plant, Cell Environ.* 31, 1735–1743. doi:10.1111/j.1365-3040.2008.01873.x

Liu, L., Xia, W., Li, H., Zeng, H., Wei, B., Han, S., et al. (2018). Salinity Inhibits Rice Seed Germination by Reducing α-Amylase Activity via Decreased Bioactive Gibberellin Content. *Front. Plant Sci.* 9, 1–9. doi:10.3389/fpls.2018.00275

Liu, Q., Kasuga, M., Sakuma, Y., Abe, H., Miura, S., Yamaguchi-Shinozaki, K., et al. (1998). Two Transcription Factors, DREB1 and DREB2, with an EREBP/AP2 DNA Binding Domain Separate Two Cellular Signal Transduction Pathways in Drought- and Low-Temperature-Responsive Gene Expression, Respectively, in Arabidopsis. *The Plant Cell* 10, 1391–1406. doi:10.1105/tpc.10.8.1391

Lu, J., Xu, Y., Wang, J., Singer, S. D., and Chen, G. (2020). The Role of Triacylglycerol in Plant Stress Response. *Plants* 9, 472. doi:10.3390/PLANTS9040472

Ma, H., and Zhao, J. (2010). Genome-wide Identification, Classification, and Expression Analysis of the Arabinogalactan Protein Gene Family in rice (Oryza Sativa L.). *J. Exp. Bot.* 61, 2647–2668. doi:10.1093/JXB/ERQ104

Ma, X., Liang, W., Gu, P., and Huang, Z. (2016). Salt Tolerance Function of the Novel C2H2-type Zinc finger Protein TaZNF in Wheat. *Plant Physiol. Biochem.* 106, 129–140. doi:10.1016/j.plaphy.2016.04.033

Maas, E. V. (2019). "Salt Tolerance of Plants," in *CRC Handbook of Plant Science in Agriculture Volume II*, 57–76.

Mao, X., Zhang, H., Tian, S., Chang, X., and Jing, R. (2010). TaSnRK2.4, an SNF1-type Serine/threonine Protein Kinase of Wheat (*Triticum aestivum* L.), Confers Enhanced Multistress Tolerance in Arabidopsis. *J. Exp. Bot.* 61, 683–696. doi:10.1093/jxb/erp331

Meng, C., Quan, T.-Y., Li, Z.-Y., Cui, K.-L., Yan, L., Liang, Y., et al. (2017). Transcriptome Profiling Reveals the Genetic Basis of Alkalinity Tolerance in Wheat. *BMC Genomics* 18, 1–14. doi:10.1186/s12864-016-3421-8

Minhas, P. S., Qadir, M., and Yadav, R. K. (2019). Groundwater Irrigation Induced Soil Sodification and Response Options. *Agric. Water Manage.* 215, 74–85. doi:10.1016/J.AGWAT.2018.12.030

Mittal, S., Banduni, P., Mallikarjuna, M. G., Rao, A. R., Jain, P. A., Dash, P. K., et al. (2018). Structural, Functional, and Evolutionary Characterization of Major Drought Transcription Factors Families in maize. *Front. Chem.* 6, 177. doi:10.3389/fchem.2018.00177

Msimbira, L. A., and Smith, D. L. (2020). The Roles of Plant Growth Promoting Microbes in Enhancing Plant Tolerance to Acidity and Alkalinity Stresses. *Front. Sustain. Food Syst.* 4, 106. doi:10.3389/FSUFS.2020.00106

Muchhal, U. S., Pardo, J. M., and Raghothama, K. G. (1996). Phosphate Transporters from the Higher Plant *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci.* 93, 10519–10523. doi:10.1073/PNAS.93.19.10519

Mudge, S. R., Rae, A. L., Diatloff, E., and Smith, F. W. (2002). Expression Analysis Suggests Novel Roles for Members of the Pht1 Family of Phosphate Transporters in Arabidopsis. *Plant J.* 31, 341–353. doi:10.1046/J.1365-313X.2002.01356.X

Muthamilarasan, M., Bonthala, V. S., Mishra, A. K., Khandelwal, R., Khan, Y., Roy, R., et al. (2014). C2H2 Type of Zinc finger Transcription Factors in Foxtail Millet Define Response to Abiotic Stresses. *Funct. Integr. Genomics* 14, 531–543. doi:10.1007/s10142-014-0383-2

Olmos, E., García De La Garma, J., Gomez-Jimenez, M. C., and Fernandez-Garcia, N. (2017). Arabinogalactan Proteins Are Involved in Salt-Adaptation and Vesicle Trafficking in Tobacco By-2 Cell Cultures. *Front. Plant Sci.* 8, 1–14. doi:10.3389/fpls.2017.01092

Ouyang, B., Yang, T., Li, H., Zhang, L., Zhang, Y., Zhang, J., et al. (2007). Identification of Early Salt Stress Response Genes in Tomato Root by Suppression Subtractive Hybridization and Microarray Analysis. *J. Exp. Bot.* 58, 507–520. doi:10.1093/JXB/ERL258

Pandey, A., Misra, P., Alok, A., Kaur, N., Sharma, S., Lakhwani, D., et al. (2016). Genome-Wide Identification and Expression Analysis of Homeodomain Leucine Zipper Subfamily IV (HDZ IV) Gene Family from Musa Accuminata. *Front. Plant Sci.* 7, 20. doi:10.3389/FPLS.2016.00020

Péret, B., Clément, M., Nussaume, L., and Desnos, T. (2011). Root Developmental Adaptation to Phosphate Starvation: Better Safe Than Sorry. *Trends Plant Sci.* 16, 442–450. doi:10.1016/J.TPLANTS.2011.05.006

Prakash, B., Selvaraj, S., Murthy, M. R. N., Sreerama, Y. N., Rajagopal Rao, D., and Gowda, L. R. (1996). Analysis of the Amino Acid Sequences of Plant Bowman-Birk Inhibitors. *J. Mol. Evol.* 42, 560–569. doi:10.1007/BF02352286

Qadir, M., Quillérou, E., Nangia, V., Murtaza, G., Singh, M., Thomas, R. J., et al. (2014). Economics of Salt-Induced Land Degradation and Restoration. *Nat. Resour. Forum* 38, 282–295. doi:10.1111/1477-8947.12054

Qian, Y., Zhang, T., Yu, Y., Gou, L., Yang, J., Xu, J., et al. (2021). Regulatory Mechanisms of bHLH Transcription Factors in Plant Adaptive Responses to Various Abiotic Stresses. *Front. Plant Sci.* 12, 1143. doi:10.3389/fpls.2021.677611

Qiu, S., Ma, N., Che, S., Wang, Y., Peng, X., Zhang, G., et al. (2014). Repression of OsEXPA3 Expression Leads to Root System Growth Suppression in Rice. *Crop Sci.* 54, 2201–2213. doi:10.2135/CROPSCI2013.11.0746

Rahman, F., Hassan, M., Hanano, A., Fitzpatrick, D. A., McCarthy, C. G. P., and Murphy, D. J. (2018). Evolutionary, Structural and Functional Analysis of the Caleosin/peroxygenase Gene Family in the Fungi. *BMC Genomics* 19, 1–24. doi:10.1186/s12864-018-5334-1

Ramadas, S., Kiran Kumar, T. M., and Pratap Singh, G. (2020). Wheat Production in India: Trends and Prospects. *Recent Adv. Grain Crop Res.* doi:10.5772/INTECHOPEN.86341

Rausch, C., and Bucher, M. (2002). Molecular Mechanisms of Phosphate Transport in Plants. *Planta* 216, 23–37. doi:10.1007/S00425-002-0921-3

Rombauts, S., Dehais, P., Van Montagu, M., and Rouze, P. (1999). PlantCARE, a Plant Cis-Acting Regulatory Element Database. *Nucleic Acids Res.* 27, 295–296. doi:10.1093/NAR/27.1.295

Rouached, H., Arpat, A. B., and Poirier, Y. (2010). Regulation of Phosphate Starvation Responses in Plants: Signaling Players and Cross-Talks. *Mol. Plant* 3, 288–299. doi:10.1093/MP/SSP120

Rudrabhatla, P., and Rajasekharan, R. (2002). Developmentally Regulated Dual-Specificity Kinase from Peanut that Is Induced by Abiotic Stresses. *Plant Physiol.* 130, 380–390. doi:10.1104/pp.005173

Shah, W. H., Rasool, A., Saleem, S., Mushtaq, N. U., Tahir, I., Hakeem, K. R., et al. (2021). Understanding the Integrated Pathways and Mechanisms of Transporters, Protein Kinases, and Transcription Factors in Plants under Salt Stress. *Int. J. Genomics* 2021, 1–16. doi:10.1155/2021/5578727

Shan, L., Li, C., Chen, F., Zhao, S., and Xia, G. (2008). A Bowman-Birk Type Protease Inhibitor Is Involved in the Tolerance to Salt Stress in Wheat. *Plant Cel Environ* 31, 1128–1137. doi:10.1111/j.1365-3040.2008.01825.x

Sheoran, P., Kumar, A., Singh, A., Kumar, A., Parjapat, K., Sharma, R., et al. (2021). Pressmud Alleviates Soil Sodicity Stress in a rice-wheat Rotation: Effects on Soil Properties, Physiological Adaptation and Yield-related Traits. *Land Degrad. Dev.* 32, 2735–2748. doi:10.1002/LDR.3953

Sheshadri, S. A., Nishanth, M. J., and Simon, B. (2016). Stress-Mediated Cis-Element Transcription Factor Interactions Interconnecting Primary and Specialized Metabolism in Planta. *Front. Plant Sci.* 7, 1–23. doi:10.3389/fpls.2016.01725

Shi, H., and Zhu, J.-K. (2002). Regulation of Expression of the Vacuolar Na+/H+ Antiporter Gene AtNHX1 by Salt Stress and Abscisic Acid. *Plant Mol. Biol.* 50, 543–550. doi:10.1023/A:1019859319617

Shin, H., Shin, H.-S., Dewbre, G. R., and Harrison, M. J. (2004). Phosphate Transport inArabidopsis: Pht1;1 and Pht1;4 Play a Major Role in Phosphate Acquisition from Both Low- and High-Phosphate Environments. *Plant J.* 39, 629–642. doi:10.1111/J.1365-313X.2004.02161.X

Shinozaki, K., and Yamaguchi-Shinozaki, K. (2007). Gene Networks Involved in Drought Stress Response and Tolerance. *J. Exp. Bot.* 58, 221–227. doi:10.1093/jxb/erl164

Shukla, V., Kaur, M., Aggarwal, S., Bhati, K. K., Kaur, J., Mantri, S., et al. (2016). Tissue Specific Transcript Profiling of Wheat Phosphate Transporter Genes and its Association with Phosphate Allocation in Grains. *Sci. Rep.* 6, 1–12. doi:10.1038/srep39293

Singh, A. K., Chaurasia, S., Kumar, S., Singh, R., Kumari, J., Yadav, M. C., et al. (2018). Identification, Analysis and Development of Salt Responsive Candidate Gene Based SSR Markers in Wheat. *BMC Plant Biol.* 18, 1–15. doi:10.1186/S12870-018-1476-1

Singh, K., Kulshreshtha, N., Chatrath, R., Sharma, S. K., and Yaduvanshi, N. (2010). KRL 3–4 (IC408331; INGR09087), a Wheat (*Triticum aestivum*) Germplasm

with Salt Tolerance, Water Logging Tolerance, Red Grain, Low Sodium Uptake under Salinity. *Singh, KN Kulshreshtha, Neeraj Chatrath, Ravish Sharma, Surinder Kumar Yaduvanshi, NPS* 23. Available at: http://www.indianjournals.com/ijor.aspx?target=ijor:ijpgr&volume=23&issue=3&article=038 (Accessed September 14, 2021).

Singh, R. P., Jha, P., and Jha, P. N. (2015). The Plant-Growth-Promoting Bacterium Klebsiella Sp. SBP-8 Confers Induced Systemic Tolerance in Wheat (*Triticum aestivum*) under Salt Stress. *J. Plant Physiol.* 184, 57–67. doi:10.1016/j.jplph.2015.07.002

Smith, S., and Read, D. (2008). *Mycorrhizal Symbiosis*. Academic Press

Subramaniam, Y., and Masron, T. A. (2021). Food Security and Environmental Degradation: Evidence from Developing Countries. *GeoJournal* 86, 1141–1153. doi:10.1007/S10708-019-10119-W

Ticconi, C., and Abel, S. (2004). Short on Phosphate: Plant Surveillance and Countermeasures. *Trends Plant Sci.* 9, 548–555. doi:10.1016/J.TPLANTS.2004.09.003

Tran, L.-S. P., 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[W]. *Plant Cell* 16, 2481–2498. doi:10.1105/tpc.104.022699

Umezawa, T., Yoshida, R., Maruyama, K., Yamaguchi-Shinozaki, K., and Shinozaki, K. (2004). SRK2C, a SNF1-Related Protein Kinase 2, Improves Drought Tolerance by Controlling Stress-Responsive Gene Expression in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci.* 101, 17306–17311. doi:10.1073/pnas.0407758101

UNCCD (2017). *Best Practices for American Indian and Alaska Native Data Collection*. 2nd Edn Bonn, Germany: UN Convention to Combat Desertifcation.

Verma, P. K., Verma, S., Tripathi, R. D., Pandey, N., and Chakrabarty, D. (2021). CC-type Glutaredoxin, OsGrx_C7 Plays a Crucial Role in Enhancing protection against Salt Stress in rice. *J. Biotechnol.* 329, 192–203. doi:10.1016/J.JBIOTEC.2021.02.008

Voorrips, R. E. (2002). MapChart: Software for the Graphical Presentation of Linkage Maps and QTLs. *J. Hered.* 93, 77–78. Available at: https://academic.oup.com/jhered/article-abstract/93/1/77/2187477 (Accessed October 22, 2021). doi:10.1093/jhered/93.1.77

Wang, L., Li, Z., Qian, W., Guo, W., Gao, X., Huang, L., et al. (2011). The Arabidopsis Purple Acid Phosphatase AtPAP10 Is Predominantly Associated with the Root Surface and Plays an Important Role in Plant Tolerance to Phosphate Limitation. *Plant Physiol.* 157, 1283–1299. doi:10.1104/PP.111.183723

Wang, L., Qin, L., Liu, W., Zhang, D., and Wang, Y. (2014). A Novel Ethylene-Responsive Factor fromTamarix Hispida, ThERF1, Is a GCC-Box- and DRE-Motif Binding Protein that Negatively Modulates Abiotic Stress Tolerance inArabidopsis. *Physiol. Plantarum* 152, 84–97. doi:10.1111/ppl.12159

Wang, R. R. C., Li, X. M., Hu, Z. M., Zhang, J. Y., Larson, S. R., Zhang, X. Y., et al. (2003). Development of Salinity-Tolerant Wheat Recombinant Lines from a Wheat Disomic Addition Line Carrying aThinopyrum junceumChromosome. *Int. J. Plant Sci.* 164, 25–33. doi:10.1086/344556

Wang, X.-F., Su, J., Yang, N., Zhang, H., Cao, X.-Y., and Kang, J.-F. (2017). Functional Characterization of Selected Universal Stress Protein from Salvia Miltiorrhiza (SmUSP) in *Escherichia coli*. *Genes* 8, 224. doi:10.3390/genes8090224

Wang, Y. M., Yang, Q., Liu, Y. J., and Yang, H. L. (2016). Molecular Evolution and Expression Divergence of the Aconitase (ACO) Gene Family in Land Plants. *Front. Plant Sci.* 7, 1–12. doi:10.3389/fpls.2016.01879

Wild, R., Gerasimaite, R., Jung, J.-Y., Truffault, V., Pavlovic, I., Schmidt, A., et al. (2016). Control of Eukaryotic Phosphate Homeostasis by Inositol Polyphosphate Sensor Domains. *Science* 352, 986–990. doi:10.1126/SCIENCE.AAD9858

Wu, Y., and Cosgrove, D. J. (2000). Adaptation of Roots to Low Water Potentials by Changes in Cell wall Extensibility and Cell wall Proteins. *J. Exp. Bot.* 51, 1543–1553. doi:10.1093/jexbot/51.350.1543

Xiong, J., Liu, L., Ma, X., Li, F., Tang, C., Li, Z., et al. (2020). Characterization of Ptaos1 Promoter and Three Novel Interacting Proteins Responding to Drought in Poncirus Trifoliata. *Ijms* 21, 4705–4715. doi:10.3390/ijms21134705

Xing, Q., Liao, J., Cao, S., Li, M., Lv, T., and Qi, H. (2020). CmLOX10 Positively Regulates Drought Tolerance Through Aasmonic Acid -mediated Stomatal Closure in Oriental Melon (Cucumis melo var. makuwa Makino). *Sci. Rep.* 10, 1–14. doi:10.1038/s41598-020-74550-7

Yaduvanshi, N. P. S., Setter, T. L., Sharma, S. K., Singh, K. N., and Kulshreshtha, N. (2012). Influence of Waterlogging on Yield of Wheat (*Triticum aestivum*), Redox Potentials, and Concentrations of Microelements in Different Soils in India and Australia. *Soil Res.* 50, 489–499. doi:10.1071/SR11266

Yu, D., Zhang, L., Zhao, K., Niu, R., Zhai, H., and Zhang, J. (2017). VaERD15, a Transcription Factor Gene Associated with Cold-Tolerance in Chinese Wild Vitis Amurensis. *Front. Plant Sci.* 8, 297. doi:10.3389/FPLS.2017.00297

Yu, S., Ligang, C., Liping, Z., and Diqiu, Y. (2010). Overexpression of OsWRKY72 Gene Interferes in the Abscisic Acid Signal and Auxin Transport Pathway of Arabidopsis. *J. Biosci.* 35, 459–471. doi:10.1007/s12038-010-0051-1

Yuenyong, W., Chinpongpanich, A., Comai, L., and Chadchawan, S. (2018). Downstream Components of the Calmudulin Signalling Pathway in the Rice Salt Stress Response Revealed by Transcriptome Profiling and Target Identification. *BMC Plant Biol.* 18, 1–23.

Zhang, H., Ma, F., Wang, X., Liu, S., Saeed, U. H., Hou, X., et al. (2020). CmLOX10 Positively Regulates Drought Tolerance Through Aasmonic Acid -mediated Stomatal Closure in Oriental Melon (Cucumis melo var. makuwa Makino). *Front. Plant Sci.* 11. doi:10.3389/fpls.2020.00014

Zhou, M. L., Ma, J. T., Pang, J. F., Zhang, Z. L., Tang, Y. X., Wu, Y. M., et al. (2010). Regulation of Plant Stress Response by Dehydration Responsive Element Binding (DREB) Transcription Factors. *African J. Biotechnol.* 9, 9255–9269. doi:10.4314/ajb.v9i54

Zhu, J.-K., Bressan, R., and Hasegawa, P. (1993). Loss of Arabinogalactan-Proteins from the Plasma Membrane of NaCl-Adapted Tobacco Cells. *Planta* 190, 221–226. *1993 1902* 190. doi:10.1007/BF00196614

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.

Copyright © 2022 Prasad, Mittal, Kumar, Chauhan, Sahu, Kumar, Singh, Yadav and Singh. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genetics

REVIEW
published: 16 February 2022
doi: 10.3389/fgene.2022.818727

An Epigenetic Alphabet of Crop Adaptation to Climate Change

Francesco Guarino[1], Angela Cicatelli[1], Stefano Castiglione[1], Dolores R. Agius[2], Gul Ebru Orhun[3], Sotirios Fragkostefanakis[4], Julie Leclercq[5,6], Judit Dobránszki[7], Eirini Kaiserli[8], Michal Lieberman-Lazarovich[9], Merike Sõmera[10], Cecilia Sarmiento[10], Cristina Vettori[11], Donatella Paffetti[12], Anna M. G. Poma[13], Panagiotis N. Moschou[14,15,16], Mateo Gašparović[17], Sanaz Yousefi[18], Chiara Vergata[19], Margot M. J. Berger[20], Philippe Gallusci[20], Dragana Miladinović[21] and Federico Martinelli[19]**

[1]Dipartimento di Chimica e Biologia "A. Zambelli", Università Degli Studi di Salerno, Salerno, Italy, [2]Centre of Molecular Medicine and Biobanking, University of Malta, Msida, Malta, [3]Bayramic Vocational College, Canakkale Onsekiz Mart University, Canakkale, Turkey, [4]Department of Molecular and Cell Biology of Plants, Goethe University, Frankfurt, Germany, [5]CIRAD, UMR AGAP, Montpellier, France, [6]AGAP, Univ Montpellier, CIRAD, INRA, Institut Agro, Montpellier, France, [7]Centre for Agricultural Genomics and Biotechnology, FAFSEM, University of Debrecen, Debrecen, Hungary, [8]Institute of Molecular, Cell and Systems Biology, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, United Kingdom, [9]Plant Sciences Institute, Agricultural Research Organization Volcani Center, Rishon LeZion, Israel, [10]Department of Chemistry and Biotechnology, Tallinn University of Technology, Tallinn, Estonia, [11]Institute of Biosciences and Bioresources (IBBR), National Research Council (CNR), Sesto Fiorentino, Italy, [12]Department of Agriculture, Food, Environment and Forestry (DAGRI), University of Florence, Florence, Italy, [13]Department of Clinical Medicine, Public Health, Life and Environmental Sciences, University of L'Aquila, Aquila, Italy, [14]Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology—Hellas, Heraklion, Greece, [15]Department of Biology, University of Crete, Heraklion, Greece, [16]Department of Plant Biology, Uppsala BioCenter, Swedish University of Agricultural Sciences and Linnean Center for Plant Biology, Uppsala, Sweden, [17]Chair of Photogrammetry and Remote Sensing, Faculty of Geodesy, University of Zagreb, Zagreb, Croatia, [18]Department of Horticultural Science, Bu-Ali Sina University, Hamedan, Iran, [19]Department of Biology, University of Florence, Sesto Fiorentino, Italy, [20]UMR Ecophysiologie et Génomique Fonctionnelle de la Vigne, Université de Bordeaux, INRAE, Bordeaux Science Agro, Bordeaux, France, [21]Institute of Field and Vegetable Crops, National Institute of Republic of Serbia, Novi Sad, Serbia*

Crop adaptation to climate change is in a part attributed to epigenetic mechanisms which are related to response to abiotic and biotic stresses. Although recent studies increased our knowledge on the nature of these mechanisms, epigenetics remains under-investigated and still poorly understood in many, especially non-model, plants, Epigenetic modifications are traditionally divided into two main groups, DNA methylation and histone modifications that lead to chromatin remodeling and the regulation of genome functioning. In this review, we outline the most recent and interesting findings on crop epigenetic responses to the environmental cues that are most relevant to climate change. In addition, we discuss a speculative point of view, in which we try to decipher the "epigenetic alphabet" that underlies crop adaptation mechanisms to climate change. The understanding of these mechanisms will pave the way to new strategies to design and implement the next generation of cultivars with a broad range of tolerance/resistance to stresses as well as balanced agronomic traits, with a limited loss of (epi)genetic variability.

Keywords: abiotic stresses, adaptation, climate change, epigenetics, environmental stresses, epigenetic code

OPEN ACCESS

Edited by:
Vijay Gahlaut, Institute of Himalayan Bioresource Technology (CSIR), India

Reviewed by:
Rajeev Ranjan, Purdue University, United States
Namisha Sharma, National Institute of Plant Genome Research (NIPGR), India

****Correspondence:***
Dragana Miladinović
draganavas@yahoo.com
Federico Martinelli
federico.martinelli@unifi.it

Specialty section:
This article was submitted to Plant Genomics, a section of the journal Frontiers in Genetics

Received: *19 November 2021*
Accepted: *28 January 2022*
Published: *16 February 2022*

Citation:
Guarino F, Cicatelli A, Castiglione S, Agius DR, Orhun GE, Fragkostefanakis S, Leclercq J, Dobránszki J, Kaiserli E, Lieberman-Lazarovich M, Sõmera M, Sarmiento C, Vettori C, Paffetti D, Poma AMG, Moschou PN, Gašparović M, Yousefi S, Vergata C, Berger MMJ, Gallusci P, Miladinović D and Martinelli F (2022) An Epigenetic Alphabet of Crop Adaptation to Climate Change. Front. Genet. 13:818727. doi: 10.3389/fgene.2022.818727

1 EPIGENETICS – BEYOND THE CLASSIC GENETIC ALPHABET

The term "epigenetics" derives from "epigenesis," coined by the physician and physiologist William Harvey at around 1650, for the conception of development as a gradual process of increasing complexity from initially homogeneous material present in the egg of different animals. This idea was originally proposed by Aristotle (Van Speybroeck et al., 2002). However, this concept deeply changed over time, and in 1942 the embryologist Conrad Waddington introduced the term "epigenetics" into modern biology defining it as "the whole complex of developmental processes" that lies between "genotype and phenotype" (Waddington, 1942).

In recent years, our understanding of the role of epigenetic mediated responses to environmental stimuli, especially to stresses, has greatly improved (Mladenov et al., 2021). Environmental stress factors, due to climate change, affect plant growth and pose a growing threat to sustainable agriculture and food security (Altieri et al., 2015). These factors include intense drought periods, excessive rainfalls eventually causing flooding, extreme temperatures, and heat waves, among others (Schiermeier, 2018). Although the acute responses of crops to single stresses are considered individually and in single occurrence are extensively studied, stresses typically occur in a chronic or recurring way and mostly in a combined manner. Recent studies suggest that plants have "a stress memory" that is guiding, or supervising in a way, their adaptation to chronic, recurring, and combined environmental stresses (Walter et al., 2013). In general, irrespective of whether environmental stimuli are chronic or not (such as drought, hyperosmotic, salinity, heat, pathogens, etc.), they can induce diverse epigenetic mechanisms, where key genes, such as Dicer-like 4 (DCL4) and Retrotransposon-like 1 (RTL1), play an important role.

Epigenetic mechanisms involved in plant responses to environmental stresses are not encoded by the classical four-letter genetic alphabet (Faltýnek et al., 2020). Hence, epigenetic modifications are usually chemically expressed by expanding the standard four-letter genetic alphabet by the addition of a special mark to a letter (nucleobases), thus creating a specific "epigenetic alphabet" (**Figure 1**).

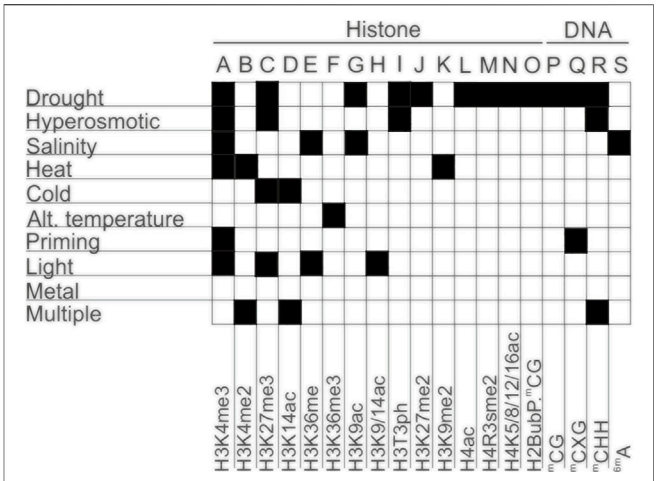


FIGURE 1 | Deciphering the alphabet of epigenetic responses to the environmental stresses in plants. Different types of epigenetic modifications in response to different abiotic and biotic stresses. A-O—Histone modifications; P-R—Cytosine methylation; S—Adenine methylation.

1.1 From A to S—"Basic" Epigenetic Alphabet

1.1.1 A-O: Histone Variants and Histone Post-Transcriptional Modifications

Chromatin structuring and remodeling, which are key regulatory processes for controlling the accessibility of genes to the transcriptional machinery, play an important role in plant responses to climate change (Song et al., 2021) (**Figure 2**).

The basic functional unit of the chromatin is the nucleosome, which consists of a histone octamer made of two copies of each of the histone H2A, H2B, H3, and H4 wound by 147 bp of DNA. The histone H1 binds to the "linker DNA" comprising 20–80 nucleotides that separate two nucleosomes (Annunziato, 2008). Canonical histones, except for H4, have minor variants which can be incorporated into the nucleosome throughout the cell cycle. Canonical histones and their variants differ only by a few amino acid residues, but their exchange in the nucleosome can modulate the exposure of DNA and regulation of transcription by directly influencing the chromatin structure. For example, the H2A.Z variant located in gene bodies ensures the repression of heat and osmotic stress-related genes in the absence of stress, while eviction of H2A.Z allows their transcriptional induction upon stress (Cortijo et al., 2017; Sura et al., 2017). Mutants of the SWR1-like chromatin remodeling complex which are impaired in H2A.Z installment show enhanced resistance to pathogens, highlighting H2A.Z importance for adaptive response to both abiotic and biotic stresses (March-Díaz et al., 2008). Another example is provided by the stress-inducible H1 variant H1.3 in *Arabidopsis thaliana*, which modulates stomata under non-stress or light and water-limited conditions (Rutowicz et al., 2015).

In addition, post-translational modifications of histones may lead to changes of chromatin structure and packaging and modify the accessibility of cis-regulatory elements to transcription factors and associated protein complexes (Zhang X. et al., 2020). Among the 26 histone post-translational modifications (HPTMs) described in the literature (Zhao et al., 2015), two have been intensively studied in the context of the response to stress, namely acetylation and methylation, while recent work suggest that ubiquitination and phosphorylation are also involved in this process. As all other HPTMs, these marks are established by histone writers complexes such as histone acetyltransferases (HAT), methyltransferases (HMT), kinases, and ubiquitinases, and removed by "erasers" including deacetylases (HDA), demethylases (HDM), phosphatases, and de-ubiquitinases (Xu et al., 2017; Maeji and Nishimura, 2018). Acetylation which occurs on lysine residues (K) on histones H3 and H4 respectively at positions 9, 14, 18, 23, and 27, and positions 5, 8, 12, 16, and 20, neutralizes the positive charge of histones

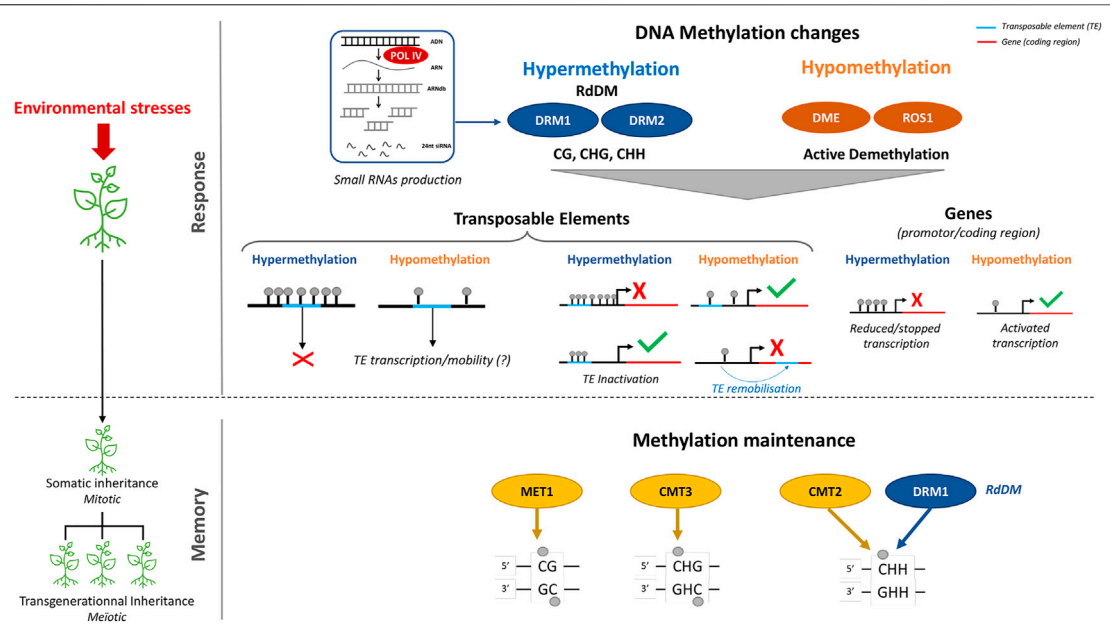


FIGURE 2 | Roles of DNA-methylation in environmental stress responses and memories in plants. Changes in DNA-methylation landscape are part of the response of plants to environmental stresses. De novo methylation, which is targeted at specific loci by small-RNAs, is established by the RNA-dependent-DNA-Methylation pathway (RdDM) whereas, DNA demethylation at specific loci requires functional DNA Glycosylase Lyase also called DNA demethylase such as Repressor of Silencing 1 (ROS1). Modification of DNA methylation patterns at genes may result in changes in gene expression level leading to gene induction or repression. In addition, stress induced DNA methylation variations may occur at transposable elements (TEs) and determine their inactive or active state. When hypomethylated and transcriptionally active, TEs may indirectly influence the expression of genes located in their vicinity, whereas their hypermethylation has the reverse effect. Additionally, the mobility of TEs may generate new regulatory patterns or mutations leading to loss of gene function when their insertion occurs in genes. Maintenance of stress induced patterns of DNA methylation through cell division (Mitosis or meiosis), results in an epigenetic memory. This memory requires the context-specific DNA-methyltransferases METHYLTRANSFERASE-1 (MET1), CHROMOMETHYLASE-3 (CMT3) for CG and CHG sequence context, respectively. Methylation in the CHH sequence context is maintained by CMT2 or by the RdDM pathway in heterochromatic and euchromatic regions, respectively.

thereby weakening the interaction between histones and DNA. In contrast, deacetylation has the opposite effect and results in chromatin condensation (Shahbazian and Grunstein, 2007; Prakash and Fournier, 2018). Consistently, histone acetylation has been associated with active gene expression (reviewed in Hu Y. et al., 2019). Several studies have demonstrated that the abundance and/or distribution of acetylated histones change in plants facing abiotic stresses or pathogen attacks (reviewed in Hu Y. et al., 2019; Lu et al., 2018; Park et al., 2018). Furthermore mutants affected in either of these enzymatic activities present altered responses to abiotic stresses (reviewed in Hu Y. et al., 2019). Among the HAT, several studies have shown that the GCN5 protein plays a central role in coordinating the response to heat and salt stress in *A. thaliana* (Hu et al., 2015; Zheng et al., 2018). Inversely, deacetylation which leads transcriptional repression affects the transcriptome landscape under abiotic stress conditions (Park et al., 2018).

The mono-, di- or tri-methylation of histone tails, which occurs at arginine (R) or lysine (K), alters the hydrophobicity of histone side chains thereby the interaction with reader proteins and the transcriptional machinery. However, R and K methylation has diverse effects on chromatin organization and gene expression depending on the position of the modified amino acid (Lämke and Bäurle, 2017). Asymmetric H4R3me2 (dimethylation (me2) of the third arginine (R3) of Histone 4 (H4), H3K4me3, H3K36me2/3 are associated with active transcription, while symmetric H4R3me2, H3K9me2/3, and H3K27me3, that exist symmetrically on the two copies of identical histones in the same nucleosome, correlate with transcriptional silencing (Bobadilla and Berr, 2016; Ueda and Seki, 2020). Numerous works that either analyzed the dynamics of histone methylation or the reponse of mutants affected in HMT aor HDMT activities have now shown the importance of histone methylation in the development and responses to stresses (reviewed in Ueda and Seki, 2020). Importantly, data suggest that the removal of repressive methylation marks is necessary for some stresses to unlock the expression of stress related genes (Shen Y. et al., 2014; Huang et al., 2019b).

In contrast to acetylation and methylation, histone ubiquitination and phosphorylation have been only sparsely studied. Monoubiquitination of histone lysine H2Bub is considered an active mark in salt and drought stress response (Chen et al., 2020). However, the mono-ubiquitination of histones is an important HPTM that occurs on histones H2A and H2B at lysine K121 and K143 respectively (March and Farrona, 2018). Whereas H2B mono-ubiquitination (H2B ub) marks active genes in association with methylation at K4 and K36 of histone H3, H2Aub, which is established by the PRC1 (Polycomb Repressive Complex 1) upon recruitment at H3K27me3 marks established by the PRC2, maintain the

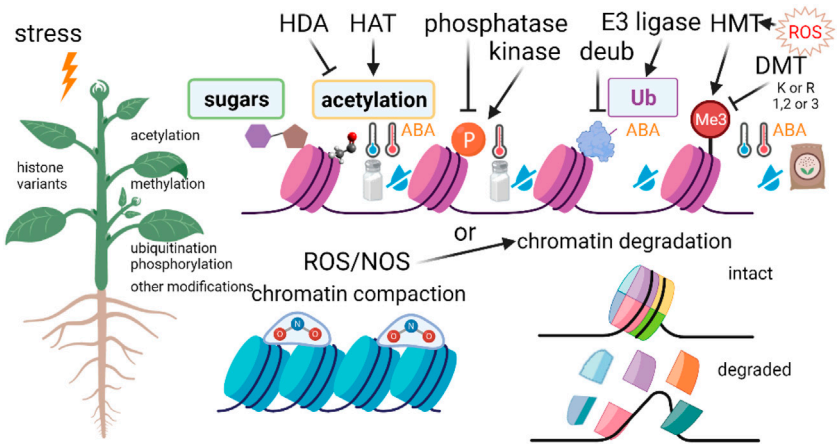


FIGURE 3 | Histone modifications in response to environmental stresses. HAT, histone acetyltransferase; HDA, histone deacetylase; HMT, histone methyltransferase; ROS, reactive oxygen species; ABA, abscisic acid.

chromatin in a closed state and is associated with the repression of gene expression (March and Farrona, 2018). It is only recently that a possible role H2Bub in the response to drought stress was established in cotton (Chen et al., 2019). and in *A. thaliana* and rice, respectively, by regulating cutin biosynthesis (Ménard et al., 2014; Patwari et al., 2019) and ABA signaling (Ma et al., 2019).

Finally, the phosphorylation of histone H3 which can occur on threonine and arginine has been essentially studied in the context of the cell cycle (Houben et al., 2007) and its putative role in stress responses is not well understood so far. However, phosphorylation at H3T3 is increased in pericentromeric regions under drought conditions to repress transcription, acting in an antagonistic manner to H3K4me3 (Wang Z. et al., 2015).

In addition to the response to stresses, HPTMs have also been implicated in stress memory such as histone methylation which can be maintained for a relatively long period in primed plants (Lämke, and Bäurle, 2017).

1.1.2 P-S: DNA Methylation

DNA methylation plays an important role in the regulation of gene expression and plant reaction to environmental stresses (Kumar and Mohapatra, 2021a) (**Figure 3**), In plants, DNA methylation predominantly occurs by the addition of a methyl group to the fifth position of the pyrimidine ring of cytosine bases or the sixth position of the purine ring of adenine bases, which is referred to as 5-methylcytosine [5 mC] or N^6-methyladenine [6 mA] DNA methylation, respectively (Liu and He, 2020).

1.1.2.1 Cytosine Methylation

Cytosine methylation in plants occurs in the two symmetrical q sequences contexts CG, and CHG, and in the non-symmetrical one CHH (where H is A, C, or T). Cytosine methylation is high at heterochromatic regions (centromeres, transposable elements (TE), other repetitive elements) is involved in their transcriptional silencing (Transcriptional Gene Silencing, TGS). In contrast, methylation levels are low in euchromatic regions (Liu and He, 2020).

Mechanisms involved in the regulation of DNA methylation depend on the sequence contexts and therefore occur following three different mechanisms. In the CG context, the DNA methyl transferase (DNMT) MET1 together with additional cofactors including VARIANT IN METHYLATION (VIM), and decrease IN DNA METHYLATION 1 (DDM1) methylate the unmethylated cytosine incorporated during DNA replication in the newly synthesized DNA strand (Law and Jacobsen, 2010), whereas the chromomethylase 3 (CMT3) and to a lower extend CMT2 will fulfill a similar function at CHG motives (Du et al., 2012). Methylation maintenance at CHH asymmetrical motives requires reinstalling methylation at the newly synthesized unmethylated DNA strand by the Domain Rearranged methyltransferase 2 (DRM2) and the RNA-dependent DNA methylation (RdDM) in euchromatic regions, or CMT2 methyltransferase in heterochromatic regions enriched in histone H1 where the RdDM pathway is inhibited (Zemach et al., 2013; Zhang H. et al., 2018). *De novo* methylation that occurs at non-methylated sites in any sequence context is mediated through the RdDM pathway, and requires small interfering RNAs (siRNAs), scaffold RNAs and several additional proteins (Zhang H. et al., 2018).

In addition, active DNA demethylation is gaining significant attention because it is involved in many biological processes in plants, and in the response to various stresses (Liu and Lang, 2020). 5 mC can be either passively removed by simple dilution after DNA replication, or actively eliminated by specific enzymes, namely the 5-methylcytosine glycosylase-lyase that belongs to the DEMETER (DME)/REPRESSOR OF SILENCING (ROS1) family in *A, thaliana* and DEMETER-LIKE (DML). These DMLs are bi-functional enzymes exercising the 5 mC creating an abasic site, likely repaired by unknown DNA polymerases and ligases activities. The whole process results in a net loss of cytosine methylation (Law and Jacobsen, 2010).

Variations in DNA methylation can occur during inbreeding, plant aging, and in the response to different stresses (Quadrana and Colot, 2016; Zhang Q. et al., 2018). These variations may underlay phenotypic variation (Noshay and Springer, 2021) as demonstrated by analyzing Epigenetic Inbred Lines (EpiRILs) in *A, thaliana* (Quadrana and Colot, 2016). In addition, EpiRILs allowed demonstrating that differences in the epigenetic landscape of plants can lead to a significant plastic response to stresses (Kooke et al., 2015), as DNA methylation changes were observed in stressed plants (reviewed in Zhang Q. et al., 2018). In addition, impairing enzymes involved in DNA methylation leads to variable survival in response to stresses, highlighting the fundamental role of DNA methylation in the plant responses to stresses (Yao et al., 2012; Shen X. et al., 2014; Wibowo et al., 2016).

1.1.2.2 Adenine Methylation

The recent discovery that adenine can also be methylated although at very low rates, add another layer of complexity to the epigenetic processes affecting plant genomic DNA (reviewed in Chachar et al., 2021). As for cytosine methylation, 6 mA DNA methyltransferases have been identified, as well as associated demethylases (Chachar et al., 2021). Interestingly, 6 mA associates with active gene expression, which contrasts with the main function associated with cytosine methylation (Zhang Q. et al., 2018; Liang et al., 2018). Recent evidence also suggests that rice plants with various levels of 6 mA present variable responses under different abiotic stresses, suggesting a potential role of 6 mA in the plant responses to environmental stimuli (Zhang Q. et al., 2018).

6 mA is directly involved in heterochromatin regulation in mouse embryonic cells. It also participates in the regulation of mRNA encoding HAT or HMT, and is involved in the recruitment of histone modifiers during transcription thereby affecting the deposition of specific epigenetic marks in animals (reviewed in Kan et al., 2021). However, further studies are needed to elucidate if 6 mA has similar roles in plants.

1.2 From T to Z?—RNA-Mediated Epigenetic Modifications

The regulation of gene expression in response to stresses, both at the transcriptional and post-transcriptional levels, confers plasticity, and adaptability to plants (Wang et al., 2017; Song X. et al., 2019). This regulation is induced by small and long non-coding RNAs (sRNAs 20–24 nt, and lncRNAs > 200 nt), thus adding new letters to the "epigenetic alphabet," beyond ones created by histone modifications and DNA methylation (**Figure 1**).

Non-coding RNA may be involved in the constitutive repression of transposon elements. MicroRNAs (miRNAs) are essential factors able to determine the specificity of post-transcriptional regulations. They originate from the cleavage of endogenous transcripts of miRNA (MIR) genes by DICER-LIKE 1 (DCL1). Loaded into AGO1, miRNAs regulate the gene expression by degradation or translational repression of target mRNAs. Although miRNAs are primarily involved in the PTGS regulation of gene expression, recent evidence suggests that they may also participate in epigenetic pathways, although indirectly. For example, modulation of miRNA populations was suggested to shape the epigenetic memory of stresses by modulating the expression of epigenetic regulators in Norway spruce (Yakovlev and Fossdal, 2017). In addition, miRNA may play direct although minor roles in DNA methylation through the non-canonical RdDM pathway (Cuerda-Gil and Slotkin, 2016), including in response to environmental stimuli (Iwasaki et al., 2019).

Small interfering RNAs (siRNAs) arise from the processing of intermediate double-stranded RNAs synthesized by RNA-dependent RNA polymerases (RDRs) (Song X. et al., 2019). Among them, TE-derived siRNAs are produced upon transcription and/or transpositional reactivation of TEs in response to stress (Hou et al., 2019). The plant-specific RNA polymerase IV generates single-stranded siRNA precursors, converted into double-stranded RNAs (dsRNAs) by RDR2. These dsRNAs are processed by DCL3 for producing 24-nt mature siRNA and loaded preferentially into AGO4 (Law and Jacobsen, 2010). At the same loci, another plant-specific RNA polymerase V generates non-coding transcripts allowing the recruitment of the siRNA-AGO4 complex through sequence complementarity, as well as DRM2 (Du et al., 2015). Consequently, *de novo* DNA methylation occurs at different loci (Erdmann and Picard, 2020) in all cytosine sequence contexts (Law and Jacobsen, 2010). Distinct Dicer-type nucleases are involved in miRNAs/siRNAs production. They are subsequently recruited by distinct proteins of the AGO family (Iki, 2017), which act together within the miRNA-induced silencing complex to target complementary sequences of coding and non-coding RNAs (Song X. et al., 2019).

LncRNAs regulate gene expression at the epigenetic, transcriptional, post-transcriptional, translational, and post-translational levels (Sun et al., 2018; Zhang et al., 2019; Wu et al., 2020). LncRNAs are transcribed by RNA polymerase II, III, IV, and V, and have specific spatial structures and spatiotemporal expression patterns. They are divided into five categories according to their position in the genome, next to or far from protein-coding genes: sense, antisense, bidirectional, intronic (incRNA), and large intergenic lncRNA. Many plant lncRNAs are differentially expressed by abiotic and biotic stresses (Wang et al., 2018; Yin et al., 2019) and were suggested to play an important role in this context (Urquiaga et al., 2021).

Recent technical advances have revealed widespread and sparse modification of mRNAs, providing an additional layer of complexity to the regulation of gene expression. Prevalent mRNA modifications, namely the N^6-methyladenosine (m6A) and 5-methylcytidine (m5C), are modulated by specific writers (RNA methyltransferase, e.g., AlkB), readers, and erasers (RNA demethylase). The writer complex, also known as "methylosome," adds m6A at conserved sites and comprises a catalytic heterodimer METTL3/METTL14; MTA in *A. thaliana*, associated with the regulatory proteins FIP37 (FKBP12 INTERACTING PROTEIN 37)) and VIRILIZER. The corresponding mutants are embryo lethal (Zhong et al., 2008; Shen et al., 2016; Růžička et al., 2017). Furthermore, m6A

stabilizes transcripts required for salt and osmotic stress response (Anderson et al., 2018), suggesting roles for m6A beyond development. Polymethylated mRNAs (i.e. carrying many m6A modifications) facilitate inter/intramolecular interactions, a property referred to as "multivalency." Multivalency enables m6A mRNAs to participate in assemblies comprising proteins, RNAs, and metabolites called "biomolecular condensates" due to their capacity to concentrate molecules. Condensate formation may rely on liquid-liquid phase separation (LLPS), whereby a solution de-mixes into two or more distinct phases (Huang et al., 2021). Proteins with intrinsically disordered regions (IDR) in many cases can promote LLPS. For example, the IDR-enriched YTH domain proteins EVOLUTIONARILY CONSERVED C-TERMINAL REGION2/3/4 (ECT2/3/4), which can read m6A, modulate leaf development and localize in the biomolecular condensates known as stress granules (SGs) (Kosmacz et al., 2019). Yet, there is a lack of understanding of ECT functions in the SGs. SGs form rapidly upon stress onset to readjust the transcriptome by degrading or storing mRNAs and thus optimizing adaptation (Gutierrez-Beltran et al., 2015). As has been shown for animals, ECTs may regulate SGs formation during stress (Fu and Zhuang, 2020), thereby adjusting the transcriptome landscape indirectly by recruiting m6A-modified RNA molecules in SGs, whereby they are kept inert.

Several studies suggest that the m6A writers AlkB homologs (ALKHB) regulate stress responses due to their gene expression levels modulation upon stress (Hu H. et al., 2019). *A. thaliana* has 13 ALKHB proteins, and ALKBH9B demethylates m6A and affects viral spread (Martínez-Pérez et al., 2017), while ALKBH10B influences flowering by controlling SQUAMOSA PROMOTER-BINDING PROTEIN-LIKE (SPL) 3, SPL9, and FLOWERING LOCUS T mRNA levels (Duan et al., 2017). This link is indirect and merits further investigation.

Apart from m6A, we know little about other mRNA modifications in plants. Recent evidence suggests a link between m5C and RNA mobility. Mobile mRNAs are rich in m5C (Yang et al., 2019). Yet, the molecular machinery involved in recognizing and distributing m5C mRNA is still unknown. Furthermore, there is evidence, mainly from animal systems, that mRNA modification also plays a direct role in epigenetics (Kumar and Mohapatra, 2021b; Kan et al., 2021). Whether similar roles for epi-modification of plant mRNAs exist in plants requires further investigation.

2 EPIGENETIC ALPHABET—(DE)CODING THE STRESS RESPONSE

Climate change is altering the environments in which all organisms develop and thrive. Plant species, as sessile organisms, can adjust to these novel conditions through phenotypic plasticity, natural selection and eventually can change habitat to follow their optimal growing conditions, these possibilities being not mutually exclusive. Epigenetic modifications that occur in plants are also part of their response to changes in their environment. Those epigenetic changes are adding to natural mutations, with epigenetic marks creating an enlarged version of the genetic alphabet, thus increasing the variety of phenotypes within the stress-affected habitat (Faltýnek et al., 2020). When heritable to the progeny, they become a certain kind of "norm," enabling us to further decode stress response in crop of interest and apply it for resilience improvement.

In the context of stress responses, these histone-modifying complexes are directed by stress-induced transcription factors to their appropriate targets. For example, the COMPASS H3K4 methyltransferase complex is recruited by bZIP transcription factors and brings about methylation of H3K4 (Song et al., 2015). Furthermore, among the numerous messengers, such as calcium, redox signaling, membrane integrity, G-proteins, mitogen-activated protein kinases (MAPKs), plant stress hormones (salicylic, jasmonic and abscisic acid, ethylene) that modulate the response of plants to stresses, the Reactive Oxygen Species (ROS) and the Reactive Nitrogen Species (RNS) have received increasing attention over the last decade as they are key players of the integrated responses of plants to these stresses, in addition to their fundamental functions in plant development (Sewelam et al., 2016; Huang et al., 2019a; Kumar et al., 2020). Indeed, the plant responses to different abiotic stresses, such as heat, chilling, excessive light, drought, ozone exposure, UV-B irradiation, osmotic shock, heavy metals, and organic pollutants involves a rapid oxidative burst that leads to the generation and/or accumulation of oxidants such as ROS and RNS. These reactive species are essential signaling systems that participate to multiple processes, necessary to adjust the metabolism or physiology either at the whole plant or tissue level or in a specific subcellular compartment (Waszczak et al., 2018).

Redox intermediates play also a critical role in the regulation of epigenetic mechanisms in response to plant stresses. They govern DNA methylation levels: increases in ROS caused DNA hypomethylation both in tobacco (Choi and Sano, 2007; Poborilova et al., 2015) and pea (Berglund et al., 2017). Similarly, in rice, RNS caused a heritable hypomethylation (Ou et al., 2015). In addition, redox intermediates often regulate enzymes involved in histone methylation and acetylation (Ojima et al., 2012). In maize, ROS, generated by heat stress, induced histone hyperacetylation (Wang P. et al., 2015). Heritable changes induced by the environment have been shown in *Linum usitatissimum* L. (Cullis, 1986), in *Mesembryanthemum crystallinum* L. (Bohnert et al., 1995), and *Brassica nigra* L. (Waters and Schaal, 1996). Since the '90s hypermethylation of heterochromatic loci has been reported in tobacco, either in response to osmotic stress (Kovařik et al., 1997) or in silenced genes in transgenic plants (Meyer et al., 1992; Meyer and Heidmann, 1994). On the contrary, hypomethylation has been documented in chicory root tips (Demeulemeester et al., 1999) and *A. thaliana* (Finnegan et al., 1998) when exposed to low temperature. Epigenetic changes have been observed in tissue cultures, while methylation polymorphisms have been frequently observed during the propagation of tissue cultures at the level of repeated sequence (Smulders et al., 1995) and may contribute to somaclonal variation (Kaeppler et al., 2000). All these variations in genome methylation might be part of the plant's adaptation

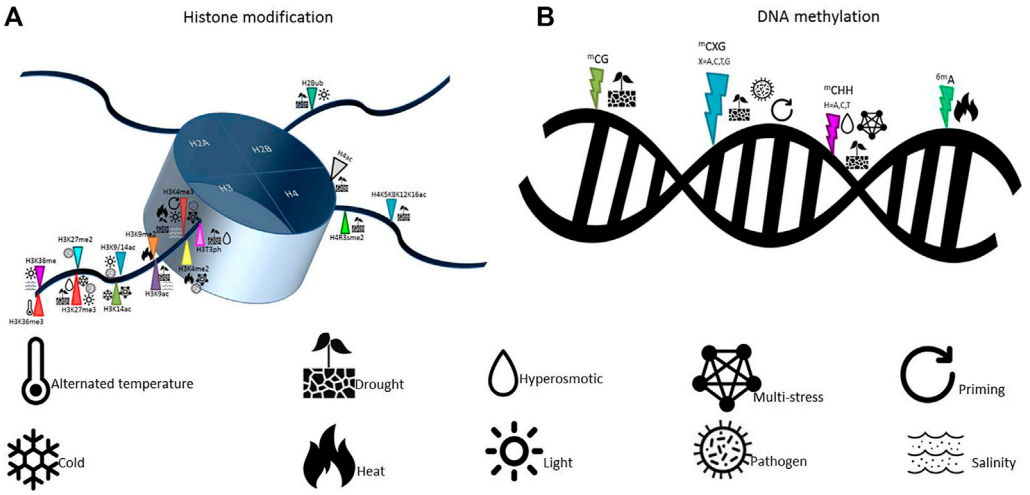


FIGURE 4 | Epigenetic mechanisms involved in plant response to stress. Histone modifications **(A)** include acetylation/deacetyaltion and methylation/demethylation, while DNA methylation **(B)** includes cytosine methylation and adenine methylation processes.

mechanisms to abiotic stresses (Martienssen and Richards, 1995; Kovarík et al., 1997).

Furthermore, under abiotic constraints, plants show multiple alterations in their sRNAome, leading to changes in the accumulation of individual sRNAs or through their specific induction in stress conditions, as shown in annual plants (Liu et al., 2017; Pan et al., 2017; He et al., 2019) and cultivated perennials plants (Leclercq et al., 2020). The respective proportions of the different sRNA classes may also be modified in response to stress, adjusting a genome-wide gene expression reprogramming to different sRNA-dependent regulation mechanisms. Few examples are known in plants responding to developmental or abiotic cues. A switch between 24 and 21 nt sRNA has been observed in rubber trees upon the occurrence of the stress-induced Tapping Panel Dryness syndrome (Gébelin et al., 2013), as well as in apple trees during the vegetative-to-floral transition with changes within 24 nt sRNA population (Guo et al., 2017). In cereals, different cultivars showed altered abundance in miRNAs contents which was associated with differences in stress sensitivity and in the modulation of a wide set of genes referable to drought tolerance (Bakhshi et al., 2017; Fard et al., 2017). A summary of the most recent epigenetic modifications, in response to different types of stresses, as well as the alphabet of epigenetic responses to the environmental stresses in plants are presented in **Figures 1**, **4**; **Table 1**, respectively.

2.1 Drought Stress

Water availability is one of the most important and prevalent environmental cues which affect the growth, development, and productivity of plants, and ultimately, their survival. Many known epigenetic regulations were shown to play a significant role in acclimation and adaptation to drought stress (Varotto et al., 2020) Concerning climatic change, one of the most important abiotic stresses, water scarcity, is becoming increasingly critical for the survival of plants and crop productivity and yield. Moreover, in long-lived perennial tree species, the water balance in the organ tissues is crucial for growth, survival, and reproductive capability, and influences their distribution along with the environment and climate gradient (Jenkins et al., 2018). Therefore, understanding how plants respond to water stress/scarcity would allow us to inform breeders to select new varieties more resilient to this kind of stress. Hence, many studies on epigenetic changes associated with water stress/scarcity or drought have occurred in recent years both in crops and woody species (Ashapkin et al., 2020; Varotto et al., 2020; Kapazoglou et al., 2021).

MYB96 transcription factor was identified in *A. thaliana* (Lee and Seo, 2019), and AREB1 in poplar as a HAT recruiter in response to drought stress, which affects the acetylation of the H3K9 and thereby the expression of NAC genes (Li et al., 2018). HDA6 and HDA15 regulate genes participating in the jasmonate signaling network and production of Rho of plants (RHO GTPases) by deacetylation of H3K914ac and H4K5K8K12K16ac, respectively, (Jiang et al., 2020). HDA9 is one of the major histone deacetylases, which regulates the expression of drought-responsive genes in *A. thaliana* (Lee and Seo, 2019). The increased expression of BdHD1 in *Brachypodium* caused lower acetylation of H3K9 affecting 230 genes and leading to an abscisic acid hypersensitive phenotype (Song J. et al., 2019). H3K4 methylation is widespread histone methylation in response to drought stress. In *A. thaliana*, the lower levels of H3K4me3 or H4R3sme2 cause increased drought stress tolerance (Liu et al., 2018), while H3T3ph, the phosphorylation of H3 at the threonine of the pericentromeric part, causes an increase in osmotic tolerance (Wang Z. et al., 2015). Monoubiquitination of H2A and H2B is also related to drought tolerance. In fact, H2Bub acts on changing abscisic acid signaling and wax biosynthesis and thereby enhancing drought tolerance not only in *A. thaliana* but in cotton and rice, as well (Chen et al., 2019).

TABLE 1 | Most recent examples of epigenetic modifications in plants in response to different types of abiotic stresses. Duration of epigenetic state, type of epigenetic modification, key proteins involved (M: mediators; T: their targets).

Species	Stress condition	Epigenetic state duration	Epigenetic and post-transcriptional modifications			Key proteins involved	References
			Histone modifications	DNA methylation	Non-coding RNA		
Drought/Osmotic							
Arabidopsis thaliana	Four cycles of 2 h dehydration/22 h rehydration	Up to 5 days	H3K4me3	—	—	T: RD29B, RAB18	Ding et al. (2012)
Arabidopsis thaliana	One or two cycles of 2 h dehydration/22 h rehydration	5 days	H3K4me3, H3K27me3	—	—	M: CLF	Liu et al. (2014)
Arabidopsis thaliana	1–8 h dehydration	Up to 5 h after rehydration from a 4 h stress	H3K9Ac	—	—	T: RD20, RD29A	Kim et al. (2012)
Oryza sativa	7–9 days without watering	ND[a]	—	—	mir162b	T: TRE1	Tian et al. (2015)
Oryza sativa	7–9 days without watering in seedlings or during panicle development	ND	—	—	miR164	T: NACs	Fang et al. (2014)
Arabidopsis thaliana	–0.5 MPa PEG for 10 days	ND	—	—	miR393	T: TIR1, AFB2	Chen et al. (2012)
Arabidopsis thaliana	12 days without watering	12 days	H4R3sme2	—	—	M: CAU1/PRMT5/SKB1 T: CAS	Fu et al. (2013)
Arabidopsis thaliana	30% PEG for 7 days	7 days	H3T3ph	—	—	M: MLK1/2 T: pericentromeric regions	Wang et al. (2015a)
Arabidopsis thaliana	14 days without watering	14 days	H4Ac	—	—	T: PDC, ALDH2B7	Kim et al. (2017)
Arabidopsis thaliana	7 days without watering	14 days	H3K4me3	—	—	M: ATX4, ATX5 T: AHG3	Liu et al. (2018)
Gossypium hirsutum	7 days in MS medium with up to 40% PEG	12 days	H2Bub H3K4me3	—	—	M: AtHUB2 T: GhDREB	Chen et al. (2019)
Arabidopsis thaliana	14 days without watering	14 days	H3Kac, H4ac	—	—	M: HDA15, MYB96 T: RHO gtpase	Lee and Seo, (2019)
Populus trichocarpa	7 days without watering	7 days	H3K9ac	—	—	M: AREB1-ADA2b-GCN5 T: PtrNAC006, PtrNAC007, PtrNAC120	Li et al., 2019
Oryza sativa	5–7 days without watering	5–7 days	H2Bub1	—	—	M: OsHUB2, OsOTLD1 T: MODD	Ma et al. (2019)
Zea mays	Water content threshold of 25% of the available water for 10 days	7 days	H3K4me3, H3K9ac	—	—	T: ZEP1, NCAD6, AP2/EREBP, NAC, MADS4, MADS15	Forestan et al. (2020)
Glycine max	13 days without watering	—	—	—	miR169g	T: NFY	Ni et al. (2013)
Oryza sativa	7 days without watering	—	—	—	miR162b	T: TRE1	Tian et al. (2015)
Arabidopsis thaliana	11–15 days without watering	—	—	—	miR168a	M: AGO1	Li et al. (2012)
Arabidopsis thaliana	14 days without watering	—	—	—	miR396a, miR396b	T: GRF	Liu et al. (2009)
Hyperosmotic							
Arabidopsis thaliana	Priming with 50 mM NaCl, 10 days recovery, 14 days in 80 mM NaCl	10 days	H3K27me3	—	—	T: SOS5, LRP1, SCARECROW	Sani et al. (2013)
Medicago truncatula	204 mM NaCl for 1 week	ND	H3K4me2	mCHH			Yaish et al. (2018)
Salt							
Arabidopsis thaliana	Priming in 100 mM NaCl for 24 h, recovery for 48 h, 200 mM NaCl	5 days	H3K4me3	—	—	M: HY5 T: P5CS1	Feng et al. (2016)
Arabidopsis thaliana	200 mM NaCl for 6 h	ND	H4R3sme2	—	—	M: SKB1 T: LSM4	Zhang et al. (2011)
Medicago sativa	200 mM NaCl up to 24 h	ND	H3K9Ac	—	—	T: MsMYB4	Dong et al. (2020)

(Continued on following page)

TABLE 1 | (*Continued*) Most recent examples of epigenetic modifications in plants in response to different types of abiotic stresses. Duration of epigenetic state, type of epigenetic modification, key proteins involved (M: mediators; T: their targets).

Species	Stress condition	Epigenetic state duration	Epigenetic and post-transcriptional modifications			Key proteins involved	References
			Histone modifications	DNA methylation	Non-coding RNA		
Glycine max	200 mM NaCl, 4 h	ND	HDAC	—	miR482bd-5	HEC1	Cadavid et al. (2020)
Arabidopsis thaliana Heat	100–150 mM NaCl for 9 days	—	—	—	miR393	T: TIR1, AFB2	Chen et al. (2015)
Arabidopsis thaliana	Acclimation: 37°C for 1 h, 23°C for 90 min, and 44°C for 45 min	3 days	H3K4me2 H3K4me3	—	—	M: HsfA2 T: HSPs	Lämke et al. (2016)
Arabidopsis thaliana	Acclimation: 37°C for 1 h, 23°C for 90 min, and 44°C for 45 min	3 days	Histone occupancy	—	—	M: FGT1 T: HSPs	Brzezinka et al. (2016)
Brassica rapa	42°C for 3 h per day for 7 days	Transgenerational	—	—	miR168	AGO1	Bilichak et al. (2015)
Arabidopsis thaliana	Basal: 44°C for 50 min Acquired: 37°C for 1 h, 2 h 22°C, 44°C for 3.5 h	ND —	—	—	TAS1 (tasiRNA)	HTT1/2	Li Shuxia et al. (2014)
Arabidopsis thaliana	37°C for 24 h	Transgenerational	—	—	siRNA	—	Ito et al. (2011)
Arabidopsis thaliana	37°C for 2 h	ND	—	—	miR408	CSD1, CSD2, GST-U25, CCS1, SAP12	Ma et al. (2015)

[a]*ND, No data available.*

In maize, the levels of H3K4me3, H3K27me3, and H3K9ac were investigated after exposure to prolonged drought conditions and in the recovery period. Modifications of H3K4me3 and H3K9ac were found to be positively correlated with the gene transcription level. Histone modifications of H3K4me3, H3K9ac indeed serve as a blueprint for stress memory. Transcription levels of stress-responsive genes for abscisic acid synthesis and signaling pathways were either maintained longer, even after the recovery period (example: *ZEP1*, *NCAD6*, *AP2/EREBP*, *NAC*), or some responsive genes stored the signal for a delayed response (example: *MADS4,15*) (Forestan et al., 2020).

Modifications of DNA methylation of genes, promoters, or transcription factors in response to drought stress have been detected in various plant species (Akhter et al., 2021; Czajka et al., 2021). Under drought stress conditions, plants exhibit dynamic and variable methylation levels, however the methylation changes are not always related to known transcriptome regulation associated with that stress. In mulberry plants, the overall methylation level of plants subjected to drought stress was 8.64% higher than that of well-watered ones (Li et al., 2020), while about 29% of DNA methylation processes were detected to be irreversible in rice plants exposed to drought stress (Wang et al., 2010). The degree, level and polymorphism of DNA methylation were different in wheat (Duan et al., 2020) or rice (Zheng et al., 2013; Wang et al., 2016) varieties differing in their response to water deficit. In general, drought increases the level of DNA methylation in non-adapted plants, however, if plants were previously adapted to the stress, the DNA methylation level decreased (Rendina González et al., 2018). Under drought conditions, hypermethylation events occur in the drought-susceptible genotypes while drought-tolerant genotypes present hypomethylation behavior (Gayacharan and Joel, 2013). DNA methylation can persist through some generations leading to transgenerational plasticity of the offspring (Herman and Sultan, 2016). Zheng et al. (2017) found that rice exposed to drought conditions had several stable methylation changes in stress-responsive genes which were passed on to progeny for multiple generations (Zheng et al., 2017).

The BRAHMA-type ATP-dependent chromatin remodeling factors (CHR12 or SW1/SNF2) play an essential regulating role in response to drought in *A. thaliana* (Han et al., 2012). Thousands of regulatory RNAs were identified in response to drought stress including miRNA, hc-siRNA, sRNA, and lncRNA-mediated regulation of gene expression and post-transcriptional modifications in several crops (Jha et al., 2020).

2.2 Salinity

High salinity causes ion toxicity and hyperosmotic stress, which inhibit plant development and productivity (Wani et al., 2020). Although the involvement of DNA methylation and different histone modification marks in regulating salt tolerance was demonstrated in various crops, the specific roles of DNA methylation in salt stress responses remain to be clarified (Liu and Lang, 2020). Salt stress induces, in different plant species, opposite effects on 5 mC (methylation or demethylation) of transcriptional regulators, to differentially modulate the downstream expression of salinity-related genes. In soybean and rice, salt stress induces 5 mC demethylation at the promoter of specific transporters, associated with a higher expression and with increased tolerance to salinity stress (Zhu

et al., 2015; Zhang W. et al., 2020). In other cases, salinity stress induces an increase in 5 mC levels that may influence the expression of transporters or miRNA, thereby improving the salt tolerance (Ganie et al., 2016; Kumar et al., 2017). Most of the observed methylation/demethylation changes were stable after plant recovery, implying a possible establishment of a stress memory.

In salt-stressed rice, different DNA methylation patterns were identified in 14 zinc-finger-containing genes (Ahmad et al., 2019). Most methylation/demethylation changes were stable after recovery, implying a possible establishment of stress memory. In Foxtail millet (*Setaria italica* L.), a crop that is considered more tolerant to environmental stresses compared with other cereal crops, a strong decrease in DNA methylation levels was found in a salt-tolerant line when compared with a salt-sensitive variety under salt stress conditions. Promoter regions and coding sequences of several genes were hypomethylated including ABC transporters, WRKY transcription factors, serine-threonine protein phosphatases, and genes related to disease resistance and retrotransposon activation (Pandey et al., 2017). Methylation changes under salt stress were also observed in wheat. For example, the transcriptional level of the GAPC1 (Cytosolic glyceraldehyde-3-phosphate dehydrogenase) gene was induced under both osmotic and salinity stresses, accompanied by decreased methylation of CG and CHG cytosine residues in the promoter region of this gene. GAPC (Cytosolic glyceraldehyde-3-phosphate dehydrogenase) catalyzes a key reaction during glycolysis and was suggested to positively regulate stress responses in plants (Fei et al., 2017).

In maize seedlings, expansin genes were induced in response to salinity, leading to cell enlargement. Expansin-related genes showed elevated promoter H3K9ac levels accompanied by global accumulation of H3K9ac and H4K5ac under salt stress (Li Hui et al., 2014). Elevated levels of H3K9ac and H3K27ac marks were identified in the coding region of the peroxidase (POX) encoding gene in beet plants, which was transcriptionally activated by salt treatments. These marks were linked with high POX transcript abundance in both sugar beet (*Beta vulgaris* L.) and wild beet [*Beta vulgaris subsp. maritima* (L.) Arcang.], but the degree and the site of acetylation were different between the species and subspecies (Yolcu et al., 2016). Lastly, in alfalfa (*Medicago sativa* L.) alterations in the methylation status of the promoter region of the transcription factor MsMYB4 were detected following salinity stress. Activation of MsMYB4 was associated with an increased level of histone H3K4 trimethylation and H3K9 acetylation in it's the corresponding promoter (Dong et al., 2020).

2.3 Cold/High Temperature

Since temperature is a key factor governing plants/crops growth and development, either high or low temperatures limit their productivity and yield. The link between epigenetic processes and plant responses to non-optimal temperature conditions was demonstrated on the molecular, biochemical, and cellular levels (Ueda and Seki, 2020). Recent studies show that the expression of 29 genes in a cold-tolerant rice line was altered under cold stress, in correlation with changes in DNA methylation, mostly at promoter regions (Guo et al., 2019). Similarly, even histone modifications are involved in cold/heat stress response through gene expression tuning (Bannister and Kouzarides, 2011; Kim et al., 2015). This is due to a large number of DNase I hypersensitive sites (DHSs) induced by cold stress, translating this event into enhanced chromatin accessibility.

For instance, in plants is known that the euchromatin mark H3K4me3, which indicates the tri-methylation at the 4th lysine residue of the histone H3 protein, is commonly associated with the activation of transcription of nearby genes (Zhang et al., 2009). In contrast, H3K27me3, indicating the tri-methylation of lysine 27 on histone H3 protein, is associated with one of the major gene silencing systems in plants (Zhang et al., 2007). Genome-wide distributions of these histone modifications and their association with gene expression have been well-documented in several plant species as *A. thaliana* and potato (Zhou et al., 2010). The cold stress may induce the H3K27me3 deposition, which, in turn, has been demonstrated to be involved, for instance, in Flowering Locus C downregulation.

However, several cases of bivalent histone modifications of H3K4me3 and H3K27me3 are known and associated with cold stress related genes (about 6,500). In particular, Zeng et al. (2019) demonstrated that active genes (transcribed in both conditions) displayed enhanced chromatin accessibility upon cold storage. Upregulated genes, associated with this bivalent mark, were enriched in functions and related to the stress response, while the downregulated genes were involved in the developmental processes. The authors hypothesized that the bivalent H3K4me3-H3K27me3 mark represents, in potato tubers, a distinct chromatin environment with greater accessibility, which might facilitate the access of regulatory proteins required for gene upregulation or downregulation in response to cold stress.

Regarding heat stress, CHH methylation patterns differed between two rice lines showing different levels of heat tolerance (He et al., 2020). In barley (*Hordeum vulgare* L.), increasing air temperature by 3°C led to increased levels and altering DNA methylation patterns while in cotton (*Gossypium hirsutum* L.) prolonged heat stress led to methylation changes in the promoter of anther-expressed genes. These changes promoted a series of redox processes to support a different development program under stress conditions (Zhang X. et al., 2020).

2.4 Visible and Ultraviolet Light

Light is essential for photosynthesis and also for conveying information on environmental conditions such as wavelength composition, direction, intensity, and photoperiod. Plant photoreceptors are specialized in perceiving light stimuli ranging from ultraviolet (UV) to visible and far-red (FR) irradiation that induce downstream signaling events including major transcriptional reprogramming. There is increasing evidence of how light triggers changes in chromatin compaction, nuclear morphology as well as influencing histone modifications and gene repositioning (Perrella et al., 2020). Plant photoreceptors and downstream signaling components interact and modulate the action of chromatin remodeling enzymes and transcriptional regulators that confer light-induced chromatin changes through the deposition of epigenetic marks. Early studies

revealed that histone acetylation is associated with the induction of gene expression in response to light. These initial observations were further verified by physiological and molecular experiments on mutants of histone acetyltransferase (GCN5, HAF2) and deacetylase enzymes (HDA15, HDA6) (Bourbousse et al., 2020; Perrella et al., 2020). Furthermore, the role of histone H2B mono-ubiquitination was linked to light-induced activation of gene expression in light-grown *A. thaliana* seedlings by facilitating the activity of RNA Polymerase II (Bourbousse et al., 2012). Changes in histone methylation levels have been also associated with shade avoidance responses, which are triggered by a decrease in the R/FR ratio due to canopy coverage. Shade induces growth-promoting genes leading to the elongation of hypocotyls, stems, and petioles in search of light (Martínez-García and Moreno-Romero, 2020). An increase in H3K4me3 and H3K36me3 levels and recruitment of the histone methylation reader MRG2 on growth-promoting loci such as *YUCCA8* has been reported to mediate shade-induced physiological responses in *A. thaliana*.

Studies in *A. thaliana* and maize have shown that UV-B induces an increase in histone H3 and H4 acetylation levels (Casati et al., 2006; Casati et al., 2008; Campi et al., 2012). Members of multiple histone acetyltransferase families, such as HAM1, HAM2, HAC1, HAG3, and HAF1, have also been shown to regulate different UV-B signaling responses (Fina et al., 2017). Furthermore, UV-B can negatively regulate the transcript levels of the Polycomb Repressive Complex two components MSII and CURLY LEAF that control H3K27me3 deposition on the flowering regulating loci *MIR156* and *FLC* (Dotto et al., 2018). As a result, UV-B leads to delayed flowering in *A. thaliana* (Dotto et al., 2018). The UV-B receptor UVR8 also plays a role in controlling a UV-B-dependent increase in the acetylation status of histone H3 lysine K9 and K14 on target genes (Velanis et al., 2016). Furthermore, UVR8 regulates DNA methylation by directly associating and inhibiting DRM2 (Jiang et al., 2021). In addition to regulating plant development, UV-irradiation induces DNA damage responses leading to changes in chromatin and epigenome dynamics. However, the exact molecular interplay among DNA-damage repair and acclimation responses to high light and temperature requires further investigation (Molinier, 2017).

2.5 Heavy Metals and Metalloids

Among the abiotic stresses affecting plant wellness, heavy metal (HM) contamination represents a serious threat also to humans and animals. In plants, exposure to excessive amounts of both essential and non-essential HMs induced toxic effects, activating a broad array of alterations (Edelstein and Ben-Hur, 2018). In this relation, many recent studies suggest that climate change has both a direct and indirect effect on HM leaching and bioavailability (Fan and Shibata, 2015; Xia et al., 2016). The raising temperatures and the related increase in atmospheric CO_2 levels, which indirectly increase chemical weathering due to both temperature and lower pH, lead to the release of metals in the earth's crust and soils (Whitehead et al., 2009; Benítez-Gilabert et al., 2010). On the other hand, precipitation has an impact on surface runoff, river discharge, and thus indirectly on river water quality. Surface runoff is an important carrier of contaminants from the surrounding land (brownfields) to the receiving surface water. The consequences of these effects are leading to degradation in water and sediment quality that could have negative impact on the ecosystems. Growing evidence highlights important roles in plant adaptation to highly HM and metalloids contaminated environments of epigenetic variations, often responsible for modulating gene expression (Cicatelli et al., 2014; Kumar, 2018). This phenomenon is mediated by a complex interplay among different molecular factors: changes in DNA methylation patterns, histone modifications and chromatin remodeling (Dutta et al., 2018). A recent study suggests that DNA demethylation is one of the molecular strategies adopted by *Arundo donax* L. plants to counteract the stress caused by soil arsenic pollution (Guarino et al., 2020).

Many HM-related RNAs have been identified and several findings are indicating their important role as trans-acting epigenetic signals, involved in specific gene regulatory networks activated in response to HM stress in plants (Ding et al., 2020). For instance, aluminum can induce a hypomethylation of the NtGPDL gene in tobacco. This carries the information for an aluminum stress-activated glycerophosphodiesterase (Choi and Sano, 2007). In wheat, hypermethylation was obtained with the highest concentrations of aluminum and hypomethylation with the lowest ones (Hossein Pour, 2019). Gallo-Franco et al. (2020) calculated the methylation level of 26 genes from the IR64, Nipponbare, and Pokkali varieties of rice plants using data in the scientific literature and the Rice SNP Seek database. All three varieties were hypermethylated with the highest levels in the Nipponbare variety, and the ART1 and STAR1 genes were differently methylated. These genes encode and regulate the transcription of transmembrane proteins useful for aluminum detoxification. Similarly, Gullì et al. (2018) showed that a specimen of *Noccaea caerulescens* grown in an area with nickel-rich soil showed a genome that was more methylated than the control. The overexpressed genes were the *MET1* DNA methyltransferase, the *HDA8* histone deacetylase, and the *DRM2* DNA methyltransferase involved in RdDM. All three were upregulated from 3 to 16-fold.

2.6 Nanomaterials

A large number of new materials is produced for human use. Their environmental dispersal under climate change conditions has led to increased pollution and risk to the health of plants, animals, and humans. Waste dispersed in the environment undergoes degradation processes that cause the dispersion of nanoparticles and pollutants with varying toxicological characteristics (Nejdl et al., 2020).

Nanomaterials are particles smaller than 100 nm that can be of natural or artificial origin. The first category includes those formed by natural processes such as volcanic activities and air particles while the second includes those synthesized for biomedical and industrial purposes. Studying the effects of nanomaterials on the plants is therefore critical to understanding the impact of the pollutant on the ecosystem. Nanoplastics as new pollutants can get adsorbed by plants. The main route of plant intoxication is the root route with the uptake of nanoparticles from the polluted soil (Deng et al., 2014). Roots

can take up nanoparticles symplastically or apoplastically. In the former, internalization occurs *via* endocytosis or *via* aquaporins, the number of which affects the uptake (Rico et al., 2011). In the latter, nanoparticles cross the spaces between cells, if their size is smaller (Zhao et al., 2012). Absorption can also occur through leaves but to lesser extent (Deng et al., 2014). Other general effects include the production of ROS resulting in increased lipid peroxidation, DNA degradation, and cell death (Tarrahi et al., 2021). So far, few studies have addressed the possible involvement of epigenetic processes in the response of plants to these types of stress. However, the observation that ROS are part of the cell response to nanomaterials could suggest that ROS-mediated epigenetic regulation is also involved.

Carbon-based nanomaterials are to date used for a lot of industrial purposes and studied for their nanotoxicology in plants (Marmiroli and White, 2016). It has been shown how carbon nanotubes can pierce the root walls of plants and enter both apoplastically and symplastically (Tripathi et al., 2017). Once adsorbed, they can reach organelles such as mitochondria and chloroplasts and especially in the nucleus of plant cells (Jordan et al., 2018). In *Allium cepa* L. several variations were observed depending on the concentration of the Multi-Walled Nanomaterials (MWNM) used. The cutting sequence of the restriction enzyme *Hpa* II was found to be hypomethylated at low concentration and hypermethylated at high concentration (Ghosh et al., 2015). Single-wall and multi-wall carbon nanotubes were found to promote rice root growth, by eliciting similar molecular pathways and epigenetic regulation (Yan et al., 2016).

Contrasting results were obtained for silver nanoparticles (AgNPs), a large family of materials used from the home appliance industry to the cosmetic industry. AgNPs formulation Argovit™ showed no cyto- or genotoxic damage or epigenetic effects in *A. cepa* (Casillas-Figueroa et al., 2020). However, in *A. thaliana*, the evaluated concentrations resulted in increased expression of genes involved in glutathione biosynthesis, glutathione S- transferase, and glutathione reductase (Nair and Chung, 2014).

3 FROM ALPHABET TO SYNTAX – RECOMMENDATIONS FOR THE FUTURE

Enormous progress has been made in understanding the role of epigenetic regulation in crop response to different stresses. However, in order to make step and translate an "epigenetic alphabet" into "epigenetic syntax" and evolve from "experimental" to "classical" methodology in crop breeding, epigenetics needs to overcome four main challenges:

(1) Need for improved experimental procedures, especially in sequencing technology (longer reads, deeper single-cell sequencing, more efficient sample preparation kits, sequencing portable device improvements);
(2) Need for improved workflows of data analysis, as epigenomic data are currently dispersed, obtained with different methodologies and approaches. Indeed, there is an urgent need of defining and delivering approved methodological standards for both wet-lab and *in silico* analysis. The first steps are made in this direction with solutions offered to improve data workflow systems with cloud services and use of open data for bioinformatics research (Rezaul Karim et al., 2018) and development of standardized workflows for epigenetic data such as ARPEGGIO (Milosavljevic et al., 2021);
(3) Need for enhanced knowledge on crop species at all epigenetic levels as well as interactions between epigenetic machinery and other TF or DNA binding proteins to gain insight into the interactions between epigenome and changes in DNA sequences. Future directions to hasten application of epigenetic modifications in crop breeding strategies for specific agronomical traits have been proposed by several authors (Gallusci et al., 2017; Varotto et al., 2020; Kakoulidou et al., 2021), and need to be applied on wider scale in order to transfer knowledge from model plants to crops;
(4) Need to better integrate epigenomic data with other "omics" data, since epigenomic data are difficult to match with data obtained at other "omics" levels. This highlights the need for agreeing which standards and workflows have to be followed in experiments comprising different "omics" analyses. Hence, constructive and methodological guidelines on how to perform multi-omics data integration (MOI) in plants are needed. Studies of Jamil et al. (2020) who propose three levels of MOI—element-based, pathway-based and mathematical-based integration and Grabowski and Rappsilber (2019) who provide practical guide on how to move from data to insight while using easily accessible data sources, could be good models for future work in "omics" data integration.

Overcoming above-mentioned challenges will facilitate: *i*) elucidation of the role of other mechanisms, besides chromatin-based mechanisms, in somatic and inter-generational stress memory and understanding if there is a universal mechanism of stress memory or if different cases of stress memory are modulated in a different way; *ii*) demonstrating if targeted, gene-specific epigenome or epi-transcriptome modifications anticipated responses to stresses, that will allow the identification of key regulatory mechanisms for tailored responses to the new challenges driven by climate change; *iii*) understanding how epigenetic changes can produce new stable phenotypes in a few generations, allowing the plant survival in their natural habitats; *iv*) clarification of the role of chromatin structure modifications in hypersensitivity reaction, contributing to plasticity and plant adaptation in a world context of climate change; *v*) clarification of the role of RdDM machinery, together with other DNA methylation mechanisms targeting and often silencing repetitive elements, highly represented in the plant genome and *vi*) identification of the difference between correlation and causality, that is if a chromatin regulator is required for a particular stress response, it does not necessarily imply that it modulates the stress response, as it may be a passive response affecting gene expression, rather than being an endogenous regulation of the process. Consequently, silencing of a chromatin regulator may cause a stress response not through the action of stress-responsive genes, but indirectly due to phenotypic, metabolic, and developmental modifications.

The interdisciplinary effort of scientists involved in plant biology and crop improvement in resolving the above-mentioned issues and gaining new insights into epigenetics mechanisms involved in plant stress response should pave the way for further understanding of an epigenetic alphabet of plants and its translation into epigenetic syntaxes for further exploitation of epigenetic variation in crop breeding for climate resilience.

AUTHOR CONTRIBUTIONS

Conceptualization: FM; Original draft preparation: FG, AC, SC, DRGA, GEO, SF, JL, JD, EK, MLL, MS, CS, CV, DP, AMGP, PNM, MG, SY, CV, PG, DM, and FM; Review and editing: all authors. All authors have read and agreed to the submitted version of the manuscript.

FUNDING

This article is based upon work from COST Action EPI-CATCH (CA19125), supported by COST (European Cooperation in Science and Technology), www.cost.eu. JD was also supported by the Thematic Excellence Programme (TKP2021-EGA-20) of the Ministry for Innovation and Technology in Hungary, within the framework of the Biotechnology thematic program of the University of Debrecen, and DM by activities of Center of Excellence for Innovations in Breeding of Climate-Resilient Crops—Climate Crops, Institute of Field and Vegetable Crops, Novi Sad, Serbia, as well as by Ministry of Education, Science and Technological Development of the Republic of Serbia, grant number: 451-03-9/2021-14/200032, and the Science Fund of the Republic of Serbia, through IDEAS project "Creating climate smart sunflower for future challenges" (SMARTSUN) grant number 7732457.

REFERENCES

Ahmad, F., Farman, K., Waseem, M., Rana, R. M., Nawaz, M. A., Rehman, H. M., et al. (2019). Genome-wide Identification, Classification, Expression Profiling and DNA Methylation (5mC) Analysis of Stress-Responsive ZFP Transcription Factors in rice (Oryza Sativa L.). *Gene* 718, 144018. doi:10.1016/j.gene.2019.144018

Akhter, Z., Bi, Z., Ali, K., Sun, C., Fiaz, S., Haider, F. U., et al. (2021). In Response to Abiotic Stress, DNA Methylation Confers EpiGenetic Changes in Plants. *Plants* 10, 1096. doi:10.3390/plants10061096

Altieri, M. A., Nicholls, C. I., Henao, A., and Lana, M. A. (2015). Agroecology and the Design of Climate Change-Resilient Farming Systems. *Agron. Sustain. Dev.* 35 (3), 869–890. doi:10.1007/s13593-015-0285-2

Anderson, S. J., Kramer, M. C., Gosai, S. J., Yu, X., Vandivier, L. E., Nelson, A. D. L., et al. (2018). N6-Methyladenosine Inhibits Local Ribonucleolytic Cleavage to Stabilize mRNAs in Arabidopsis. *Cel Rep.* 25 (5), 1146–1157.e1143. doi:10.1016/j.celrep.2018.10.020

Annunziato, A. (2008). DNA Packaging: Nucleosomes and Chromatin. *Nat. Educat* 1 (1), 26.

Ashapkin, V. V., Kutueva, L. I., Aleksandrushkina, N. I., and Vanyushin, B. F. (2020). Epigenetic Mechanisms of Plant Adaptation to Biotic and Abiotic Stresses. *Ijms* 21 (20), 7457. doi:10.3390/ijms21207457

Bakhshi, B., Fard, E. M., Gharechahi, J., Safarzadeh, M., Nikpay, N., Fotovat, R., et al. (2017). The Contrasting microRNA Content of a Drought Tolerant and a Drought Susceptible Wheat Cultivar. *J. Plant Physiol.* 216, 35–43. doi:10.1016/j.jplph.2017.05.012

Bannister, A. J., and Kouzarides, T. (20112011). Regulation of Chromatin by Histone Modifications. *Cell Res* 21, 381–395. doi:10.1038/cr.2011.22

Benítez-Gilabert, M., Alvarez-Cobelas, M., and Angeler, D. G. (2010). Effects of Climatic Change on Stream Water Quality in Spain. *Climatic Change* 103 (3), 339–352. doi:10.1007/s10584-009-9778-9

Berglund, T., Wallström, A., Nguyen, T.-V., Laurell, C., and Ohlsson, A. B. (2017). Nicotinamide; Antioxidative and DNA Hypomethylation Effects in Plant Cells. *Plant Physiol. Biochem.* 118, 551–560. doi:10.1016/j.plaphy.2017.07.023

Bilichak, A., Ilnytskyy, Y., Wã³ycicki, R., Kepeshchuk, N., Fogen, D., and Kovalchuk, I. (2015). The Elucidation of Stress Memory Inheritance in Brassica Rapa Plants. *Front. Plant Sci.* 6, 1–20. doi:10.3389/fpls.2015.00005

Bobadilla, R., and Berr, A. (2016). "Histone Methylation-A Cornerstone for Plant Responses to Environmental Stresses," in *Abiotic and Biotic Stress in Plants - Recent Advances and Future Perspectives*. Editors A.K. Shanker and C. Shanker (London, UK: IntechOpen), 31–61. doi:10.5772/61733

Bohnert, H. J., Nelson, D. E., and Jensen, R. G. (1995). Adaptations to Environmental Stresses. *Plant Cell* 7 (7), 1099–1111. doi:10.1105/tpc.7.7.1099

Bourbousse, C., Ahmed, I., Roudier, F., Zabulon, G., Blondet, E., Balzergue, S., et al. (2012). Histone H2B Monoubiquitination Facilitates the Rapid Modulation of Gene Expression during Arabidopsis Photomorphogenesis. *PLOS Genet.* 8 (7), e1002825. doi:10.1371/journal.pgen.1002825

Bourbousse, C., Barneche, F., and Laloi, C. (2020). Plant Chromatin Catches the Sun. *Front. Plant Sci.* 10 (1728). doi:10.3389/fpls.2019.01728

Brzezinka, K., Altmann, S., Czesnick, H., Nicolas, P., Gorka, M., Benke, E., et al. (2016). Arabidopsis FORGETTER1 Mediates Stress-Induced Chromatin Memory through Nucleosome Remodeling. *Elife* 5, 1–23. doi:10.7554/elife.17061

Cadavid, I. C., da Fonseca, G. C., and Margis, R. (2020). HDAC Inhibitor Affects Soybean miRNA482bd Expression under Salt and Osmotic Stress. *J. Plant Physiol.* 253, 153261. doi:10.1016/j.jplph.2020.153261

Campi, M., D'Andrea, L., Emiliani, J., and Casati, P. (2012). Participation of Chromatin-Remodeling Proteins in the Repair of Ultraviolet-B-Damaged DNA. *Plant Physiol.* 158 (2), 981–995. doi:10.1104/pp.111.191452

Casati, P., Campi, M., Chu, F., Suzuki, N., Maltby, D., Guan, S., et al. (2008). Histone Acetylation and Chromatin Remodeling Are Required for UV-B-dependent Transcriptional Activation of Regulated Genes in maize. *Plant Cell* 20 (4), 827–842. doi:10.1105/tpc.107.056457

Casati, P., Stapleton, A. E., Blum, J. E., and Walbot, V. (2006). Genome-wide Analysis of High-Altitude maize and Gene Knockdown Stocks Implicates Chromatin Remodeling Proteins in Response to UV-B. *Plant J.* 46 (4), 613–627. doi:10.1111/j.1365-313X.2006.02721.x

Casillas-Figueroa, F., Arellano-García, M. E., Leyva-Aguilera, C., Ruíz-Ruíz, B., Luna Vázquez-Gómez, R., Radilla-Chávez, P., et al. (2020). Argovit Silver Nanoparticles Effects on Allium cepa: Plant Growth Promotion without Cyto Genotoxic Damage. *Nanomaterials* 10 (7), 1386. doi:10.3390/nano10071386

Chachar, S., Liu, J., Zhang, P., Riaz, A., Guan, C., and Liu, S. (2021). Harnessing Current Knowledge of DNA N6-Methyladenosine from Model Plants for Non-model Crops. *Front. Genet.* 12, 668317. doi:10.3389/fgene.2021.668317

Chen, H., Feng, H., Zhang, X., Zhang, C., Wang, T., and Dong, J. (2019). An Arabidopsis E3 Ligase HUB2 Increases Histone H2B Monoubiquitination and Enhances Drought Tolerance in Transgenic Cotton. *Plant Biotechnol. J.* 17 (3), 556–568. doi:10.1111/pbi.12998

Chen, H., Li, Z., and Xiong, L. (2012). A Plant microRNA Regulates the Adaptation of Roots to Drought Stress. *FEBS Lett.* 586, 1742–1747. doi:10.1016/j.febslet.2012.05.013

Chen, K., Tang, W.-S., Zhou, Y.-B., Xu, Z.-S., Chen, J., Ma, Y.-Z., et al. (2020). Overexpression of GmUBC9 Gene Enhances Plant Drought Resistance and Affects Flowering Time via Histone H2B Monoubiquitination. *Front. Plant Sci.* 11. doi:10.3389/fpls.2020.555794

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 Cel Physiol* 56, 73–83. doi:10.1093/pcp/pcu149

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. Genet. Genomics* 277 (5), 589–600. doi:10.1007/s00438-007-0209-1

Cicatelli, A., Todeschini, V., Lingua, G., Biondi, S., Torrigiani, P., and Castiglione, S. (2014). Epigenetic Control of Heavy Metal Stress Response in Mycorrhizal versus Non-mycorrhizal poplar Plants. *Environ. Sci. Pollut. Res.* 21 (3), 1723–1737. doi:10.1007/s11356-013-2072-4

Cortijo, S., Charoensawan, V., Brestovitsky, A., Buning, R., Ravarani, C., Rhodes, D., et al. (2017). Transcriptional Regulation of the Ambient Temperature Response by H2A.Z Nucleosomes and HSF1 Transcription Factors in Arabidopsis. *Mol. Plant* 10 (10), 1258–1273. doi:10.1016/j.molp.2017.08.014

Cuerda-Gil, D., and Slotkin, R. K. (2016). Non-canonical RNA-Directed DNA Methylation. *Nat. Plants* 2 (11), 16163. Erratum in: Nat Plants. 2016 Dec 12;3: 16211. doi:10.1038/nplants.2016.163

Cullis, C. A. (1986). "Plant DNA Variation and Stress," in *Genetics, Development, and Evolution*. Editors J.P. Gustafson, G.L. Stebbins, and F.J. Ayala (Boston, MA: Springer). doi:10.1007/978-1-4684-5137-5_6

Czajka, K., Mehes-Smith, M., and Nkongolo, K. (2021). DNA Methylation and Histone Modifications Induced by Abiotic Stressors in Plants. *Genes Genom*. doi:10.1007/s13258-021-01191-z

Demeulemeester, M. A. C., Van Stallen, N., and De Proft, M. P. (1999). Degree of DNA Methylation in Chicory (Cichorium Intybus L.): Influence of Plant Age and Vernalization. *Plant Sci.* 142, 101–108. doi:10.1016/s0168-9452(99)00010-2

Deng, Y.-q., White, J. C., and Xing, B.-s. (2014). Interactions between Engineered Nanomaterials and Agricultural Crops: Implications for Food Safety. *J. Zhejiang Univ. Sci. A.* 15 (8), 552–572. doi:10.1631/jzus.A1400165

Ding, Y., Ding, L., Xia, Y., Wang, F., and Zhu, C. (2020). Emerging Roles of microRNAs in Plant Heavy Metal Tolerance and Homeostasis. *J. Agric. Food Chem.* 68 (7), 1958–1965. doi:10.1021/acs.jafc.9b07468

Ding, Y., Fromm, M., and Avramova, Z. (2012). Multiple Exposures to Drought 'train' Transcriptional Responses in Arabidopsis. *Nat. Commun.* 3. doi:10.1038/ncomms1732

Dong, W., Gao, T., Wang, Q., Chen, J., Lv, J., and Song, Y. (2020). Salinity Stress Induces Epigenetic Alterations to the Promoter of MsMYB4 Encoding a Salt-Induced MYB Transcription Factor. *Plant Physiol. Biochem.* 155, 709–715. doi:10.1016/j.plaphy.2020.08.015

Dotto, M., Gómez, M. S., Soto, M. S., and Casati, P. (2018). UV-B Radiation Delays Flowering Time through Changes in the PRC2 Complex Activity and miR156 Levels in *Arabidopsis thaliana*. *Plant Cel Environ* 41 (6), 1394–1406. doi:10.1111/pce.13166

Du, J., Johnson, L. M., Jacobsen, S. E., and Patel, D. J. (2015). DNA Methylation Pathways and Their Crosstalk with Histone Methylation. *Nat. Rev. Mol. Cel Biol* 16 (9), 519–532. doi:10.1038/nrm4043

Du, J., Zhong, X., Bernatavichute, Y. V., Stroud, H., Feng, S., Caro, E., et al. (2012). Dual Binding of Chromomethylase Domains to H3K9me2-Containing Nucleosomes Directs DNA Methylation in Plants. *Cell* 151 (1), 167–180. doi:10.1016/j.cell.2012.07.034

Duan, H.-C., Wei, L.-H., Zhang, C., Wang, Y., Chen, L., Lu, Z., et al. (2017). ALKBH10B Is an RNA N6-Methyladenosine Demethylase Affecting Arabidopsis Floral Transition. *Plant Cell* 29 (12), 2995–3011. doi:10.1105/tpc.16.00912

Duan, H., Li, J., Zhu, Y., Jia, W., Wang, H., Jiang, L., et al. (2020). Responsive Changes of DNA Methylation in Wheat (*Triticum aestivum*) under Water Deficit. *Sci. Rep.* 10, 7938. doi:10.1038/s41598-020-64660-7

Dutta, S., Mitra, M., Agarwal, P., Mahapatra, K., De, S., Sett, U., et al. (2018). Oxidative and Genotoxic Damages in Plants in Response to Heavy Metal Stress and Maintenance of Genome Stability. *Plant Signaling Behav.* 13 (8, 49), 1. doi:10.1080/15592324.2018.1460048

Edelstein, M., and Ben-Hur, M. (2018). Heavy Metals and Metalloids: Sources, Risks and Strategies to Reduce Their Accumulation in Horticultural Crops. *Scientia Horticulturae* 234, 431–444. doi:10.1016/j.scienta.2017.12.039

Erdmann, R. M., and Picard, C. L. (2020). RNA-directed DNA Methylation. *Plos Genet.* 16 (10), e1009034. doi:10.1371/journal.pgen.1009034

Faltýnek, D., Lacková, L., and Owsianková, H. (2020). Once Again about the Hapax Grammar: Epigenetic Linguistics. *Linguist Front.* 3 (1), 23–27. doi:10.2478/lf-2019-0002

Fan, M., and Shibata, H. (2015). Simulation of Watershed Hydrology and Stream Water Quality under Land Use and Climate Change Scenarios in Teshio River Watershed, Northern Japan. *Ecol. Indicators* 50, 79–89. doi:10.1016/j.ecolind.2014.11.003

Fang, Y., Xie, K., and Xiong, L. (2014). Conserved miR164-Targeted NAC Genes Negatively Regulate Drought Resistance in rice. *J. Exp. Bot.* 65, 2119–2135. doi:10.1093/jxb/eru072

Fard, E. M., Bakhshi, B., Farsi, M., Kakhki, A. M., Nikpay, N., Ebrahimi, M. A., et al. (2017). MicroRNAs Regulate the Main Events in rice Drought Stress Response by Manipulating the Water Supply to Shoots. *Mol. Biosyst.* 13 (11), 2289–2302. doi:10.1039/C7MB00298J

Fei, Y., Xue, Y., Du, P., Yang, S., and Deng, X. (2017). Expression Analysis and Promoter Methylation under Osmotic and Salinity Stress of TaGAPC1 in Wheat (*Triticum aestivum* L). *Protoplasma* 254 (2), 987–996. doi:10.1007/s00709-016-1008-5

Feng, X. J., Li, J. R., Qi, S. L., Lin, Q. F., Jin, J. B., and Hua, X. J. (2016). Light Affects Salt Stress-Induced Transcriptional Memory ofP5CS1inArabidopsis. *Proc. Natl. Acad. Sci. USA* 113, E8335–E8343. doi:10.1073/pnas.1610670114

Fina, J. P., Masotti, F., Rius, S. P., Crevacuore, F., and Casati, P. (2017). HAC1 and HAF1 Histone Acetyltransferases Have Different Roles in UV-B Responses in Arabidopsis. *Front. Plant Sci.* 8, 1179. doi:10.3389/fpls.2017.01179

Finnegan, E. J., Genger, R. K., Kovac, K., Peacock, W. J., and Dennis, E. S. (1998). DNA Methylation and the Promotion of Flowering by Vernalization. *Proc. Natl. Acad. Sci.* 95 (10), 5824–5829. doi:10.1073/pnas.95.10.5824

Forestan, C., Farinati, S., Zambelli, F., Pavesi, G., Rossi, V., and Varotto, S. (2020). Epigenetic Signatures of Stress Adaptation and Flowering Regulation in Response to Extended Drought and Recovery in *Zea mays*. *Plant Cel Environ* 43 (1), 55–75. doi:10.1111/pce.13660

Fu, Y.-L., Zhang, G.-B., Lv, X.-F., Guan, Y., Yi, H.-Y., and Gong, J.-M. (2013). Arabidopsis Histone Methylase CAU1/PRMT5/SKB1 Acts as an Epigenetic Suppressor of the Calcium Signaling Gene Cas to Mediate Stomatal Closure in Response to Extracellular Calcium. *Plant Cell* 25, 2878–2891. doi:10.1105/tpc.113.113886

Fu, Y., and Zhuang, X. (2020). m6A-binding YTHDF Proteins Promote Stress Granule Formation. *Nat. Chem. Biol.* 16 (9), 955–963. doi:10.1038/s41589-020-0524-y

Gallo-Franco, J. J., Sosa, C. C., Ghneim-Herrera, T., and Quimbaya, M. (20202025). Epigenetic Control of Plant Response to Heavy Metal Stress: A New View on Aluminum Tolerance. *Front. Plant Sci.* 11. doi:10.3389/fpls.2020.602625

Gallusci, P., Dai, Z., Génard, M., Gauffretau, A., Leblanc-Fournier, N., Richard-Molard, C., et al. (2017). Epigenetics for Plant Improvement: Current Knowledge and Modeling Avenues. *Trends Plant Sci.* 22, 610–623. doi:10.1016/j.tplants.2017.04.009

Ganie, S. A., Borgohain, M. J., Kritika, K., Talukdar, A., Pani, D. R., and Mondal, T. K. (2016). Assessment of Genetic Diversity of Saltol QTL Among the rice (Oryza Sativa L.) Genotypes. *Physiol. Mol. Biol. Plants* 22, 107–114. doi:10.1007/s12298-016-0342-6

Gayarcharan, A., and Joel, J. (2013). Epigenetic Responses to Drought Stress in Rice (*Oryza Sativa* L.). *Physiol. Mol. Biol. Plants* 19 (3), 379–387. doi:10.1007/s12298-013-0176-4

Gebelin, V., Leclercq, J., KuswanhadiArgout, X., Argout, X., Chaidamsari, T., Hu, S., et al. (2013). The Small RNA Profile in Latex from Hevea Brasiliensis Trees Is Affected by Tapping Panel Dryness. *Tree Physiol.* 33 (10), 1084–1098. doi:10.1093/treephys/tpt076

Ghosh, M., Bhadra, S., Adegoke, A., Bandyopadhyay, M., and Mukherjee, A. (2015). MWCNT Uptake in Allium cepa Root Cells Induces Cytotoxic and Genotoxic Responses and Results in DNA Hyper-Methylation. *Mutat. Research/Fundamental Mol. Mech. Mutagenesis* 774, 49–58. doi:10.1016/j.mrfmmm.2015.03.004

Grabowski, P., and Rappsilber, J. (2019). A Primer on Data Analytics in Functional Genomics: How to Move from Data to Insight? *Trends Biochem. Sci.* 44, 21–32. doi:10.1016/j.tibs.2018.10.010

Guarino, F., Miranda, A., Castiglione, S., and Cicatelli, A. (2020). Arsenic Phytovolatilization and Epigenetic Modifications in *Arundo donax* L. Assisted by a PGPR Consortium. *Chemosphere* 251, 126310. doi:10.1016/j.chemosphere.2020.126310

Gullì, M., Marchi, L., Fragni, R., Buschini, A., and Visioli, G. (2018). Epigenetic Modifications Preserve the hyperaccumulatorNoccaea Caerulescensfrom Ni Geno-Toxicity. *Environ. Mol. Mutagen.* 59 (6), 464–475. doi:10.1002/em.22191

Guo, X., Ma, Z., Zhang, Z., Cheng, L., Zhang, X., and Li, T. (2017). Small RNA-Sequencing Links Physiological Changes and RdDM Process to Vegetative-To-Floral Transition in Apple. *Front. Plant Sci.* 8 (873). doi:10.3389/fpls.2017.00873

Guo, Z., Liu, C., Xiao, W., Wang, R., Zhang, L., Guan, S., et al. (2019). Comparative Transcriptome Profile Analysis of Anther Development in Reproductive Stage of Rice in Cold Region under Cold Stress. *Plant Mol. Biol. Rep.* 37 (3), 129–145. doi:10.1007/s11105-019-01137-6

Gutierrez-Beltran, E., Moschou, P. N., Smertenko, A. P., and Bozhkov, P. V. (2015). Tudor Staphylococcal Nuclease Links Formation of Stress Granules and Processing Bodies with mRNA Catabolism in Arabidopsis. *Plant Cell* 27 (3), 926–943. doi:10.1105/tpc.114.134494

Han, S.-K., Sang, Y., Rodrigues, A., Wu, M.-F., Rodriguez, P. L., and Wagner, D. (2012). The SWI2/SNF2 Chromatin Remodeling ATPase BRAHMA Represses Abscisic Acid Responses in the Absence of the Stress Stimulus in Arabidopsis. *Plant Cell* 24 (12), 4892–4906. doi:10.1105/tpc.112.105114

He, C., Zhang, H.-Y., Zhang, Y.-X., Fu, P., You, L.-L., Xiao, W.-B., et al. (2020). Cytosine Methylations in the Promoter Regions of Genes Involved in the Cellular Oxidation Equilibrium Pathways Affect rice Heat Tolerance. *BMC Genomics* 21 (1), 560. doi:10.1186/s12864-020-06975-3

He, J., Jiang, Z., Gao, L., You, C., Ma, X., Wang, X., et al. (2019). Genome-Wide Transcript and Small RNA Profiling Reveals Transcriptomic Responses to Heat Stress. *Plant Physiol.* 181 (2), 609–629. doi:10.1104/pp.19.00403

Herman, J. J., and Sultan, S. E. (2016). DNA Methylation Mediates Genetic Variation for Adaptive Transgenerational Plasticity. *Proc. R. Soc. B.* 283, 20160988. doi:10.1098/rspb.2016.0988

Hossein Pour, A., Özkan, G., Balpinar Nalci, Ö., and Halílŏglu, K. (2019). Estimation of Genomic Instability and DNA Methylation Due to Aluminum (Al) Stress in Wheat (*Triticum aestivum* L.) Using iPBS and CRED-iPBS Analyses. *Turk J. Bot.* 43 (1), 27–37. doi:10.3906/bot-1804-23

Hou, J., Lu, D., Mason, A. S., Li, B., Xiao, M., An, S., et al. (2019). Non-coding RNAs and Transposable Elements in Plant Genomes: Emergence, Regulatory Mechanisms and Roles in Plant Development and Stress Responses. *Planta* 250 (1), 23–40. doi:10.1007/s00425-019-03166-7

Houben, A., Demidov, D., Caperta, A. D., Karimi, R., Agueci, F., and Vlasenko, L. (2007). Phosphorylation of Histone H3 in Plants-A Dynamic Affair. *Biochim. Biophys. Acta (Bba) - Gene Struct. Expr.* 1769 (5-6), 308–315. doi:10.1016/j.bbaexp.2007.01.002

Hu, H., Tian, M., Ding, C., and Yu, S. (2019b). The C/EBP Homologous Protein (CHOP) Transcription Factor Functions in Endoplasmic Reticulum Stress-Induced Apoptosis and Microbial Infection. *Front. Immunol.* 9 (3083). doi:10.3389/fimmu.2018.03083

Hu, Y., Lu, Y., Zhao, Y., and Zhou, D.-X. (2019a). Histone Acetylation Dynamics Integrates Metabolic Activity to Regulate Plant Response to Stress. *Front. Plant Sci.* 10, 1236. doi:10.3389/fpls.2019.01236

Hu, Z., Song, N., Zheng, M., Liu, X., Liu, Z., Xing, J., et al. (2015). Histone Acetyltransferase GCN 5 Is Essential for Heat Stress-responsive Gene Activation and Thermotolerance in Arabidopsis. *Plant J.* 84, 1178–1191. doi:10.1111/tpj.13076

Huang, H., Ullah, F., Zhou, D.-X., Yi, M., and Zhao, Y. (2019a). Mechanisms of ROS Regulation of Plant Development and Stress Responses. *Front. Plant Sci.* 10. doi:10.3389/fpls.2019.00800

Huang, S., Zhang, A., Jin, J. B., Zhao, B., Wang, T. J., Wu, Y., et al. (2019b). Arabidopsis Histone H3K4 Demethylase JMJ 17 Functions in Dehydration Stress Response. *New Phytol.* 223, 1372–1387. doi:10.1111/nph.15874

Huang, S., Zhu, S., Kumar, P., and MacMicking, J. D. (2021). A Phase-Separated Nuclear GBPL Circuit Controls Immunity in Plants. *Nature* 594 (7863), 424–429. doi:10.1038/s41586-021-03572-6

Iki, T. (2017). Messages on Small RNA Duplexes in Plants. *J. Plant Res.* 130 (1), 7–16. doi:10.1007/s10265-016-0876-2

Ito, H., Gaubert, H., Bucher, E., Mirouze, M., Vaillant, I., and Paszkowski, J. (2011). An siRNA Pathway Prevents Transgenerational Retrotransposition in Plants Subjected to Stress. *Nature* 472, 115–119. doi:10.1038/nature09861

Iwasaki, M., Hyvärinen, L., Piskurewicz, U., and Lopez-Molina, L. (2019). Non-canonical RNA-Directed DNA Methylation Participates in Maternal and Environmental Control of Seed Dormancy. *eLife* 8, e37434. doi:10.7554/eLife.37434

Jamil, I. N., Remali, J., Azizan, K. A., Nor Muhammad, N. A., Arita, M., Goh, H.-H., et al. (2020). Systematic Multi-Omics Integration (MOI) Approach in Plant Systems Biology. *Front. Plant Sci.* 11, 944. doi:10.3389/fpls.2020.00944

Jenkins, M. F., White, E. P., and Hurlbert, A. H. (2018). The Proportion of Core Species in a Community Varies with Spatial Scale and Environmental Heterogeneity. *PeerJ* 6, e6019. doi:10.7717/peerj.6019

Jha, U. C., Nayyar, H., Jha, R., Khurshid, M., Zhou, M., Mantri, N., et al. (2020). Long Non-coding RNAs: Emerging Players Regulating Plant Abiotic Stress Response and Adaptation. *BMC Plant Biol.* 20 (1), 466. doi:10.1186/s12870-020-02595-x

Jiang, J., Ding, A. B., Liu, F., and Zhong, X. (2020). Linking Signaling Pathways to Histone Acetylation Dynamics in Plants. *J. Exp. Bot.* 71 (17), 5179–5190. doi:10.1093/jxb/eraa202

Jiang, J., Liu, J., Sanders, D., Qian, S., Ren, W., Song, J., et al. (2021). UVR8 Interacts with De Novo DNA Methyltransferase and Suppresses DNA Methylation in Arabidopsis. *Nat. Plants* 7 (2), 184–197. doi:10.1038/s41477-020-00843-4

Jordan, J. T., Singh, K. P., and Cañas-Carrell, J. E. (2018). Carbon-based Nanomaterials Elicit Changes in Physiology, Gene Expression, and Epigenetics in Exposed Plants: A Review. *Curr. Opin. Environ. Sci. Health* 6, 29–35. doi:10.1016/j.coesh.2018.07.007

Kaeppler, S. M., Kaeppler, H. F., and Rhee, Y. (2000). Epigenetic Aspects of Somaclonal Variation in Plants. *Plant Mol. Biol.* 43 (2), 179–188. doi:10.1023/A:1006423110134

Kakoulidou, I., Avramidou, E. V., Baránek, M., Brunel-Muguet, S., Farrona, S., Johannes, F., et al. (2021). Epigenetics for Crop Improvement in Times of Global Change. *Biology* 10, 766. doi:10.3390/biology10080766

Kan, R. L., Chen, J., and Sallam, T. (2022). Crosstalk between Epitranscriptomic and Epigenetic Mechanisms in Gene Regulation. *Trends Genet.* 38, 182–193. doi:10.1016/j.tig.2021.06.014

Kapazoglou, A., Tani, E., Avramidou, E. V., Abraham, E. M., Gerakari, M., Megariti, S., et al. (2021). Epigenetic Changes and Transcriptional Reprogramming upon Woody Plant Grafting for Crop Sustainability in a Changing Environment. *Front. Plant Sci.* 11 (2160). doi:10.3389/fpls.2020.613004

Karim, M. R., Michel, A., Zappa, A., Baranov, P., Sahay, R., and Rebholz-Schuhmann, D. (2018). Improving Data Workflow Systems with Cloud Services and Use of Open Data for Bioinformatics Research. *Brief Bioinformat* 19, 1035–1050. doi:10.1093/bib/bbx039

Kim, J.-M., Sasaki, T., Ueda, M., Sako, K., and Seki, M. (2015). Chromatin Changes in Response to Drought, Salinity, Heat, and Cold Stresses in Plants. *Front. Plant Sci.* 6, 114. doi:10.3389/fpls.2015.00114

Kim, J.-M., To, T. K., Ishida, J., Matsui, A., Kimura, H., and Seki, M. (2012). Transition of Chromatin Status during the Process of Recovery from Drought Stress in Arabidopsis Thaliana. *Plant Cel Physiol* 53, 847–856. doi:10.1093/pcp/pcs053

Kim, J.-M., To, T. K., Matsui, A., Tanoi, K., Kobayashi, N. I., Matsuda, F., et al. (2017). Acetate-mediated Novel Survival Strategy against Drought in Plants. *Nat. Plants* 3. doi:10.1038/nplants.2017.97

Kooke, R., Johannes, F., Wardenaar, R., Becker, F., Etcheverry, M., Colot, V., et al. (2015). Epigenetic Basis of Morphological Variation and Phenotypic Plasticity in *Arabidopsis thaliana*. Arabidopsis thaliana. *Plant Cel* 27 (2), 337–348. doi:10.1105/tpc.114.133025

Kosmacz, M., Gorka, M., Schmidt, S., Luzarowski, M., Moreno, J. C., Szlachetko, J., et al. (2019). Protein and Metabolite Composition of Arabidopsis Stress Granules. *New Phytol.* 222 (3), 1420–1433. doi:10.1111/nph.15690

Kovarík, A., Matyásek, R., Leitch, A., Gazdová, B., Fulnecek, J., and Bezdek, M. (1997). Variability in CpNpG Methylation in Higher Plant Genomes. *Gene* 204 (1-2), 25–33. doi:10.1016/s0378-1119(97)00503-9

Kovařík, A., Koukalová, B., Bezděk, M., and Opatrn´, Z. (1997). Hypermethylation of Tobacco Heterochromatic Loci in Response to Osmotic Stress. *Theor. Appl. Genet.* 95 (1), 301–306. doi:10.1007/s001220050563

Kumar, S., Beena, A. S., Awana, M., and Singh, A. (2017). Salt-Induced Tissue-specific Cytosine Methylation Downregulates Expression of HKT Genes in Contrasting Wheat (*Triticum aestivum* L.) Genotypes. *DNA Cel Biol.* 36 (4), 283–294. doi:10.1089/dna.2016.3505

Kumar, S. (2018). Epigenomics of Plant Responses to Environmental Stress. *Epigenomes* 2 (1), 6. doi:10.3390/epigenomes2010006

Kumar, S., and Mohapatra, T. (2021b). Deciphering Epitranscriptome: Modification of mRNA Bases Provides a New Perspective for Post-transcriptional Regulation of Gene Expression. *Front. Cel Dev. Biol.* 9, 628415. doi:10.3389/fcell.2021.628415

Kumar, S., and Mohapatra, T. (2021a). Dynamics of DNA Methylation and its Functions in Plant Growth and Development. *Front. Plant Sci.* 12, 596236. doi:10.3389/fpls.2021.596236

Lämke, J., and Bäurle, I. (2017). Epigenetic and Chromatin-Based Mechanisms in Environmental Stress Adaptation and Stress Memory in Plants. *Genome Biol.* 18 (1), 124. doi:10.1186/s13059-017-1263-6

Lämke, J., Brzezinka, K., Altmann, S., and Bäurle, I. (2016). A Hit-And-Run Heat Shock Factor Governs Sustained Histone Methylation and Transcriptional Stress Memory. *EMBO J.* 35, 162–175. doi:10.15252/embj

Law, J. A., and Jacobsen, S. E. (2010). Establishing, Maintaining and Modifying DNA Methylation Patterns in Plants and Animals. *Nat. Rev. Genet.* 11 (3), 204–220. doi:10.1038/nrg2719

Leclercq, J., Wu, S., Farinas, B., Pointet, S., Favreau, B., Vignes, H., et al. (2020). Post-transcriptional Regulation of Several Biological Processes Involved in Latex Production in *Hevea Brasiliensis*. *PeerJ* 8, e8932. doi:10.7717/peerj.8932

Lee, H. G., and Seo, P. J. (2019). MYB96 Recruits the HDA15 Protein to Suppress Negative Regulators of ABA Signaling in Arabidopsis. *Nat. Commun.* 10 (1), 1713. doi:10.1038/s41467-019-09417-1

Li, H., Soriano, M., Cordewener, J., Muiño, J. M., Riksen, T., Fukuoka, H., et al. (2014). The Histone Deacetylase Inhibitor Trichostatin a Promotes Totipotency in the Male Gametophyte. *Plant Cell* 26 (1), 195–209. doi:10.1105/tpc.113.116491

Li, R., Hu, F., Li, B., Zhang, Y., Chen, M., Fan, T., et al. (2020). Whole Genome Bisulfite Sequencing Methylome Analysis of mulberry (Morus alba) Reveals Epigenome Modifications in Response to Drought Stress. *Sci. Rep.* 10, 8013. doi:10.1038/s41598-020-64975-5

Li, S., Liu, J., Liu, Z., Li, X., Wu, F., and He, Y. (2014). Heat-Induced Tas1 Target1 Mediates Thermotolerance via Heat Stress Transcription Factor A1a-Directed Pathways in Arabidopsis. *Plant Cell* 26, 1764–1780. doi:10.1105/tpc.114.124883

Li, S., Lin, Y.-C. J., Wang, P., Zhang, B., Li, M., Chen, S., et al. (2018). The AREB1 Transcription Factor Influences Histone Acetylation to Regulate Drought Responses and Tolerance in *Populus trichocarpa*. *Plant Cell* 31 (3), 663–686. doi:10.1105/tpc.18.00437

Li, S., Lin, Y. J., Wang, P., Zhang, B., Li, M., Chen, S., et al. (2019). The AREB1 Transcription Factor Influences Histone Acetylation to Regulate Drought Responses and Tolerance in Populus Trichocarpa. *Plant Cell* 31 (3), 663–686.

Li, W., Cui, X., Meng, Z., Huang, X., Xie, Q., Wu, H., et al. (2012). Transcriptional Regulation of Arabidopsis MIR168a and ARGONAUTE1 Homeostasis in Abscisic Acid and Abiotic Stress Responses. *Plant Physiol.* 158, 1279–1292. doi:10.1104/pp.111.188789

Liang, Z., Geng, Y., and Gu, X. (2018). Adenine Methylation: New Epigenetic Marker of DNA and mRNA. *Mol. Plant* 11, 1219–1221. doi:10.1016/j.molp.2018.08.001

Liu, D., Song, Y., Chen, Z., and Yu, D. (2009). Ectopic Expression of miR396 suppressesGRFtarget Gene Expression and Alters Leaf Growth inArabidopsis. *Physiol. Plant* 136, 223–236. doi:10.1111/j.1399-3054.2009.01229.x

Liu, J., and He, Z. (2020). Small DNA Methylation, Big Player in Plant Abiotic Stress Responses and Memory. *Front. Plant Sci.* 11, 595603. doi:10.3389/fpls.2020.595603

Liu, M., Yu, H., Zhao, G., Huang, Q., Lu, Y., and Ouyang, B. (2017). Profiling of Drought-Responsive microRNA and mRNA in Tomato Using High-Throughput Sequencing. *BMC Genomics* 18 (1), 481. doi:10.1186/s12864-017-3869-1

Liu, N., Fromm, M., and Avramova, Z. (2014). H3K27me3 and H3K4me3 Chromatin Environment at Super-induced Dehydration Stress Memory Genes of Arabidopsis Thaliana. *Mol. Plant* 7, 502–513. doi:10.1093/mp/ssu001

Liu, R., and Lang, Z. (2020). The Mechanism and Function of Active DNA Demethylation in Plants. *J. Integr. Plant Biol.* 62 (1), 148–159. doi:10.1111/jipb.12879

Liu, Y., Zhang, A., Yin, H., Meng, Q., Yu, X., Huang, S., et al. (2018). Trithorax-group Proteins ARABIDOPSIS TRITHORAX4 (ATX4) and ATX5 Function in Abscisic Acid and Dehydration Stress Responses. *New Phytol.* 217 (4), 1582–1597. doi:10.1111/nph.14933

Lu, Y., Xu, Q., Liu, Y., Yu, Y., Cheng, Z.-Y., Zhao, Y., et al. (2018). Dynamics and Functional Interplay of Histone Lysine Butyrylation, Crotonylation, and Acetylation in rice under Starvation and Submergence. *Genome Biol.* 19, 144. doi:10.1186/s13059-018-1533-y

Ma, C., Burd, S., and Lers, A. (2015). miR408is Involved in Abiotic Stress Responses in Arabidopsis. *Plant J.* 84, 169–187. doi:10.1111/tpj.12999

Ma, S., Tang, N., Li, X., Xie, Y., Xiang, D., Fu, J., et al. (2019). Reversible Histone H2B Monoubiquitination Fine-Tunes Abscisic Acid Signaling and Drought Response in Rice. *Mol. Plant* 12 (2), 263–277. doi:10.1016/j.molp.2018.12.005

Maeji, H., and Nishimura, T. (2018). Epigenetic Mechanisms in Plants. *Adv. Biol. Res.* 88, 21–47. doi:10.1016/bs.abr.2018.09.014

March, E., and Farrona, S. (2018). Plant Deubiquitinases and Their Role in the Control of Gene Expression through Modification of Histones. *Front. Plant Sci.* 8, 2274. doi:10.3389/fpls.2017.02274

March-Díaz, R., García-Domínguez, M., Lozano-Juste, J., León, J., Florencio, F. J., and Reyes, J. C. (2008). Histone H2A.Z and Homologues of Components of the SWR1 Complex Are Required to Control Immunity in Arabidopsis. *Plant J.* 53 (3), 475–487. doi:10.1111/j.1365-313X.2007.03361.x

Marmiroli, N., and White, J. C. (2016). Editorial: Nanotoxicology and Environmental Risk Assessment of Engineered Nanomaterials (ENMs) in Plants. *Front. Plant Sci.* 7 (1370). doi:10.3389/fpls.2016.01370

Martienssen, R., and Richards, E. J. (1995). DNA Methylation in Eukaryotes. *Curr. Opin. Genet. Develop* 5 (2), 234–242. doi:10.1016/0959-437X(95)80014-X

Martínez-García, J. F., and Moreno-Romero, J. (2020). Shedding Light on the Chromatin Changes that Modulate Shade Responses. *Physiol. Plantarum* 169 (3), 407–417. doi:10.1111/ppl.13101

Martínez-Pérez, M., Aparicio, F., López-Gresa, M. P., Bellés, J. M., Sánchez-Navarro, J. A., and Pallás, V. (2017). Arabidopsis m6A Demethylase Activity Modulates Viral Infection of a Plant Virus and the m6A Abundance in its Genomic RNAs. *Proc. Natl. Acad. Sci. USA* 114 (40), 10755–10760. doi:10.1073/pnas.1703139114

Ménard, R., Verdier, G., Ors, M., Erhardt, M., Beisson, F., and Shen, W.-H. (2014). Histone H2B Monoubiquitination Is Involved in the Regulation of Cutin and Wax Composition in *Arabidopsis thaliana*. *Plant Cel Physiol* 55 (2), 455–466. doi:10.1093/pcp/pct182

Meyer, P., and Heidmann, I. (1994). Epigenetic Variants of a Transgenic Petunia Line Show Hypermethylation in Transgene DNA: an Indication for Specific Recognition of Foreign DNA in Transgenic Plants. *Mol. Gen. Genet.* 243 (4), 390–399. doi:10.1007/bf00280469

Meyer, P., Linn, F., Heidmann, I., Meyer, H., Niedenhof, I., and Saedler, H. (1992). Endogenous and Environmental Factors Influence 35S Promoter Methylation of a maize A1 Gene Construct in Transgenic Petunia and its Colour Phenotype. *Mol. Gen. Genet.* 231 (3), 345–352. doi:10.1007/bf00292701

Milosavljevic, S., Kuo, T., Decarli, S., Mohn, L., Sese, J., Shimizu, K. K., et al. (2021). ARPEGGIO: Automated Reproducible Polyploid EpiGenetic GuIdance workflOw. *BMC Genomics* 22, 547. doi:10.1186/s12864-021-07845-2

Mladenov, V., Fotopoulos, V., Kaiserli, E., Karalija, E., Maury, S., Baranek, M., et al. (2021). Deciphering the Epigenetic Alphabet Involved in Transgenerational Stress Memory in Crops. *Ijms* 22 (13), 7118. doi:10.3390/ijms22137118

Molinier, J. (2017). Genome and Epigenome Surveillance Processes Underlying UV Exposure in Plants. *Genes* 8 (11), 316. doi:10.3390/genes8110316

Nair, P. M. G., and Chung, I. M. (2014). Assessment of Silver Nanoparticle-Induced Physiological and Molecular Changes in *Arabidopsis thaliana*. *Environ. Sci. Pollut. Res.* 21 (14), 8858–8869. doi:10.1007/s11356-014-2822-y

Nejdl, L., Zemankova, K., Havlikova, M., Buresova, M., Hynek, D., Xhaxhiu, K., et al. (2020). UV-induced Nanoparticles-Formation, Properties and Their Potential Role in Origin of Life. *Nanomaterials* 10 (8), 1529. doi:10.3390/nano10081529

Ni, Z., Hu, Z., Jiang, Q., and Zhang, H. (2013). GmNFYA3, a Target Gene of miR169, Is a Positive Regulator of Plant Tolerance to Drought Stress. *Plant Mol. Biol.* 82, 113–129. doi:10.1007/s11103-013-0040-5

Noshay, J. M., and Springer, N. M. (2021). Stories that Can't Be Told by SNPs; DNA Methylation Variation in Plant Populations. *Curr. Opin. Plant Biol.* 61, 101989. doi:10.1016/j.pbi.2020.101989

Ojima, Y., Suryadarma, P., Tsuchida, K., and Taya, M. (2012). Accumulation of Pyruvate by Changing the Redox Status in *Escherichia coli*. *Biotechnol. Lett.* 34 (5), 889–893. doi:10.1007/s10529-011-0842-y

Ou, X., Zhuang, T., Yin, W., Miao, Y., Wang, B., Zhang, Y., et al. (2015). DNA Methylation Changes Induced in Rice by Exposure to High Concentrations of the Nitric Oxide Modulator, Sodium Nitroprusside. *Plant Mol. Biol. Rep.* 33 (5), 1428–1440. doi:10.1007/s11105-014-0843-9

Pan, C., Ye, L., Zheng, Y., Wang, Y., Yang, D., Liu, X., et al. (2017). Identification and Expression Profiling of microRNAs Involved in the Stigma Exsertion under High-Temperature Stress in Tomato. *BMC Genomics* 18 (1), 843. doi:10.1186/s12864-017-4238-9

Pandey, G., Yadav, C. B., Sahu, P. P., Muthamilarasan, M., and Prasad, M. (2017). Salinity Induced Differential Methylation Patterns in Contrasting Cultivars of

Foxtail Millet (Setaria Italica L.). *Plant Cel Rep* 36 (5), 759–772. doi:10.1007/s00299-016-2093-9

Park, J., Lim, C. J., Shen, M., Park, H. J., Cha, J.-Y., Iniesto, E., et al. (2018). Epigenetic Switch from Repressive to Permissive Chromatin in Response to Cold Stress. *Proc. Natl. Acad. Sci. USA* 115 (23), E5400–E5409. doi:10.1073/pnas.1721241115

Patwari, P., Salewski, V., Gutbrod, K., Kreszies, T., Dresen-Scholz, B., Peisker, H., et al. (2019). Surface Wax Esters Contribute to Drought Tolerance in Arabidopsis. *Plant J.* 98, 727–744. doi:10.1111/tpj.14269

Perrella, G., Zioutopoulou, A., Headland, L. R., and Kaiserli, E. (2020). The Impact of Light and Temperature on Chromatin Organization and Plant Adaptation. *J. Exp. Bot.* 71 (17), 5247–5255. doi:10.1093/jxb/eraa154

Poborilova, Z., Ohlsson, A. B., Berglund, T., Vildova, A., Provaznik, I., and Babula, P. (2015). DNA Hypomethylation Concomitant with the Overproduction of ROS Induced by Naphthoquinone Juglone on Tobacco BY-2 Suspension Cells. *Environ. Exp. Bot.* 113, 28–39. doi:10.1016/j.envexpbot.2015.01.005

Prakash, K., and Fournier, D. (2018). Evidence for the Implication of the Histone Code in Building the Genome Structure. *BioSystems* 164, 49–59. doi:10.1016/j.biosystems.2017.11.005

Quadrana, L., and Colot, V. (2016). Plant Transgenerational Epigenetics. *Annu. Rev. Genet.* 50 (50), 467–491. doi:10.1146/annurev-genet-120215-035254

Rendina González, A. P., Preite, V., Verhoeven, K. J. F., and Latzel, V. (2018). Transgenerational Effects and Epigenetic Memory in the Clonal Plant Trifolium Repens. *Front. Plant Sci.* 9, 1677. doi:10.3389/fpls.2018.01677

Rico, C. M., Majumdar, S., Duarte-Gardea, M., Peralta-Videa, J. R., and Gardea-Torresdey, J. L. (2011). Interaction of Nanoparticles with Edible Plants and Their Possible Implications in the Food Chain. *J. Agric. Food Chem.* 59 (8), 3485–3498. doi:10.1021/jf104517j

R. M., S. K., Wang, Y., Zhang, X., Cheng, H., Sun, L., He, S., et al. (2020). Redox Components: Key Regulators of Epigenetic Modifications in Plants. *Ijms* 21, 1419. doi:10.3390/ijms21041419

Rutowicz, K., Puzio, M., Halibart-Puzio, J., Lirski, M., Kroteń, M. A., Kotliński, M., et al. (2015). A Specialized Histone H1 Variant Is Required for Adaptive Responses to Complex Abiotic Stress and Related DNA Methylation in Arabidopsis. *Plant Physiol.* 169 (3), 00493–02101. doi:10.1104/pp.15.00493

Růžička, K., Zhang, M., Campilho, A., Bodi, Z., Kashif, M., Saleh, M., et al. (2017). Identification of Factors Required for M 6 A mRNA Methylation in Arabidopsis Reveals a Role for the Conserved E3 Ubiquitin Ligase HAKAI. *New Phytol.* 215 (1), 157–172. doi:10.1111/nph.14586

Sani, E., Herzyk, P., Perrella, G., Colot, V., and Amtmann, A. (2013). Hyperosmotic Priming of Arabidopsis Seedlings Establishes a Long-Term Somatic Memory Accompanied by Specific Changes of the Epigenome. *Genome Biol.* 14. doi:10.1186/gb-2013-14-6-r59

Schiermeier, Q. (2018). Droughts, Heatwaves and Floods: How to Tell when Climate Change Is to Blame. *Nature* 560 (7716), 20–22. doi:10.1038/d41586-018-05849-9

Sewelam, N., Kazan, K., and Schenk, P. M. (2016). Global Plant Stress Signaling: Reactive Oxygen Species at the Cross-Road. *Front. Plant Sci.* 7, 187. doi:10.3389/fpls.2016.00187

Shahbazian, M. D., and Grunstein, M. (2007). Functions of Site-specific Histone Acetylation and Deacetylation. *Annu. Rev. Biochem.* 76, 75–100. doi:10.1146/annurev.biochem.76.052705.162114

Shen, L., Liang, Z., Gu, X., Chen, Y., Teo, Z. W., Hou, X., et al. (2016). N(6)-Methyladenosine RNA Modification Regulates Shoot Stem Cell Fate in Arabidopsis. *Dev. Cel* 38 (2), 186–200. doi:10.1016/j.devcel.2016.06.008

Shen, X., De Jonge, J., Forsberg, S. K. G., Pettersson, M. E., Sheng, Z., Hennig, L., et al. (2014b). Natural CMT2 Variation Is Associated with Genome-wide Methylation Changes and Temperature Seasonality. *Plos Genet.* 10, e1004842. doi:10.1371/journal.pgen.1004842

Shen, Y., Conde e Silva, N., Audonnet, L., Servet, C., Wei, W., and Zhou, D.-X. (2014a). Over-expression of Histone H3K4 Demethylase Gene JMJ15 Enhances Salt Tolerance in Arabidopsis. *Front. Plant Sci.* 5, 290. doi:10.3389/fpls.2014.00290

Smulders, M. J. M., Rus-Kortekaas, W., and Vosman, B. (1995). Tissue Culture-Induced DNA Methylation Polymorphisms in Repetitive DNA of Tomato Calli and Regenerated Plants. *Theoret. Appl. Genet.* 91 (8), 1257–1264. doi:10.1007/BF00220938

Song, J., Henry, H. A. L., and Tian, L. (2019a). Brachypodium Histone Deacetylase BdHD1 Positively Regulates ABA and Drought Stress Responses. *Plant Sci.* 283, 355–365. doi:10.1016/j.plantsci.2019.03.009

Song, X., Li, Y., Cao, X., and Qi, Y. (2019b). MicroRNAs and Their Regulatory Roles in Plant-Environment Interactions. *Annu. Rev. Plant Biol.* 70, 489–525. doi:10.1146/annurev-arplant-050718-100334

Song, Z.-T., Sun, L., Lu, S.-J., Tian, Y., Ding, Y., and Liu, J.-X. (2015). Transcription Factor Interaction with COMPASS-like Complex Regulates Histone H3K4 Trimethylation for Specific Gene Expression in Plants. *Proc. Natl. Acad. Sci. USA* 112 (9), 2900–2905. doi:10.1073/pnas.1419703112

Song, Z. T., Liu, J. X., and Han, J. J. (2021). Chromatin Remodeling Factors Regulate Environmental Stress Responses in Plants. *J. Integr. Plant Biol.* 63 (3), 438–450. doi:10.1111/jipb.13064

Speybroeck, L., Waele, D., and Vijver, G. (2002). Theories in Early Embryology. *Ann. N Y Acad. Sci.* 981, 7–49. doi:10.1111/j.1749-6632.2002.tb04910.x

Sun, X., Zheng, H., and Sui, N. (2018). Regulation Mechanism of Long Non-coding RNA in Plant Response to Stress. *Biochem. Biophysical Res. Commun.* 503 (2), 402–407. doi:10.1016/j.bbrc.2018.07.072

Sura, W., Kabza, M., Karlowski, W. M., Bieluszewski, T., Kus-Slowinska, M., Pawełoszek, Ł., et al. (2017). Dual Role of the Histone Variant H2A.Z in Transcriptional Regulation of Stress-Response Genes. *Plant Cell* 29 (4), 791–807. doi:10.1105/tpc.16.00573

Tarrahi, R., Mahjouri, S., and Khataee, A. (2021). A Review on *In Vivo* and *In Vitro* Nanotoxicological Studies in Plants: A Headlight for Future Targets. *Ecotoxicology Environ. Saf.* 208, 111697. doi:10.1016/j.ecoenv.2020.111697

Tian, C., Zuo, Z., and Qiu, J.-L. (2015). Identification and Characterization of ABA-Responsive MicroRNAs in Rice. *J. Genet. Genomics* 42, 393–402. doi:10.1016/j.jgg.2015.04.008

Tripathi, D. K., ShwetaSingh, S., Singh, S., Singh, S., Pandey, R., Singh, V. P., et al. (2017). An Overview on Manufactured Nanoparticles in Plants: Uptake, Translocation, Accumulation and Phytotoxicity. *Plant Physiol. Biochem.* 110, 2–12. doi:10.1016/j.plaphy.2016.07.030

Ueda, M., and Seki, M. (2020). Histone Modifications Form Epigenetic Regulatory Networks to Regulate Abiotic Stress Response. *Plant Physiol.* 182 (1), 15–26. doi:10.1104/pp.19.00988

Urquiaga, M. C. d. O., Thiebaut, F., Hemerly, A. S., and Ferreira, P. C. G. (2021). From Trash to Luxury: The Potential Role of Plant LncRNA in DNA Methylation during Abiotic Stress. *Front. Plant Sci.* 11, 603246. doi:10.3389/fpls.2020.603246

Varotto, S., Tani, E., Abraham, E., Krugman, T., Kapazoglou, A., Melzer, R., et al. (2020). Epigenetics: Possible Applications in Climate-Smart Crop Breeding. *J. Exp. Bot.* 71 (17), 5223–5236. doi:10.1093/jxb/eraa188

Velanis, C. N., Herzyk, P., and Jenkins, G. I. (2016). Regulation of Transcription by the Arabidopsis UVR8 Photoreceptor Involves a Specific Histone Modification. *Plant Mol. Biol.* 92 (4-5), 425–443. doi:10.1007/s11103-016-0522-3

Waddington, C. H. (1942). The Epigenotype. *Endeavour* 1, 18–20.

Walter, J., Jentsch, A., Beierkuhnlein, C., and Kreyling, J. (2013). Ecological Stress Memory and Cross Stress Tolerance in Plants in the Face of Climate Extremes. *Environ. Exp. Bot.* 94, 3–8. doi:10.1016/j.envexpbot.2012.02.009

Wang, J., Meng, X., Dobrovolskaya, O. B., Orlov, Y. L., and Chen, M. (2017). Non-coding RNAs and Their Roles in Stress Response in Plants. *Genomics, Proteomics & Bioinformatics* 15 (5), 301–312. doi:10.1016/j.gpb.2017.01.007

Wang, P., Zhao, L., Hou, H., Zhang, H., Huang, Y., Wang, Y., et al. (2015a). Epigenetic Changes Are Associated with Programmed Cell Death Induced by Heat Stress in Seedling Leaves of Zea mays. *Plant Cel Physiol.* 56 (5), 965–976. doi:10.1093/pcp/pcv023

Wang, W.-S., Pan, Y.-J., Zhao, X.-Q., Dwivedi, D., Zhu, L.-H., Ali, J., et al. (2010). Drought-induced Site-specific DNA Methylation and its Association with Drought Tolerance in rice (*Oryza Sativa* L.). *J. Exp. Bot.* 62 (6), 1951–1960. doi:10.1093/jxb/erq391

Wang, W., Qin, Q., Sun, F., Wang, Y., Xu, D., Li, Z., et al. (2016). Genome-Wide Differences in DNA Methylation Changes in Two Contrasting Rice Genotypes in Response to Drought Conditions. *Front. Plant Sci.* 7, 1675. doi:10.3389/fpls.2016.01675

Wang, X., Zhi, P., Fan, Q., Zhang, M., and Chang, C. (2018). Wheat CHD3 Protein TaCHR729 Regulates the Cuticular Wax Biosynthesis Required for Stimulating Germination of Blumeria Graminis f.Sp. Tritici. *J. Exp. Bot.* 70 (2), 701–713. doi:10.1093/jxb/ery377

Wang, Z., Casas-Mollano, J. A., Xu, J., Riethoven, J.-J. M., Zhang, C., and Cerutti, H. (2015b). Osmotic Stress Induces Phosphorylation of Histone H3 at Threonine 3 in Pericentromeric Regions of *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* 112 (27), 8487–8492. doi:10.1073/pnas.1423325112

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 (4), 76. doi:10.1007/s00425-020-03366-6

Waszczak, C., Carmody, M., and Kangasjärvi, J. (2018). Reactive Oxygen Species in Plant Signaling. *Annu. Rev. Plant Biol.* 69, 209–236. doi:10.1146/annurev-arplant-042817-040322

Waters, E. R., and Schaal, B. A. (1996). Heat Shock Induces a Loss of rRNA-Encoding DNA Repeats in Brassica Nigra. *Proc. Natl. Acad. Sci.* 93 (4), 1449–1452. doi:10.1073/pnas.93.4.1449

Whitehead, P. G., Wilby, R. L., Battarbee, R. W., Kernan, M., and Wade, A. J. (2009). A Review of the Potential Impacts of Climate Change on Surface Water Quality. *Hydrological Sci. J.* 54 (1), 101–123. doi:10.1623/hysj.54.1.101

Wibowo, A., Becker, C., Marconi, G., Durr, J., Price, J., Hagmann, J., et al. (20162016). Hyperosmotic Stress Memory in *Arabidopsis* Is Mediated by Distinct Epigenetically Labile Sites in the Genome and Is Restricted in the Male Germline by DNA Glycosylase Activity. *eLife* 5, e13546. doi:10.7554/eLife.13546

Wu, L., Liu, S., Qi, H., Cai, H., and Xu, M. (2020). Research Progress on Plant Long Non-coding RNA. *Plants* 9 (4), 408. doi:10.3390/plants9040408

Xia, X. H., Wu, Q., Mou, X. L., and Lai, Y. J. (2015). Potential Impacts of Climate Change on the Water Quality of Different Water Bodies. *J. Env Inform.* 25 (2), 85–98. doi:10.3808/jei.201400263

Xu, Y., Zhang, S., LinGuo, S., Guo, Y., Deng, W., Zhang, Y., et al. (2017). WERAM: a Database of Writers, Erasers and Readers of Histone Acetylation and Methylation in Eukaryotes. *Nucleic Acids Res.* 45 (D1), D264–D270. doi:10.1093/nar/gkw1011

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 Genomics* 19, 1–17. doi:10.1186/s12864-018-4484-5

Yakovlev, I. A., and Fossdal, C. G. (2017). In Silico Analysis of Small RNAs Suggest Roles for Novel and Conserved miRNAs in the Formation of Epigenetic Memory in Somatic Embryos of Norway Spruce. *Front. Physiol.* 8, 674. doi:10.3389/fphys.2017.00674

Yan, S., Zhang, H., Huang, Y., Tan, J., Wang, P., Wang, Y., et al. (2016). Single-wall and Multi-wall Carbon Nanotubes Promote rice Root Growth by Eliciting the Similar Molecular Pathways and Epigenetic Regulation. *IET nanobiotechnol.* 10 (4), 222–229. doi:10.1049/iet-nbt.2015.0046

Yang, L., Perrera, V., Saplaoura, E., Apelt, F., Bahin, M., Kramdi, A., et al. (2019). m5C Methylation Guides Systemic Transport of Messenger RNA over Graft Junctions in Plants. *Curr. Biol.* 29 (15), 2465–2476.e2465. doi:10.1016/j.cub.2019.06.042

Yao, Q., Song, C.-X., He, C., Kumaran, D., and Dunn, J. J. (2012). Heterologous Expression and Purification of *Arabidopsis thaliana* VIM1 Protein: *In Vitro* Evidence for its Inability to Recognize Hydroxymethylcytosine, a Rare Base in Arabidopsis DNA. *Protein Expr. Purif.* 83, 104–111. doi:10.1016/j.pep.2012.03.003

Yin, C., Ramachandran, S. R., Zhai, Y., Bu, C., Pappu, H. R., and Hulbert, S. H. (2019). A Novel Fungal Effector from Puccinia Graminis Suppressing RNA Silencing and Plant Defense Responses. *New Phytol.* 222 (3), 1561–1572. doi:10.1111/nph.15676

Yolcu, S., Ozdemir, F., Güler, A., and Bor, M. (2016). Histone Acetylation Influences the Transcriptional Activation of POX in Beta Vulgaris L. And Beta Maritima L. under Salt Stress. *Plant Physiol. Biochem.* 100, 37–46. doi:10.1016/j.plaphy.2015.12.019

Zemach, A., Kim, M. Y., Hsieh, P.-H., Coleman-Derr, D., Eshed-Williams, L., Thao, K., et al. (2013). The Arabidopsis Nucleosome Remodeler DDM1 Allows DNA Methyltransferases to Access H1-Containing Heterochromatin. *Cell* 153 (1), 193–205. doi:10.1016/j.cell.2013.02.033

Zeng, Z., Zhang, W., Marand, A. P., Zhu, B., Buell, C. R., and Jiang, J. (2019). Cold Stress Induces Enhanced Chromatin Accessibility and Bivalent Histone Modifications H3K4me3 and H3K27me3 of Active Genes in Potato. *Genome Biol.* 20, 123. doi:10.1186/s13059-019-1731-2

Zhang, H., Lang, Z., and Zhu, J.-K. (2018a). Dynamics and Function of DNA Methylation in Plants. *Nat. Rev. Mol. Cel Biol* 19 (8), 489–506. doi:10.1038/s41580-018-0016-z

Zhang, Q., Liang, Z., Cui, X., Ji, C., Li, Y., Zhang, P., et al. (2018b). N6-methyladenine DNA Methylation in Japonica and Indica rice Genomes and its Association with Gene Expression, Plant Development, and Stress Responses. *Mol. Plant* 11, 1492–1508. doi:10.1016/j.molp.2018.11.005

Zhang, W., Wang, N., Yang, J., Guo, H., Liu, Z., Zheng, X., et al. (2020a). The Salt-Induced Transcription Factor GmMYB84 Confers Salinity Tolerance in Soybean. *Plant Sci.* 291, 110326. doi:10.1016/j.plantsci.2019.110326

Zhang, X., Bernatavichute, Y. V., Cokus, S., Pellegrini, M., and Jacobsen, S. E. (2009). Genome-wide Analysis of Mono-, Di- and Trimethylation of Histone H3 Lysine 4 in *Arabidopsis thaliana*. *Genome Biol.* 10 (6), R62. doi:10.1186/gb-2009-10-6-r62

Zhang, X., Clarenz, O., Cokus, S., Bernatavichute, Y. V., Pellegrini, M., Goodrich, J., et al. (2007). Whole-genome Analysis of Histone H3 Lysine 27 Trimethylation in *Arabidopsis*. *Plos Biol.* 5 (5), e129. doi:10.1371/journal.pbio.0050129

Zhang, X., Ménard, R., Li, Y., Coruzzi, G. M., Heitz, T., Shen, W.-H., et al. (2020b). Arabidopsis SDG8 Potentiates the Sustainable Transcriptional Induction of the Pathogenesis-Related Genes PR1 and PR2 during Plant Defense Response. *Front. Plant Sci.* 11 (277). doi:10.3389/fpls.2020.00277

Zhang, X., Wang, W., Zhu, W., Dong, J., Cheng, Y., Yin, Z., et al. (2019). Mechanisms and Functions of Long Non-coding RNAs at Multiple Regulatory Levels. *Ijms* 20 (22), 5573. doi:10.3390/ijms20225573

Zhang, Z., Zhang, S., Zhang, Y., Wang, X., Li, D., Li, Q., et al. (2011). ArabidopsisFloral Initiator SKB1 Confers High Salt Tolerance by Regulating Transcription and Pre-mRNA Splicing through Altering Histone H4R3 and Small Nuclear Ribonucleoprotein LSM4 Methylation. *Plant Cell* 23, 396–411. doi:10.1105/tpc.110.081356

Zhao, L., Peralta-Videa, J. R., Ren, M., Varela-Ramirez, A., Li, C., Hernandez-Viezcas, J. A., et al. (2012). Transport of Zn in a sandy Loam Soil Treated with ZnO NPs and Uptake by Corn Plants: Electron Microprobe and Confocal Microscopy Studies. *Chemic Engin J.* 184, 1–8. doi:10.1016/j.cej.2012.01.041

Zhao, Y., and Garcia, B. A. (2015). Comprehensive Catalog of Currently Documented Histone Modifications. *Cold Spring Harb Perspect. Biol.* 7 (9), a025064. doi:10.1101/cshperspect.a025064

Zheng, M., Liu, X., Lin, J., Liu, X., Wang, Z., Xin, M., et al. (2018). Histone Acetyltransferase GCN 5 Contributes to Cell wall Integrity and Salt Stress Tolerance by Altering the Expression of Cellulose Synthesis Genes. *Plant J.* 97, 587–602. doi:10.1111/tpj.14144

Zheng, X., Chen, L., Li, M., Lou, Q., Xia, H., Wang, P., et al. (2013). Transgenerational Variations in DNA Methylation Induced by Drought Stress in Two rice Varieties with Distinguished Difference to Drought Resistance. *PLoS One* 8 (11), e80253. doi:10.1371/journal.pone.0080253

Zheng, X., Chen, L., Xia, H., Wei, H., Lou, Q., Li, M., et al. (2017). Transgenerational Epimutations Induced by Multi-Generation Drought Imposition Mediate rice Plant's Adaptation to Drought Condition. *Sci. Rep.* 7, 39843. doi:10.1038/srep39843

Zhong, S., Li, H., Bodi, Z., Button, J., Vespa, L., Herzog, M., et al. (2008). MTA Is an Arabidopsis Messenger RNA Adenosine Methylase and Interacts with a Homolog of a Sex-specific Splicing Factor. *Plant Cell* 20 (5), 1278–1288. doi:10.1105/tpc.108.058883

Zhou, J., Wang, X., He, K., Charron, J.-B. F., Elling, A. A., and Deng, X. W. (2010). Genome-wide Profiling of Histone H3 Lysine 9 Acetylation and Dimethylation in *Arabidopsis* Reveals Correlation between Multiple Histone marks and Gene Expression. *Plant Mol. Biol.* 72 (6), 585–595. doi:10.1007/s11103-009-9594-7

Zhu, N., Cheng, S., Liu, X., Du, H., Dai, M., Zhou, D.-X., et al. (2015). The R2R3-type MYB Gene OsMYB91 Has a Function in Coordinating Plant Growth and Salt Stress Tolerance in rice. *Plant Sci.* 236, 146–156. doi:10.1016/j.plantsci.2015.03.023

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.

Copyright © 2022 Guarino, Cicatelli, Castiglione, Agius, Orhun, Fragkostefanakis, Leclercq, Dobránszki, Kaiserli, Lieberman-Lazarovich, Sõmera, Sarmiento, Vettori, Paffetti, Poma, Moschou, Gašparović, Yousefi, Vergata, Berger, Gallusci, Miladinović and Martinelli. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genetics

ORIGINAL RESEARCH
published: 31 March 2022
doi: 10.3389/fgene.2022.853910

RNA-Seq-Based Transcriptomics Study to Investigate the Genes Governing Nitrogen Use Efficiency in Indian Wheat Cultivars

Sarabjit Kaur[1†], M. Shamshad[2†], Suruchi Jindal[1], Amandeep Kaur[1], Satinder Singh[2], Achla sharma[2] and Satinder Kaur[1*]*

[1]School of Agricultural Biotechnology, Punjab Agricultural University, Ludhiana, India, [2]Department of Plant Breeding and Genetics, Punjab Agricultural University, Ludhiana, India

OPEN ACCESS

Edited by:
Vijay Gahlaut,
Institute of Himalayan Bioresource Technology (CSIR), India

Reviewed by:
Rajeev Ranjan,
Purdue University, United States
Santosh Kumar Upadhyay,
Panjab University, India

***Correspondence:**
Achla sharma
achla@pau.edu
Satinder Kaur
satinder.biotech@pau.edu

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Genomics, a section of the journal Frontiers in Genetics

Received: 13 January 2022
Accepted: 18 February 2022
Published: 31 March 2022

Citation:
Kaur S, Shamshad M, Jindal S, Kaur A, Singh S, sharma A and Kaur S (2022) RNA-Seq-Based Transcriptomics Study to Investigate the Genes Governing Nitrogen Use Efficiency in Indian Wheat Cultivars. Front. Genet. 13:853910. doi: 10.3389/fgene.2022.853910

High NUE (nitrogen use efficiency) has great practical significance for sustainable crop production. Wheat is one of the main cultivated crops worldwide for human food and nutrition. However, wheat grain productivity is dependent upon cultivars with high NUE in addition to the application of nitrogen fertilizers. In order to understand the molecular mechanisms exhibiting a high NUE response, a comparative transcriptomics study was carried out through RNA-seq analysis to investigate the gene expression that regulates NUE, in root and shoot tissue of N-efficient (PBW677) and N-inefficient (703) cultivars under optimum and nitrogen (N) stress. Differentially expressed gene analysis revealed a total of 2,406 differentially expressed genes (DEGs) present in both the contrasting cultivars under N stress. The efficient genotype PBW677 had considerably more abundant DEGs with 1,653 (903 roots +750 shoots) compared to inefficient cultivar PBW703 with 753 (96 roots +657 shoots). Gene ontology enrichment and pathway analysis of these DEGs suggested that the two cultivars differed in terms of adaptive mechanism. Gene enrichment analysis revealed that among the upregulated and downregulated genes the overrepresented and underrepresented gene categories belonged to biological processes like DNA binding, response to abiotic stimulus, photosynthesis, carbon fixation, carbohydrate metabolic process, nitrogen compound metabolic process, nitrate transport, and translation in cultivar PBW677, while the enriched biological processes were nucleosome assembly, chromatin remodeling, DNA packaging, lipid transport, sulfur compound metabolic process, protein modifications, and protein folding and refolding in N inefficient cultivar PBW703. We found several transcription factors (MYB, WRKY, RING finger protein, zinc finger protein, transporters, NRT1, amino acid transporters, sugar), protein kinases, and genes involved in N absorption, transportation, and assimilation to be highly expressed in high NUE cultivar PBW677. In our study, we report 13 potential candidate genes which showed alternate gene expression in the two contrasting cultivars under study. These genes could serve as potential targets for future breeding programs.

Keywords: differential gene expression, nitrogen use efficiency, wheat, transcriptome analysis, roots

INTRODUCTION

Nitrogen is one of the essential macronutrients required for plant growth, development, and reproduction. In last few decades, extensive use of N fertilizers resulted in increased biomass and crop yield. It has been estimated that by the year 2050, the application of N fertilizer will increase to 135 Tg N (Good et al., 2004). But excessive use of N causes environmental pollution, deteriorates soil health, and also leads to a higher cost of production. Thus, increased N use efficiency in plants would not only result in higher crop yield under N-stressed conditions, but also benefit farmers via higher net profit under low input and mitigate the environment risks arising due to an excess of fertilizers. Therefore, one of the main aims in agricultural research is to increase the NUE of plants which depends on plant N-uptake efficiency, N-utilization efficiency, and on the remobilization efficiency of nitrogen from dead tissue to growing plant parts (Masclaux-Daubresse et al., 2010).

Grain production in major cereals largely depends upon the application of N fertilizers and cultivars with high NUE (Hitz et al., 2017). Grain production in crop plants could be enhanced by harnessing the genetic variation for improved NUE. Nitrogen use efficiency is strongly affected by genetic as well as environmental factors (Xu et al., 2012). The key genes involved in the pathways involved in NUE can be discovered using the power of omics. Transcriptomics or RNA-seq is one such approach which can be utilized to reveal the key genes responsible for combating N stress under N-deficit conditions using contrasting genotypes or cultivars having different NUE (Kant et al., 2011). RNA-seq via next-generation sequencing platform allows the discovery of genes playing a role in pathways affecting such phenotyping traits. Many studies are available illustrating the potential of transcriptomics to decipher the role of key genes in nitrogen-dependent pathways, for example the genotypic difference in terms of nitrogen use efficiency has been studied in various crops like rice (Vijayalakshmi et al., 2015), soybean (Hao et al., 2011), sorghum (Gelli et al., 2014), and maize (Zamboni et al., 2014).

Wheat is one of the main cultivated crops worldwide (Xu et al., 2013). To meet the needs of the world's growing population, the grain yield of wheat must be increased at an average annual rate. However, wheat yield is frequently threatened by low-nitrogen stress. Therefore, improving the nitrogen use efficiency for increased biomass as well as increased yield of wheat under N-starvation conditions has great practical significance (Curci et al., 2017). Hence, efforts have been made to understand the molecular and physiological basis of plants grown under N stress, which will help in identification of a large number of differentially expressed genes (DEGs) through RNA sequence-based transcriptomics analysis (Ruuska et al., 2008).

In wheat, attempts have been made to understand the regulatory mechanisms involved in nitrogen metabolism and various studies have been conducted so far to discover the genes via transcriptome sequencing. The RNA-seq studies have been carried out taking various tissue samples, for example leaf tissue, roots, flag leaf, etc., from a single cultivar and differential gene expression was carried out to unravel the key genes and pathways (Meng et al., 2021; Sultana et al., 2021; Tiong et al., 2021).

We have conducted a comparative transcriptomics study using the next-generation sequencing platform Illumina HiSeq 500 using two bread wheat cultivars: PBW677 with high NUE and PBW703 with low NUE (Shamshad, 2019) and studied the whole genome gene expression profile in control as well nitrogen-deficient conditions in root and shoot tissue. This is the first report to be carried out using the two contrasting cultivars and finding the gene expression in root and shoot tissue in normal vs. stress (N-deficient) conditions. As per our knowledge only one study has been conducted in bread wheat taking flag leaf and second leaf tissue after 0 DPA and 10 DPA from three different cultivars having high, medium, and low NUE (Sultana et al., 2020).

MATERIALS AND METHODS

Plant Materials, Nitrogen Treatments, and Plant Sampling

Plant material included two wheat (*Triticum aestivum*) cultivars, PBW677 (PFAU/MILAN/5/CHEN/*Ae. squarrosa*//BCN/3/VEE#7/BOW/4/PASTOR) known to show efficiency at a low input of nitrogen (called N efficient) and PBW703 (called N inefficient) (BWL9250*3//*Yr10*/6*Avocet/3/BW9250*3//*Yr15*/6*Avocet) with low NUE (Shamshad, 2019). The field experiments were carried out using a split-plot design with two N treatments (control and no nitrogen) as the main plot and the two cultivars in three replications as sub plots (2.4 m × 5 m). With each subplot, seeds were sown in 12 5 cm-long rows with 20 cm spacing. The nitrogen was supplied at the rate of 120 Kg ha-1 as controlled conditions and no nitrogen was supplied to induce stress conditions. The crop was grown by following the standard agronomic packaging and practices. For RNA-seq, seeds of both the varieties were surface-sterilized and grown as two sets of three replications in a growth chamber maintained at 26/24°C with relative humidity of 75%. One set of genotypes was sown in perchlorate-contaminated soil without adding any nitrogen fertilizer as N-stress (N^-) while the other set was given nitrogen at a recommended rate of 120 kg/acre as N-control (N^+). To the N-control set, nitrogen fertilizer was applied in the form of urea 21 days after the germination of seeds. Twenty-4 hours after application of urea, root and shoot samples were collected both from N-stress and N-control sets at the same time. Each biological replicate constituted a pool of three plants and a total of three biological replicates were used. The collected eight tissues, (N-stress root, N-control root, N-stress shoot, and N-control shoot) from both genotypes were immediately frozen in liquid nitrogen and stored at −80°C.

Analysis of Growth, Yield, and N Parameters

Three plots per genotype and N treatment were available for the data collection. Leaf chlorophyll content was measured by randomly selecting five plants from each row using an SPAD meter, plant height was measured using scale in cm, and biological yield and grain yield were recorded as yield per plot

in Kg. Grain nitrogen was estimated by using the method proposed by Eastin (2008) from the grounded grain, straw, and leaf tissue samples. The total N content was measured using the distillation and titration method (Pelican Equipment, CIT Nagar, Chennai, Tamil Nadu, India) following the manufacturer's protocol. NUE, N utilization efficiency, and N harvest index were calculated according to Moll et al. (1982). Seed yield relative to total N accumulation in above-ground tillers was used to calculate NUtE, and NUE was determined as the NUpE and NUtE. NHI was analyzed using the ratio of total N in grains to tillers and grains.

RNA Extraction and Reverse Transcription

Total RNA was extracted from 100 mg of root and shoot samples using Trizol Reagent (Takara kit) as per the manufacturer's protocol. The concentration and quality of total RNA were monitored on 1.2% agarose gels as well as on a NanoDrop™ 8,000 Spectrophotometer (Thermo Scientific). Approximately 1 μg of RNA was used as a template for first-strand cDNA synthesis (Thermo Scientific Verso cDNA Kit). Oligo $(dT)_{18}$ primers were used during reverse transcription. The cDNA was stored at −20°C.

RNA-Seq Analysis and High-Quality Read Statistics

To obtain a global overview of the wheat transcriptome and gene activity at a nucleotide resolution, cDNA samples were sequenced by the Illumina Nextseq500 platform. The generated raw reads were submitted to the NCBI sequence read archives (SRA) with accession number PRJNA780342. The raw sequences were assessed for quality using FastQC version 0.11.2 (Andrews et al., 2015). Adaptors clipping and quality trimming of raw reads were performed using Trimmomatic v0.39 software (Bolger et al., 2014). Low quality reads with phred score <30 and read length <50 bp were removed.

Read Alignment and Assessment to Mapping With Reference Wheat Genome

The high-quality transcriptome reads were aligned to the indexed wheat reference genome (http://www.ensembl.org/index.html) using Burrows-Wheeler Aligner (BWA) software (BWA-0.7.17) (Li and Durbin, 2009). The reference *Triticum aestivum* genome (IWGSC RefSeq v1.0) and the associated annotations (IWGSC RefSeq v1.1) were downloaded from the International Wheat Genome Sequencing Consortium (IWGSC). Indexing of the reference genome of wheat (*Triticum aestivum*) was done using BWA-0.7.17 (Li and Durbin, 2009). The resulted output sequence alignment map (SAM) files were converted into binary alignment map (BAM) format, sorted, and indexed using samtools v1.10 (Li et al., 2009).

Differential Gene Expression Analysis

The statistical model of the Cufflinks-cuffdiff v2.2.1 package (Trapnell et al., 2012) was used to assemble and quantify differential gene expression in terms of FPKM (fragments per kilobase per million reads). Genes with log2 fold change (FC) values > 2 were considered upregulated whereas FC < −2 were considered downregulated. These genes were further categorized on the basis of statistical significance ($p < 0.05$).

Functional Annotation and Gene Ontology Term Analysis

Functional Gene Ontology was performed with OmicsBox_windows-64_1_3_11 (http://wego.genomics.org.cn/cgi-bin/wego/index.pl). The FASTA sequences of all the upregulated and downregulated genes were downloaded from the Ensemble Plants database (https://plants.ensembl.org/index.html) and used as input into OmicsBox which contains information of all the genes assigned into three main GO domains, viz., biological process, cellular component, and molecular function. All DEGs were annotated according to wheat IWGSC release 1.1, whereas the sequences lacking annotation in the wheat reference were annotated with *Arabidopsis thaliana* (**Supplementary Tables S1, S2**). Volcano plots were made using R 4.0.4 (ggplot2 package) software which relates the observed differences in gene expression associated with Cuffdiff's statistical model. Venn diagrams were constructed to represent up and downregulated DEGs using the Venny 2.1 tool (https://bioinfogp.cnb.csic.es/tools/venny/) to analyze tissue-specific and tissue-independent genes. R-based 4.0.4 software was used to make heat maps which present hierarchical clustering based on log2 fold changes to visualize the expression patterns of DEGs.

Enrichment Analysis Based on Gene Ontology Terms

To see which class or category of genes were overrepresented and underrepresented among the differentially expressed genes in response to N stress, enrichment analysis was carried out using Fisher's exact test with FDR <0.05. Enriched bar graphs were made using OmicsBox version 2.0.36.

KEGG Pathway Analysis

In order to reveal the pathways to which DEGs under N-stress belong, KEGG pathway analysis was carried out by using the FASTA sequences of all the differentially expressed genes as input to the OmicsBox 2.0.36 combined pathway analysis plugin.

Real-Time Quantitative Polymerase Chain Reaction Analysis

Real-time PCR was performed using the LightCycler 96 Roche Real-time PCR system and PowerUp™ SYBR™ Green Master Mix (applied biosystem by Thermo Fisher Scientific). Seventeen differentially expressed genes (from shoots and roots of both low N and high N) were selected for validation. Primer3 version 2.4.0 was used to design gene-specific primers and their specificity was verified using the NCBI database through the Blast tool (**Supplementary Table S5**). The 10 μl RT-qPCR reaction contained 1 μl of template cDNA (20 ng), 1 μl of forward primer, 1 μl of reverse primer, 4 μl of PowerUp™ SYBR™ Green Master Mix, and 3 μl of H_2O. PCR was run at an initial

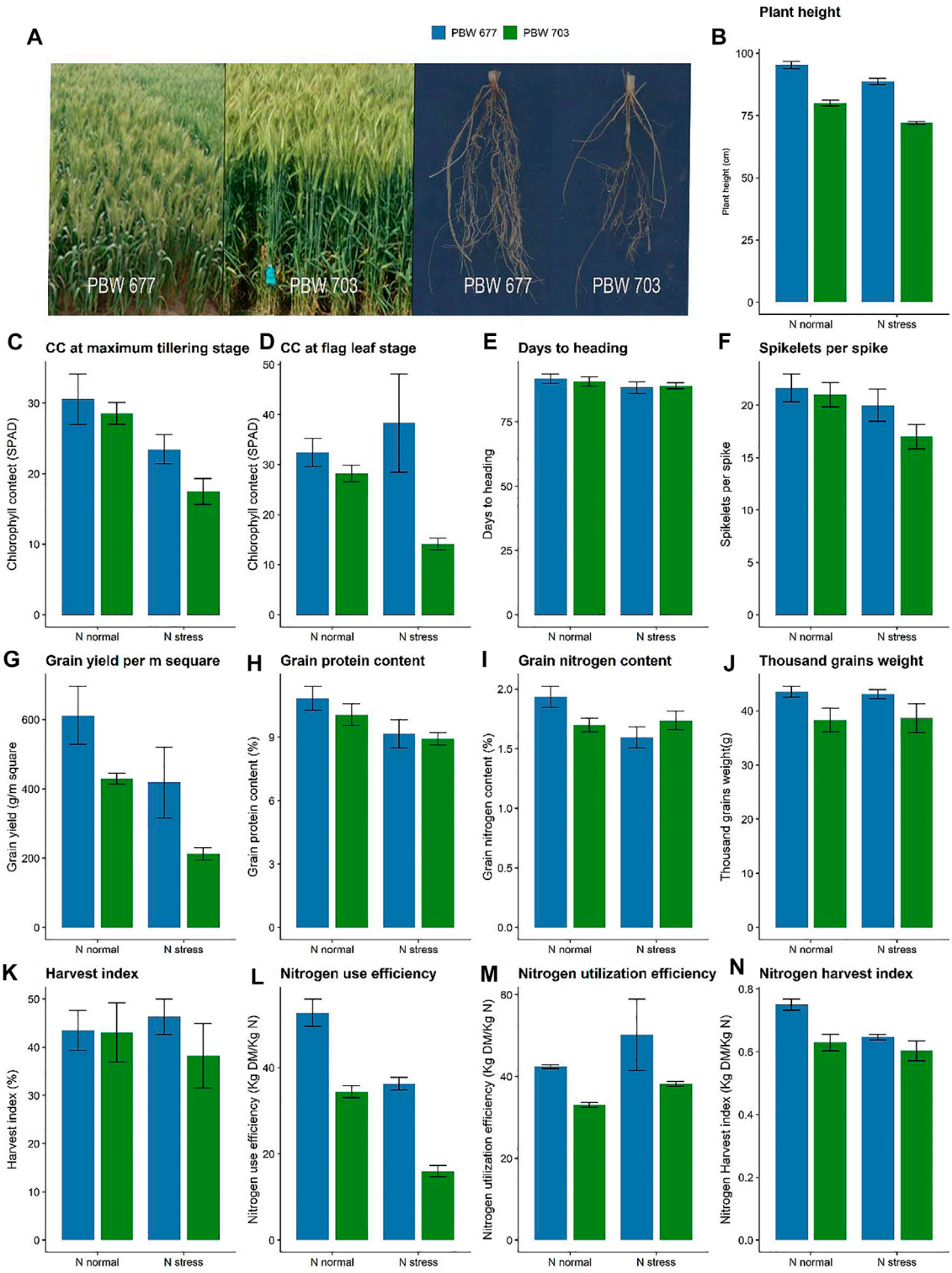


FIGURE 1 | Growth performance, yield and nitrogen use efficiency parameters of wheat cultivar PBW 677 and PBW 703 grown under N^+ and N stress conditions. **(A)** Images of plants shoot and roots **(B)** Plant height **(C)** Chlorophyll content and maximum tillering stage **(D)** Chlorophyll content at flag leaf stage **(E)** Days to heading **(F)** Spikelets per spike **(G)** Grain yield (g/m2) **(H)** Grain protein content **(I)** Grain nitrogen content **(J)** Thousand grains weight **(K)** Harvest index **(L)** Nitrogen use efficiency (NUE) **(M)** Nitrogen utilization efficiency (NUtE) **(N)** Nitrogen harvest index.

TABLE 1 | Quality of transcriptomics sequenced data of root and shoot tissues of PBW677 and PBW703 genotypes.

Sample name	Direction strand	N level	Raw read no.	Clean read no.	Clean reads	GC%
Root PBW677	Forward	N+	22,427,230	21,759,104	97.02092	56
	Reverse	N+	22,427,230	21,759,104	97.02092	56
Shoot PBW677	Forward	N+	24,320,113	23,719,546	97.53057	58
	Reverse	N+	24,320,113	23,719,546	97.53057	58
Root PBW677	Forward	N-	16,938,993	16,528,622	97.57736	55
	Reverse	N-	16,938,993	16,528,622	97.57736	55
Shoot PBW677	Forward	N-	19,982,451	19,321,385	96.69177	54
	Reverse	N-	19,982,451	19,321,385	96.69177	54
Root PBW703	Forward	N+	17,137,156	16,595,928	96.84179	54
	Reverse	N+	17,137,156	16,595,928	96.84179	54
Shoot PBW703	Forward	N+	15,424,751	15,122,142	98.03816	57
	Reverse	N+	15,424,751	15,122,142	98.03816	57
Root PBW703	Forward	N-	18,294,262	17,759,653	97.07772	57
	Reverse	N-	18,294,262	17,759,653	97.07772	57
Shoot PBW703	Forward	N-	15,692,834	14,946,663	95.24515	54
	Reverse	N-	15,692,834	14,946,663	95.24515	54

denaturation of 94°C for 3 min followed by 40 cycles of 94°C for 10 s, 60°C for 30 s, 72°C for 30 s, and a final extension at 72°C for 10 min to check the specificity of amplification. The housekeeping gene TaATP (ATP-dependent 26S proteasome regulatory) (Paolacci et al., 2009) was used as the endogenous control and all reactions were performed in triplicate. Relative gene expression was analyzed using the $2^{-\Delta\Delta Ct}$ method.

RESULTS

Plant Growth, Total Chlorophyll, and Total N Content Analysis

Both the cultivars PBW677 and PBW703 showed differences in growth and yield characteristics under contrasting N conditions (**Figure 1A**). Plant height showed no differences between the cultivars (based on cultivar) however it showed a significant difference at contrasting N levels (**Figure 1B**). The chlorophyll content showed a significant difference among the cultivars at different N levels (**Figures 1C,D**). Further no differences in the spikelets per spike were found under various N levels (**Figure 1F**). Grain and straw N were lower in PBW703 under N-stressed conditions in comparison to PBW677. The 1,000 grain weight was significantly higher for PBW677 under all the N fertilization regimes (**Figure 1J**) However, no significant difference was observed in spikelets per spike at different N levels. The nitrogen utilization and harvest index were higher in PBW677 vs. PBW703 supporting the fact that PBW677 is more efficient in acquiring, locating, and using N for grain development. Indeed, total tiller N levels as well as the NHI were significant higher in PBW677 compared to PBW703 (**Figures 1M,N**).

Transcriptomics Quality and Mapping Statistics

A total of 132 Gb 150-bp paired end (PE) reads were generated through Illumina NextSeq^{500} from eight samples (2 genotype × 2 treatment × 2 tissue) including N^- and N^+ treatment. On average, 18 million clean reads were obtained after trimming/clipping from each library with sizes ranging from 10 to 17 GB. The reads were of high quality at 98 and 95% and had a phred score of Q33. Moreover, the average GC% of each library was about 56% (**Table 1**). After mapping against the wheat reference genome (IWGSC release 1.0), varying contents of perfectly mapped reads in both roots (68.35%) and shoots/leaves (86.58%) were recorded in both genotypes (**Table 2**). Transcript profiles of the RNA-seq data were analyzed by calculating the read fragments per kilobase per million mapped reads (FPKM).

Differential Gene Expression in Response to Nitrogen Stress

Transcriptomics analysis of N-efficient (PBW677) and N-inefficient (PBW703) genotypes revealed a significant difference of adaptive response in terms of gene expression pattern when there was nitrogen stress. In PBW677, the total number of genes expressed in root tissues was 70,825, out of which 903 genes showed significant differential expression (748 genes upregulated and 155 downregulated) while in shoot tissues, a total of 61,895 genes were expressed, out of which 750 were significant differentially expressed genes (667 upregulated, 83 downregulated) (**Figure 2**; **Table 3**). In contrast, the total number of expressed genes in root tissues of PBW703 was 22,046, of which only 96 were differentially expressed genes (92 upregulated and four downregulated) and in shoot tissues, 49,121 genes were expressed, out of which 657 were differentially expressed genes (511 upregulated and 146 downregulated), as shown in **Figure 2**. **Supplementary Tables S1, S2** present the list of DEGs in root and shoot tissues of the cultivars PBW677 and PBW703, respectively. Volcano plots in **Figure 3** present the most significant upregulated and downregulated genes in root and shoot tissues of both the cultivars.

Genes Involved in Primary Nitrogen Metabolism

Genes involved in N absorption and assimilation were found to be differentially expressed in N stress-tolerant cultivar PBW677

TABLE 2 | Mapping statistics of the transcriptomics data.

Cultivar	Samples	Total mapped (%)	Perfect match (%)	Unmapped (%)
PBW677	Root N+	91.72	76.00	8.2800
	Shoot N+	98.87	83.90	1.1300
	Root N-	97.26	78.12	2.7400
	Shoot N-	99.27	82.61	0.7300
PBW703	Root N+	77.30	68.35	22.7000
	Shoot N+	96.96	84.26	3.0400
	Root N-	94.66	84.57	5.3400
	Shoot N-	97.83	86.58	2.1700

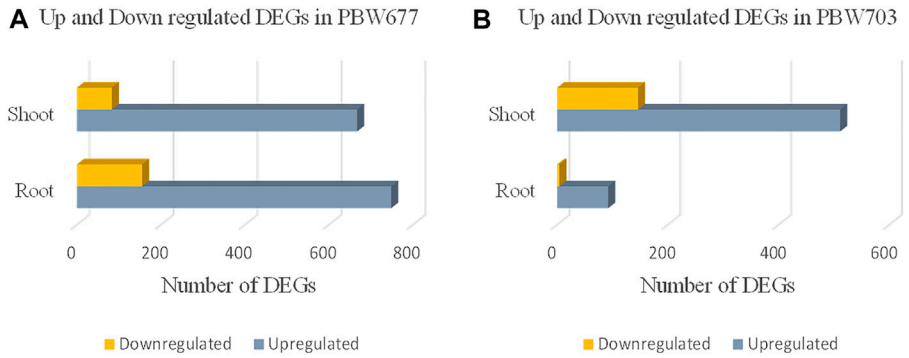


FIGURE 2 | Numbers of up and downregulated DEGs (differentially expressed genes) expressed in root and shoot tissues of **(A)** PBW677 and **(B)** PBW703.

compared to N stress-intolerant cultivar PBW703. Genes corresponding to the NRT1/PTR gene family were found to be upregulated in root tissues of PBW677 and downregulated in shoot tissues of PBW677 and PBW703. *Glutamate synthase* (GLT), amino acid transporter (AVT), *1-aminocyclopropane-1-carboxylate synthase* (ACO), IAA-amino acid hydrolase (ILL6), amino acid permease (AAP), and *Asparagine synthase* (ASNS) were upregulated in both the root and shoot tissues of PBW677 in comparison to PBW703 where these genes showed downregulation. However, some DEGs related to *serine--glyoxylate aminotransferase* (AGT), *glyoxylate aminotransferase* (GGAT1), and chloroplast/mitochondrial *Glutamine synthetase* (GLN) were upregulated in roots of PBW703 compared to PBW677. In addition, *Glutamate dehydrogenase* (GSH)-encoding genes were upregulated in shoot tissues of both the cultivars (**Figure 4A**, **Supplementary Table S3**).

Carbon Metabolism Genes

Alternate gene expression was observed for genes involved in carbon metabolism, especially involved in pathways like glycolysis, tricarboxylic acid (TCA) cycle, gluconeogenesis, photorespiration, photosynthesis, and trehalose metabolism as represented in **Figure 4B**. It has been observed that several DEGs related to glycolysis such as glyceraldehyde-3-phosphate dehydrogenase GAPA1, fructose-bisphosphate aldolase, and fructose-1,6-bisphosphatase were downregulated in both the tissues of PBW677. In contrast, in PBW703, these genes showed upregulation in root tissues. The same pattern was observed for gluconeogenesis metabolism-related genes, such as fructose-bisphosphate aldolase, sedoheptulose-1,7-bisphosphatase, and UDP-glucose 6-dehydrogenase that were found to be downregulated in roots of PBW677 but upregulated in PBW703. Similarly, the majority of DEGs involved in photosynthesis like ribulose bisphosphate carboxylase, chlorophyll a-b binding protein of LHCII, oxygen-evolving enhancer protein 2, photosystem II 5 kDa protein, photosystem I reaction center subunit III, cytochrome b6-f complex iron-sulfur subunit, and ferredoxin-NADP reductase were upregulated in both the tissues of PBW703 (**Figure 4B**, **Supplementary Table S3**). Other than this, the resulting DEGs that were involved in energy metabolic pathways other than carbon metabolism such as genes coding NADP-dependent malic enzyme and trehalose-phosphate phosphatase were found to be strongly upregulated under N stress in both tissues of PBW677 but were considerably downregulated in PBW703.

Transcription Factor (TF)-Encoding Genes

Transcription factors play a vital role as molecular switches controlling the expression of specific genes and play crucial roles in plant development, cell cycling, cell signaling, and stress responses. Due to N stress, various unique genes encoding transcription factors were differentially expressed in N stress-tolerant cultivar PBW677 corresponding to different TF families of MYB, WRKY, RING finger protein, Zinc finger protein, TIFY proteins, AP2/ERF, HSP, BTB/POZ domain,

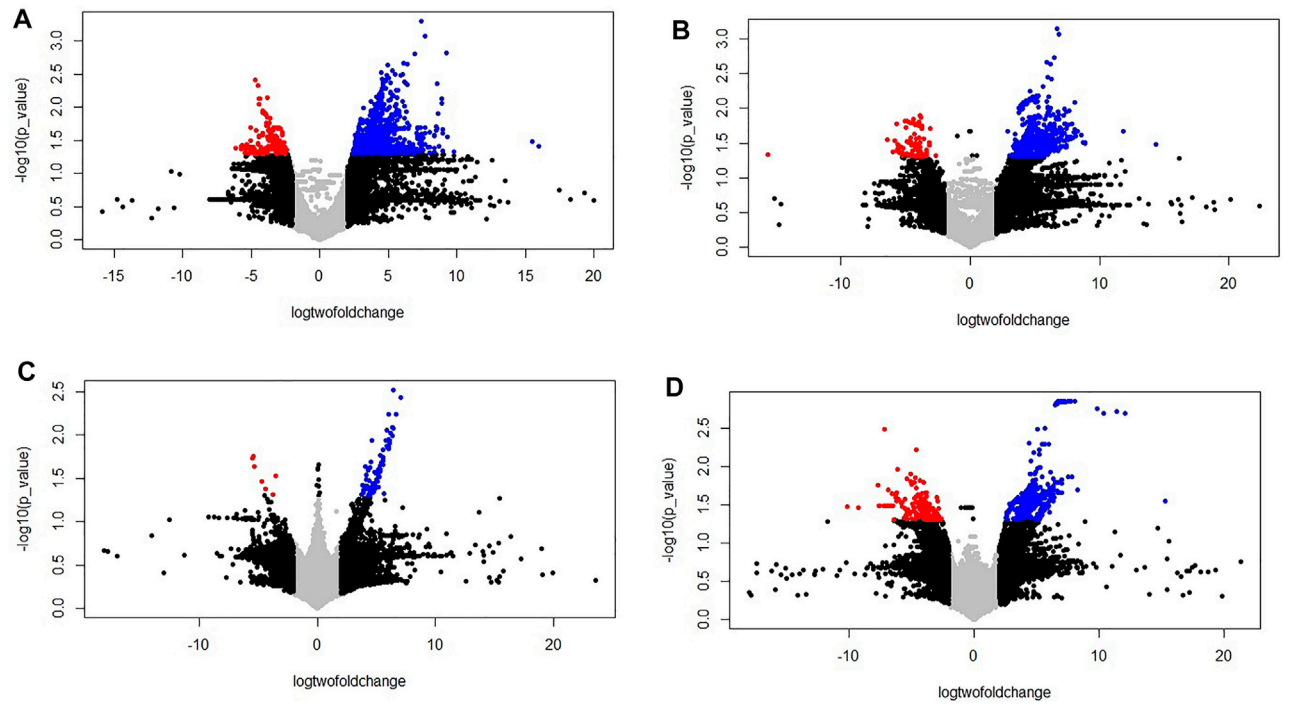


FIGURE 3 | Volcano plot of the differentially expressed genes (control vs. stressed) in the two genotypes PBW677 and PBW703 for each tissue investigated in this work. **(A)** Volcano plot of gene expression in the efficient genotype in roots and **(B)** shoot/leaves; **(C)** Volcano plot of gene expression in the inefficient genotype in roots and **(D)** shoots/leaves. *The two vertical dotted lines are twice the difference threshold (<-2 or >2), and the horizontal dotted line represents significance level of -log (*p*-value = 0.05) 1.30. Red dots indicate the downregulated genes, blue dots indicate the upregulated genes, black dots represent non-significantly differentially expressed genes surpassing the threshold absolute log2 fold change (≥2) or significantly expressed genes with lower absolute log2 fold change (≤2) than the threshold, and grey dots indicate non-significantly expressed genes with lower absolute log2 fold change (≤2) than the threshold.

TABLE 3 | Total number of differentially expressed genes (DEGs) and number of DEGs filtered on the basis log2FC≥|2| and statistical significance ($p < 0.05$) in shoot and root tissues of PBW677 and PBW703 between high (N+)/low (N-) levels.

Comparisons	Total genes	DEGs	Upregulated Genes	Downregulated Genes
PBW677 root (N+/N-)	70,825	903	748	155
PBW677 shoot (N+/N-)	61,895	750	667	83
PBW703 root (N+/N-)	22,046	96	92	4
PBW703 shoot (N+/N-)	49,121	657	511	146

NDR1/HIN1domain, bZIP, GATA, bHLH, DIVARICATA factors, Bowman-Birk type trypsin inhibitor isoform, BAG family molecular chaperone, dnaJ protein, ARF, and other families (**Supplementary Table S3**). In contrast, TF-encoding genes were identified in N stress-sensitive cultivar PBW703 and the number of TF-related DEGs in each family was less than those found in PBW677. We also observed the tissue-specific expression of several TF families. For instance, most of WRKY, HSP and AP2/ERF genes were extensively upregulated in roots only, however genes related to the BAG family molecular chaperone, ARF, DIVARICATA, GATA, and dnaJ protein were upregulated majorly in shoots of PBW703. Additionally, MYB and bHLH genes were expressed in both the tissues of PBW677 (Supplementary Fig. S1, **Supplementary Table S3**).

Transporter Encoding Genes

Genes encoding for various transporters responsible for uptake and transport of nutrients were found to be differentially expressed under N stress in both the cultivars. In general, it was observed that the genes belonging to different transporter families were shown to be unregulated in both the tissues of PBW677 in comparison to PBW703 where they were downregulated. In particular, calcium-binding protein transporter genes were upregulated in both the tissues of PBW677 in addition to the ABC transporter and SWEET transporter (Supplementary Fig. S2, **Supplementary Table S3**).

Protein Kinases Encoding Genes

In this study, various unique protein kinases were differentially expressed in both the tissues. The identified PK genes belonged to



FIGURE 4 | Heatmaps showing the expression patters of genes involved in between tissues and in both genotypes **(A)** primary nitrogen metabolism **(B)** carbon metabolism. *Colours indicate the differential gene expression in nitrogen stressed tissue; upregulated: green; downregulated: red; Do not have significant expression: white.

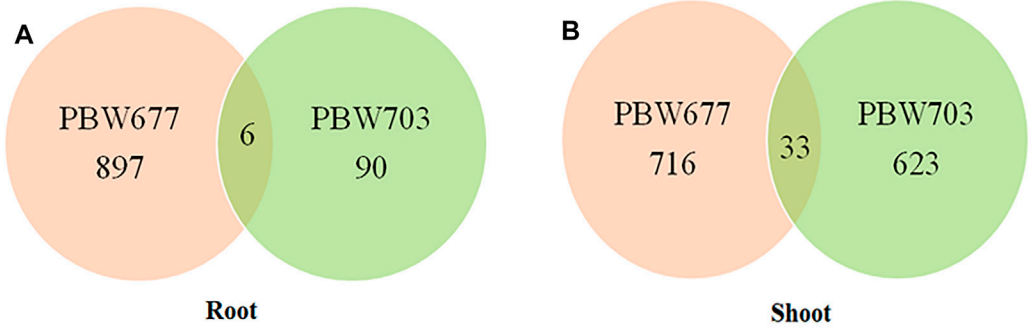


FIGURE 5 | Common DEGs (differentially expressed genes) in root and shoot tissues of **(A)** PBW677 and **(B)** PBW703.

classes CDK (cyclin-dependrnt kinase), CIPK (CBL-interacting protein kinase), cysteine-rich domains, F-box domains, LRR (leucine rich repeats), MAPK (mitogen activated protein kinase), serine threonine protein kinase, serine arginine protein kinase, and U-box domains (Supplementary Fig. S3, **Supplementary Table S3**). We found PKs belonging to MAPK, serine threonine protein kinase, CDK, CIPK, and cysteine-rich domains upregulated in roots of PBW677. Besides, most members of F-box domains and LRR were found upregulated only in shoots of PBW677.

Other Stress-Related Genes

In this study, we also observed upregulation of a number of genes related to detoxification and protection from oxidative stress, the majority of which were found in both tissues in N stress-tolerant cultivar PBW677. These mainly belong to classes of glutathione S-transferases (GSTs), cytochrome P450 (CYP450), E3 ubiquitin-protein ligase, peroxidase and phenylalanine ammonia-lyase, and aquaporin PIP1-1. Most of the genes related to glutathione S-transferases (GSTs), phenylalanine ammonia-lyase (PAL), and E3 ubiquitin-protein ligase were upregulated in both root and shoot tissues of PBW677 (**Figure 5**) (**Supplementary Figure S4**, **Supplementary Table S3**).

Common Genes Between Root and Shoot Tissues of PBW677 and PBW703

Among the total DEGs studied, common genes showing up and downregulation were detected, also the genes showing alternate expression (that is upregulation in one cultivar and downregulation in another cultivar) were identified. A list of common genes with their log fold change value and involved biological pathways is represented in **Supplementary Table S4**.

Enriched Gene Ontology Terms in Differentially Expressed Genes Under Nitrogen Stress

In cultivar PBW677, enrichment analysis using Fisher's exact test depicted the over and underrepresented genes expressed in response to nitrogen stress, as shown in **Figure 6**. Among the upregulated DEGs in root tissues of cultivar PBW677, 140 genes belonging to different GO categories (biological process, cellular component, and molecular function) were found to be overrepresented. Most of the overrepresented class of genes was found to be involved in sequence-specific DNA binding, response to temperature stimulus, abiotic stimulus, and transcription factors involved in transcription regulation and defense responses. Among the underrepresented class, only 11 genes were found, most of which belonged to the process of translation. The top 10 biological processes that were overrepresented among the upregulated DEGs expressed in root tissues of PBW677 are represented in **Table 4**. Among the downregulated DEGs in root tissues of PBW677, there were 64 genes which were overrepresented and belonged to the biological processes photosynthesis, carbon fixation, and photorespiration, as represented in **Table 5**. However, in the shoot tissues of PBW677, 82 upregulated genes were found to be overrepresented and six were underrepresented. In contrast, among the upregulated DEGs in root tissues of PBW703, 163 genes were found to be overrepresented and among them the biological processes photosynthesis, phosphorespiration, carbon fixation, and carbon metabolism, glutamine biosynthetic process, and nitrogen compound metabolic processes were highly enriched and overrepresented, as represented in **Figure 7**. In the shoot tissues of PBW703, among the upregulated DEGs (**Figure 7**), the overrepresented biological processes included nucleosome assembly, chromatin assembly, DNA packaging, chromatin remodeling, chromosome organization, response to inorganic substance, abiotic stimulus, lipid transport, sulfur compound metabolic process, glutathione metabolic process, nitrate transport, and response to nitrate with a number of 126 genes. **Table 6** represents the top 10 highly enriched biological processes overrepresented in upregulated DEGs in shoot tissues of PBW703. The underrepresented category included macromolecule modification, protein metabolic process, protein modification process, and nitrogen compound metabolic process. Protein folding and refolding and responses to heat and temperature

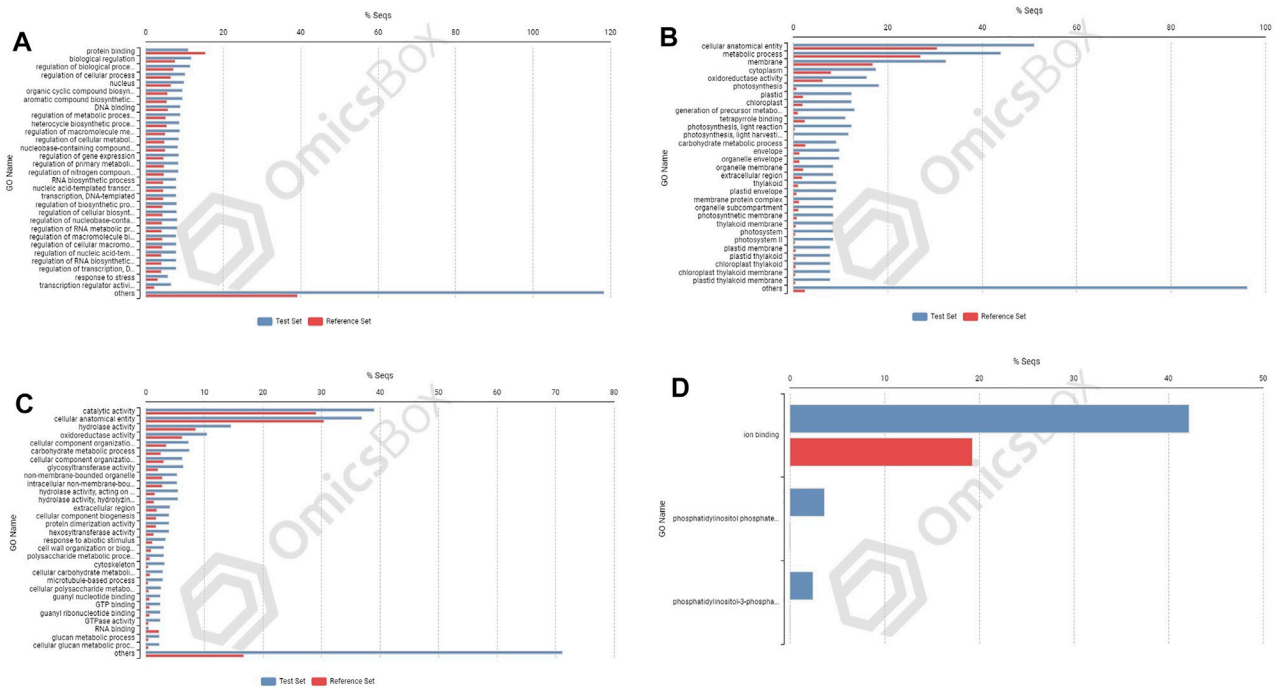


FIGURE 6 | Enrichment analysis showing genes present in roots of PBW677: **(A)** upregulated and **(B)** downregulated; and in shoot tissues: **(C)** upregulated and **(D)** downregulated.

TABLE 4 | List of top 10 biological processes with GO IDs that are overrepresented in upregulated DEGs in root tissues of PBW677.

S. No.	GO name	GO IDs	FDR
1	Response to temperature stimulus	GO:0009266	1.088379E-10
2	Response to abiotic stimulus	GO:0009628	5.055289E-8
3	Response to osmotic stress	GO:0006970	1.245807E-7
4	Response to heat	GO:0009408	1.245807E-7
5	Phosphoenol pyruvate family amino acid metabolic process	GO:1902222	1.684675E-7
6	Phenylalanine catabolic process	GO:0006559	1.684675E-7
7	Response to salt stress	GO:0009651	1.684675E-7
8	Aromatic amino acid family catabolic process	GO:0009074	3.512503E-6
9	Defense response to bacterium	GO:0042742	5.271174E-5
10	Secondary metabolite synthesis	GO:0044550	2.678367E-4

TABLE 5 | List of top 10 biological processes with GO IDs that are overrepresented in downregulated DEGs in root tissues of PBW677.

S. No.	GO name	GO IDs	FDR
1	Photosynthesis	GO:0015979	1.08140E-50
2	Protein chromophore linkage	GO:0018298	3.63567E-33
3	Photosynthesis, light harvesting	GO:0009765	1.86844E-32
4	Photorespiration	GO:0009853	1.464116E-28
5	Photosynthesis, light reaction	GO:0019684	2.379313E-23
6	Cellular metabolic compound salvage	GO:0043094	1.711553E-22
7	Photosynthesis dark reaction	GO:0019685	1.962272E-20
8	Reductive pentose phosphate cycle	GO:0019253	1.962272E-20
9	Response to abiotic stimulus	GO:0009628	8.835918E-12
10	Glutamine biosynthetic process	GO:0006542	2.613054E-3

stimulus were among the highly enriched biological processes among downregulated DEGs in shoot tissues of PBW703.

Validation of DEG Genes Using Real-Time Quantitative Polymerase Chain Reaction Analysis

To validate the expression data obtained by RNA-seq, we performed RT-qPCR analysis of the expression of the 17 selected DEGs (**Supplementary Table S5**) with both upregulated and downregulated expression in both of the tissues and in both genotypes. Results were found in agreement with the RNA-seq-based gene expression pattern with minor variations in the log2 FC values (**Figure 8**).

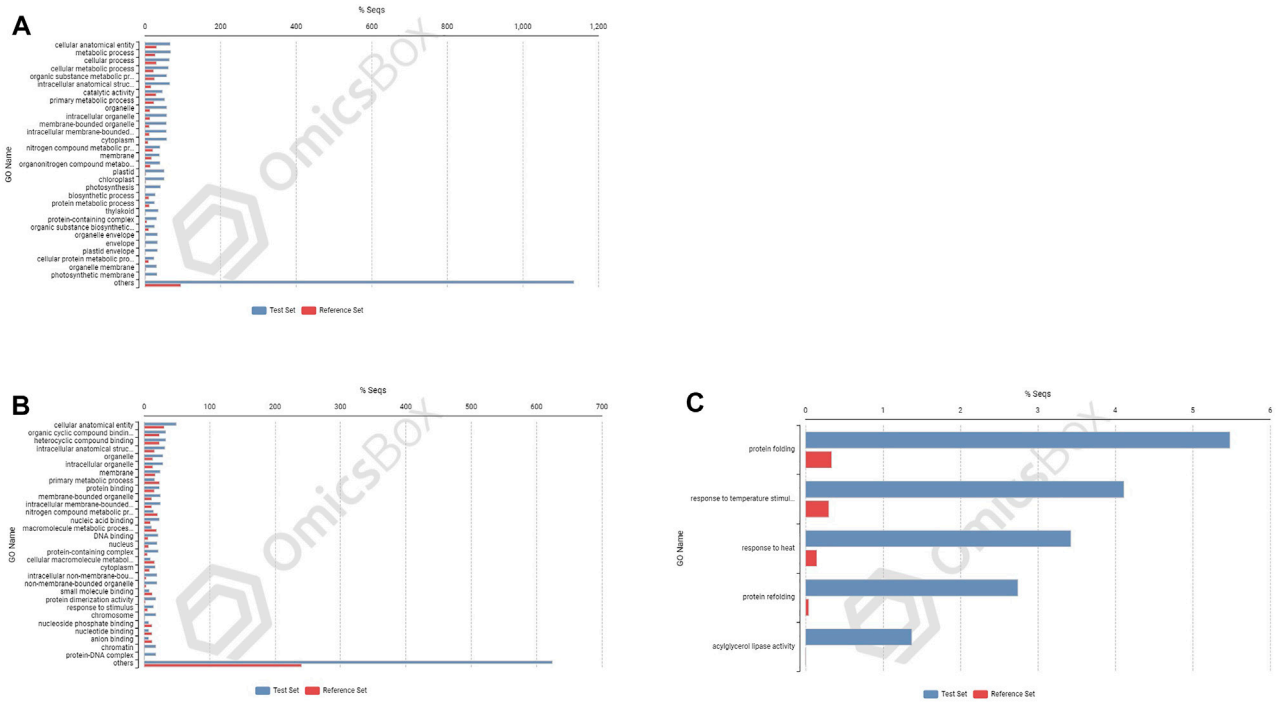


FIGURE 7 | Enrichment analysis showing genes present in roots of 703 **(A)** upregulated genes; and in shoot tissues: **(B)** upregulated and **(C)** downregulated.

TABLE 6 | List of top 10 biological processes with GO IDs that are overrepresented in upregulated DEGs in shoot tissues of PBW703.

S. No.	GO name	GO IDs	FDR
1	Nucleosome organization	GO:0034728	8.368916E-56
2	Response to stress	GO:0006950	9.207505E-18
3	Lipid transport	GO:0006869	2.889702E-7
4	Photorespiration	GO:0009853	1.464116E-28
5	Photosynthesis, light reaction	GO:0019684	2.379313E-23
6	Cellular metabolic compound salvage	GO:0043094	1.711553E-22
7	Photosynthesis dark reaction	GO:0019685	1.962272E-20
8	Reductive pentose phosphate cycle	GO:0019253	1.962272E-20
9	Response to abiotic stimulus	GO:0009628	8.835918E-12
10	Glutamine biosynthetic process	GO:0006542	2.613054E-3

DISCUSSION

To find out the underlying genes involved in NUE pathways, a transcriptomics study was conducted in control and N stress conditions in root and shoot tissues of two cultivars PBW677 and PBW703. Differential gene expression revealed the key genes involved in photosynthesis, starch and sugar metabolism, and nitrogen metabolism in both the cultivars. In limited N conditions, N primarily affected plant morphology, accelerating plant flowering time, senesces, decreasing the height, and lowering biomass and harvest index. The overall stresses of the plant were evident from the spikelets per spike while the average seed test grain weight was comparable. This indicated the plants concentrated their energy to produced viable seeds. The nitrogen use efficiency was affected by supplied N levels. The increase of NUE was affected by N regimes in winter and spring wheat (Cormier et al., 2013; Nehe et al., 2018). Under nitrogen stress, the number of upregulated as well as downregulated genes was higher in roots as compared to shoots. This is in accordance with the previous study conducted by Curci et al. (2017) on durum wheat as roots are primary organs responsive to N stress. Also, it was found that the number of DEGs was higher in PBW677 as compared to PBW703. This might be due to the fact that PBW677 responds actively in N-deficient conditions compared to PBW703 as in the case reported by Sultana et al., in 2020 where medium NUE Spitfire had more DEGs in comparison to low NUE cultivar Volcani. The 97 common genes between root and shoot tissues were mostly related to plant hormone signal transduction, MAPK signaling pathway, starch and sucrose metabolism, glutathione metabolism, and chlorophyll and porphyrin metabolism. The 33 DEGs common in shoot tissues were mainly involved in lipid transport, amino acid metabolism, photosynthesis, carotenoid biosynthesis, starch and sucrose metabolism, and nucleotide metabolism. Several previous studies indicated that when plants were subjected to nitrogen stress they exhibited a wide range of responses including molecular expression and biochemical processes involving various genes and pathways. There is a strong relation between nitrogen and carbon metabolism as has been shown in various studies (Duan et al., 2018; Naliwajski and Skłodowska, 2018). Gene enrichment analysis based on Gene Ontology is the method by which key biological processes responsible for stress adaptation can be revealed, for example an N deficiency study in durum wheat reported photosynthesis, carbon metabolism, and nitrogen metabolism as the highly enriched biological processes (Curci et al., 2017). Under N limitation, the metabolic process and oxidation–reduction process in

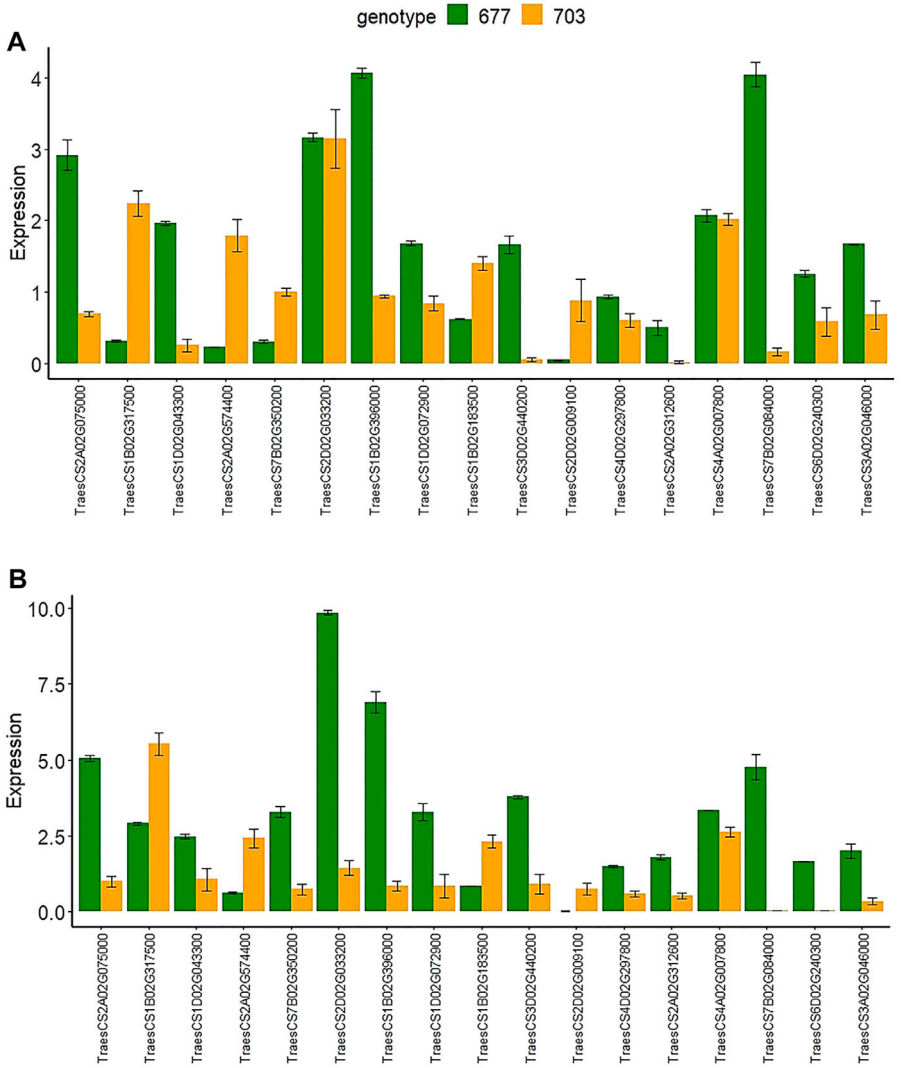


FIGURE 8 | The relative gene expression of 17 randomly selected genes examined by RT-qPCR in both genotypes PBW677 and PBW703 in **(A)** roots and **(B)** shoots.

wheat seedlings were enriched significantly (Wang et al., 2019). Under nitrogen stress, catalytic activity, binding, and the metabolic and cellular process were highly enriched GO terms in potato roots, shoots, and stolons (Tiwari et al., 2020). This study revealed that upregulated DEGs in root tissues of PBW677 were involved in pathways like DNA binding, response to abiotic stimulus, and defense response while the downregulated genes were involved in biological processes like photosynthesis, photorespiration, and carbon fixation. While in shoot tissues, the highly enriched biological processes were carbohydrate metabolic process, cell wall biogenesis, and cellular polysaccharide metabolic process. In contrast to this, the upregulated DEGs in root tissues of PBW703 belonged to enriched biological processes like photosynthesis, carbon metabolism, glutamine biosynthetic process, and nitrogen compound metabolic processes while in shoot tissues, the enriched pathways were involved in nucleosome assembly, sulfur compound metabolic process, glutathione metabolic process, nitrate transport, etc. The reason behind the upregulated transcripts belonging to the pathways related to nucleosome assembly and chromatin packaging would be the mechanisms involved towards the stability of the epigenome under N stress in PBW703 (Sudan et al., 2018). N metabolism-related genes were identified in the annotated transcriptome data showing both up as well as downregulation. For example expression of low affinity nitrate transporter NRT1 was found to be upregulated in root tissues of PBW677 and downregulated in shoot tissues of PBW677 and PBW703 under N stress which usually increases the expression of transport systems for nitrate and ammonium (Crawford and Glass, 1998). In bread wheat under N stress, a high level of expression of the NRT1/PTR family was found (Sultana et al., 2020). Also, it was shown that in the low N-treated plants, downregulation of the NRT1/PTR family was observed which is related to low grain yield and grain protein content (Léran et al., 2015). Other genes playing a role in N absorption, assimilation, and remobilization like glutamate synthase (GLT), amino acid transporter (AAP), and asparagine synthase were

upregulated in both the root and shoot tissues of PBW677 where as these genes were downregulated in PBW703. The upregulated expression of these genes can be related to the high NUE of PBW677 which is better adapted to N stress. Aminotransferases like serine glyoxylate aminotransferase (AGT), glyoxylate aminotransferase (GGAT1), and chloroplast glutamine synthetase (GLN) were upregulated in roots of PBW703 which is supported by previous reports conducted in sorghum where a transcriptomics study found more N assimilator genes in low NUE cultivars (Singh et al., 1973).

Transcription factors (TFs) are important key regulators that play a significant role in adaptation under environmental stresses (Shahzad et al., 2020). In our study we found a number of transcription factor families up and downregulated under N stress. The number of expressed TFs was higher in PBW677 than PBW703 which might be due to the fact that effective regulatory mechanisms exist under N stress in high NUE cultivar PBW677. Several TF families that were identified in response to N stress were HSF, MYB, WRKY, and ZINC finger. Transcriptomics analysis in durum wheat identified most of the WRKY family in response to N chronic stress in roots (Curci et al., 2017). This family is one of the largest families of plant-specific transcription factors that plays important roles in various abiotic stresses (drought, saline, alkali, temperature, and ultraviolet radiation (Li et al., 2020). Several MYB and bHLH TFs were reported to be involved in regulation of target genes under plant stress (Pireyre and Burow, 2015). In rice, expression of OSMYB48-1 was reported to be upregulated in tolerance toward abiotic stress probably via the regulation of stress-mediated ABA biosynthesis (Xiong et al., 2014).

Zinc finger proteins (ZEPs) enhance tolerance under abiotic stress (Zang et al., 2016). A high level of expression was observed for protein kinases in high NUE PBW677 (either up or downregulated). Protein kinases are known to play an important role in signal transduction and are regarded to be the central regulatory components to major environmental stresses such as drought, heat, cold, and pathogen attack (Wang et al., 2020). PKs play a role in hormone signaling, cell cycle growth, and nutrient signaling as well.

Nitrogen of the plant is invested in making protein and chlorophyll content of photosynthetic apparatus thus carbon metabolism and nitrogen metabolism are interconnected and carbon metabolism is dependent on nitrogen assimilation in plants (Foyer et al., 2001). In our study we found that a number of carbon metabolism genes were downregulated like genes involved in glycolysis, TCA cycle, and gluconeogenesis. This is similar to previous reports in which it was shown that N stress negatively effects plant carbohydrate metabolism (Rufty et al., 1988; Sultana et al., 2020). We found other stress-related genes to be highly expressed under N stress in high NUE cultivar PBW677 like GSTs, cytochrome P 450, E3 protein ligase, etc. Glutathione S transferases prevent cells from oxidative damage by quenching reactive molecules with the addition of glutathione (GSH) (Kumar and Trivedi, 2018).

CONCLUSION

RNA sequencing of two contrasting cultivars PBW677 and PBW703 for NUE helped in revealing candidate genes which could be utilized in future breeding programs focused on reducing the use of nitrogen fertilizers. The majority of the genes belonged to transcription factor families; protein kinases and stress-related nitrogen metabolism were found to be highly expressive in wheat cultivar PBW677 which might explain its behavior under N stress. There was a difference in highly enriched pathways responsive to nitrogen stress in the contrasting cultivars which might be the cause for their different behavior towards N stress. The 13 common genes showed alternative expression patterns in PBW677 and PBW703 and could be the potential candidates for high NUE-targeted breeding.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The RNA-seq data have been deposited in the NCBI. Submission details: NCBI Sequence Read Archive (SRA) submission: PRJNA780342.

AUTHOR CONTRIBUTIONS

SatK and AS conceived and designed the experiments, wrote the manuscript, and provided overall guidance; SK performed lab experiments, data analysis, and manuscript writing; MS and SS conducted field experiments and manuscript writing; SJ performed data analysis and drafted the manuscript; and AK took part in data analysis and drafting the manuscript.

ACKNOWLEDGMENTS

The authors acknowledge grant-in-aid support of the Department of Biotechnology (DBT), Ministry of Science and Technology, Government of India and co-funding by the Newton Bhabha Fund Project "Indo-UK Centre for improvement of nitrogen use efficiency -'Exploiting phenology and adaptation to improve nitrogen use efficiency in wheat," Grant number: BT/IN/UK-VNC/43/KV/2015-16.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.853910/full#supplementary-material

REFERENCES

Andrews, S., Krueger, F., Seconds-Pichon, A., Biggins, F., and Wingett, S. (2015). FastQC. A Quality Control Tool for High Throughput Sequence Data. Babraham Bioinformatics. Babraham Inst. Available at: https://www.bibsonomy.org/bibtex/f230a919c34360709aa298734d63dca3 (Accessed January 3, 2022).

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

Cormier, F., Faure, S., Dubreuil, P., Heumez, E., Beauchêne, K., Lafarge, S., et al. (2013). A Multi-Environmental Study of Recent Breeding Progress on Nitrogen Use Efficiency in Wheat (*Triticum A* L.). *Theor. Appl. Genet.* 126, 3035–3048. doi:10.1007/s00122-013-2191-9

Crawford, N. M., and Glass, A. D. M. (1998). Molecular and Physiological Aspects of Nitrate Uptake in Plants. *Trends Plant Sci.* 3, 389–395. doi:10.1016/S1360-1385(98)01311-9

Curci, P. L., Aiese Cigliano, R., Zuluaga, D. L., Janni, M., Sanseverino, W., and Sonnante, G. (2017). Transcriptomic Response of Durum Wheat to Nitrogen Starvation. *Sci. Rep.* 7, 1–14. doi:10.1038/s41598-017-01377-0

Duan, W., Wang, Q., Zhang, H., Xie, B., Li, A., Hou, F., et al. (2018). Comparative Study on Carbon-Nitrogen Metabolism and Endogenous Hormone Contents in normal and Overgrown Sweetpotato. *South Afr. J. Bot.* 115, 199–207. doi:10.1016/J.SAJB.2017.11.016

Eastin, E. F. (2008). Use of an AutoAnalyzer for Total Nitrogen Determination in Plants. *Commun. Soil Sci. Plant Anal.* 9, 107–113. doi:10.1080/00103627809366792

Foyer, C. H., Ferrario-Méry, S., and Noctor, G. (2001). "Interactions between Carbon and Nitrogen Metabolism," in *Plant Nitrogen*. Editors P. J. Lea and J. F. Morot-Gaudry (Berlin, Heidelberg: Springer), 237–254. doi:10.1007/978-3-662-04064-5_9

Gelli, M., Duo, Y., Konda, A. R., Zhang, C., Holding, D., and Dweikat, I. (2014). Identification of Differentially Expressed Genes between Sorghum Genotypes with Contrasting Nitrogen Stress Tolerance by Genome-Wide Transcriptional Profiling. *BMC Genomics* 15, 179. doi:10.1186/1471-2164-15-179/TABLES/6

Good, A. G., Shrawat, A. K., and Muench, D. G. (2004). Can Less Yield More? Is Reducing Nutrient Input into the Environment Compatible with Maintaining Crop Production? *Trends Plant Sci.* 9, 597–605. doi:10.1016/J.TPLANTS.2004.10.008

Hao, Q. N., Zhou, X. A., Sha, A. H., Wang, C., Zhou, R., and Chen, S. L. (2011). Identification of Genes Associated with Nitrogen-Use Efficiency by Genome-wide Transcriptional Analysis of Two Soybean Genotypes. *BMC Genomics* 12, 525. doi:10.1186/1471-2164-12-525

Hitz, K., Clark, A. J., and Van Sanford, D. A. (2017). Identifying Nitrogen-Use Efficient Soft Red winter Wheat Lines in High and Low Nitrogen Environments. *Field Crops Res.* 200, 1–9. doi:10.1016/j.fcr.2016.10.001

Kant, S., Bi, Y.-M., and Rothstein, S. J. (2011). Understanding Plant Response to Nitrogen Limitation for the Improvement of Crop Nitrogen Use Efficiency. *J. Exp. Bot.* 62, 1499–1509. doi:10.1093/JXB/ERQ297

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/BIBTEX

Léran, S., Edel, K. H., Pervent, M., Hashimoto, K., Corratgé-Faillie, C., Offenborn, J. N., et al. (2015). Nitrate Sensing and Uptake in Arabidopsis Are Enhanced by ABI2, a Phosphatase Inactivated by the Stress Hormone Abscisic Acid. *Sci. Signal.* 8 (375), ra43. doi:10.1126/SCISIGNAL.AAA4829

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., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence Alignment/Map Format and SAMtools. *Bioinformatics* 25, 2078–2079. doi:10.1093/bioinformatics/btp352

Li, W., Pang, S., Lu, Z., and Jin, B. (2020). Function and Mechanism of WRKY Transcription Factors in Abiotic Stress Responses of Plants. *Plants (Basel)* 9, 1–15. doi:10.3390/plants9111515

Masclaux-Daubresse, C., Daniel-Vedele, F., Dechorgnat, J., Chardon, F., Gaufichon, L., and Suzuki, A. (2010). Nitrogen Uptake, Assimilation and Remobilization in Plants: Challenges for Sustainable and Productive Agriculture. *Ann. Bot.* 105, 1141–1157. doi:10.1093/AOB/MCQ028

Meng, X., Liang, Z., Dai, X., Zhang, Y., Mahboub, S., Ngu, D. W., et al. (2021). Predicting Transcriptional Responses to Cold Stress across Plant Species. *Proc. Natl. Acad. Sci. U. S. A.* 118 (10), e2026330118. doi:10.1073/PNAS.2026330118

Moll, R. H., Kamprath, E. J., and Jackson, W. A. (1982). Analysis and Interpretation of Factors Which Contribute to Efficiency of Nitrogen Utilization1. *Agron. J.* 74, 562–564. doi:10.2134/AGRONJ1982.00021962007400030037X

Naliwajski, M. R., and Skłodowska, M. (2018). The Relationship between Carbon and Nitrogen Metabolism in Cucumber Leaves Acclimated to Salt Stress. *PeerJ* 6, e6043. doi:10.7717/peerj.6043

Nehe, A. S., Misra, S., Murchie, E. H., Chinnathambi, K., and Foulkes, M. J. (2018). Genetic Variation in N-Use Efficiency and Associated Traits in Indian Wheat Cultivars. *Field Crops Res.* 225, 152–162. doi:10.1016/j.fcr.2018.06.002

Paolacci, A. R., Tanzarella, O. A., Porceddu, E., and Ciaffi, M. (2009). Identification and Validation of Reference Genes for Quantitative RT-PCR Normalization in Wheat. *BMC Mol. Biol* 10, 1–27. doi:10.1186/1471-2199-10-11

Pireyre, M., and Burow, M. (2015). Regulation of MYB and bHLH Transcription Factors: A Glance at the Protein Level. *Mol. Plant* 8, 378–388. doi:10.1016/j.molp.2014.11.022

Rufty, T. W., Jr., Huber, S. C., and Volk, R. J. (1988). Alterations in Leaf Carbohydrate Metabolism in Response to Nitrogen Stress. *Plant Physiol.* 88, 725–730. doi:10.1104/PP.88.3.725

Ruuska, S. A., Lewis, D. C., Kennedy, G., Furbank, R. T., Jenkins, C. L. D., and Tabe, L. M. (2008). Large Scale Transcriptome Analysis of the Effects of Nitrogen Nutrition on Accumulation of Stem Carbohydrate Reserves in Reproductive Stage Wheat. *Plant Mol. Biol.* 66, 15–32. doi:10.1007/S11103-007-9249-5/FIGURES/7

Shahzad, R., Jamil, S., Ahmad, S., Nisar, A., Amina, Z., Saleem, S., et al. (2020). Harnessing the Potential of Plant Transcription Factors in Developing Climate-Smart Crops: Future Prospects, Challenges, and Opportunities. Preprints. doi:10.20944/PREPRINTS202010.0532.V1

Shamshad, M. (2019). *Genetic and Physiological basis of Nitrogen uptake and utilization in wheat (Triticum aestivum L.)*. PhD. thesis. Ludhiana, India: Punjab Agricultural University. Available at: https://krishikosh.egranth.ac.in/handle/1/5810137744

Singh, T., Paleg, I., and Aspinall, D. (1973). Stress Metabolism I. Nitrogen Metabolism and Growth in the Barley Plant during Water Stress. *Aust. Jnl. Bio. Sci.* 26, 45–56. doi:10.1071/BI9730045

Sudan, J., Raina, M., and Singh, R. (2018). Plant Epigenetic Mechanisms: Role in Abiotic Stress and Their Generational Heritability. *3 Biotech.* 8 (3), 172. doi:10.1007/S13205-018-1202-6

Sultana, N., Islam, S., Juhasz, A., and Ma, W. (2021). Wheat Leaf Senescence and its Regulatory Gene Network. *Crop J.* 9, 703–717. doi:10.1016/J.CJ.2021.01.004

Sultana, N., Islam, S., Juhasz, A., Yang, R., She, M., Alhabbar, Z., et al. (2020). Transcriptomic Study for Identification of Major Nitrogen Stress Responsive Genes in Australian Bread Wheat Cultivars. *Front. Genet.* 11, 583785. doi:10.3389/FGENE.2020.583785

Tiong, J., Sharma, N., Sampath, R., MacKenzie, N., Watanabe, S., Metot, C., et al. (2021). Improving Nitrogen Use Efficiency through Overexpression of Alanine Aminotransferase in Rice, Wheat, and Barley. *Front. Plant Sci.* 12, 29. doi:10.3389/FPLS.2021.628521/BIBTEX

Tiwari, J. K., Buckseth, T., Zinta, R., Saraswati, A., Singh, R. K., Rawat, S., et al. (2020). Transcriptome Analysis of Potato Shoots, Roots and Stolons under Nitrogen Stress. *Sci. Rep.* 10, 1152. doi:10.1038/s41598-020-58167-4

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

Vijayalakshmi, P., Vishnukiran, T., Ramana Kumari, B., Srikanth, B., Subhakar Rao, I., Swamy, K. N., et al. (2015). Biochemical and Physiological Characterization for Nitrogen Use Efficiency in Aromatic Rice Genotypes. *Field Crops Res.* 179, 132–143. doi:10.1016/j.fcr.2015.04.012

Wang, J., Song, K., Sun, L., Qin, Q., Sun, Y., Pan, J., et al. (2019). Morphological and Transcriptome Analysis of Wheat Seedlings Response to Low Nitrogen Stress. *Plants* 8, 98. doi:10.3390/PLANTS8040098

Wang, P., Hsu, C.-C., Du, Y., Zhu, P., Zhao, C., Fu, X., et al. (2020). Mapping Proteome-Wide Targets of Protein Kinases in Plant Stress Responses. *Proc. Natl. Acad. Sci. USA* 117, 3270–3280. doi:10.1073/pnas.1919901117

Xiong, H., Li, J., Liu, P., Duan, J., Zhao, Y., Guo, X., et al. (2014). Overexpression of OsMYB48-1, a Novel MYB-Related Transcription Factor, Enhances Drought

and Salinity Tolerance in rice. *PLoS One* 9, e92913. doi:10.1371/journal.pone.0092913

Xu, G., Fan, X., and Miller, A. J. (2012). Plant Nitrogen Assimilation and Use Efficiency. *Annu. Rev. Plant Biol.* 63, 153–182. doi:10.1146/ANNUREV-ARPLANT-042811-105532

Xu, Z., Yu, Z., and Zhao, J. (2013). Theory and Application for the Promotion of Wheat Production in China: Past, Present and Future. *J. Sci. Food Agric.* 93, 2339–2350. doi:10.1002/JSFA.6098

Zamboni, A., Astolfi, S., Zuchi, S., Pii, Y., Guardini, K., Tononi, P., et al. (2014). Nitrate Induction Triggers Different Transcriptional Changes in a High and a Low Nitrogen Use Efficiency maize Inbred Line. *J. Integr. Plant Biol.* 56, 1080–1094. doi:10.1111/JIPB.12214

Zang, D., Li, H., Xu, H., Zhang, W., Zhang, Y., Shi, X., et al. (2016). An Arabidopsis Zinc finger Protein Increases Abiotic Stress Tolerance by Regulating Sodium and Potassium Homeostasis, Reactive Oxygen Species Scavenging and Osmotic Potential. *Front. Plant Sci.* 7, 1–14. doi:10.3389/fpls.2016.01272

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.

Copyright © 2022 Kaur, Shamshad, Jindal, Kaur, Singh, sharma and Kaur. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Genetics

REVIEW
published: 12 April 2022
doi: 10.3389/fgene.2022.811732

Uncovering the Epigenetic Marks Involved in Mediating Salt Stress Tolerance in Plants

*Garima Singroha, Satish Kumar, Om Prakash Gupta, Gyanandra Pratap Singh and Pradeep Sharma**

Indian Council of Agricultural Research-Indian Institute of Wheat and Barley Research, Karnal, India

The toxic effects of salinity on agricultural productivity necessitate development of salt stress tolerance in food crops in order to meet the escalating demands. Plants use sophisticated epigenetic systems to fine-tune their responses to environmental cues. Epigenetics is the study of heritable, covalent modifications of DNA and histone proteins that regulate gene expression without altering the underlying nucleotide sequence and consequently modify the phenotype. Epigenetic processes such as covalent changes in DNA, histone modification, histone variants, and certain non-coding RNAs (ncRNA) influence chromatin architecture to regulate its accessibility to the transcriptional machinery. Under salt stress conditions, there is a high frequency of hypermethylation at promoter located CpG sites. Salt stress results in the accumulation of active histones marks like H3K9K14Ac and H3K4me3 and the downfall of repressive histone marks such as H3K9me2 and H3K27me3 on salt-tolerance genes. Similarly, the H2A.Z variant of H2A histone is reported to be down regulated under salt stress conditions. A thorough understanding of the plasticity provided by epigenetic regulation enables a modern approach to genetic modification of salt-resistant cultivars. In this review, we summarize recent developments in understanding the epigenetic mechanisms, particularly those that may play a governing role in the designing of climate smart crops in response to salt stress.

Keywords: epigenetic modifications, DNA methylation, salt stress, RNA directed DNA methylatio, histone acetylation

OPEN ACCESS

Edited by:
Abhishek Bohra, Indian Institute of Pulses Research (ICAR), India

Reviewed by:
Chenjiang You, Fudan University, China
Mukesh Choudhary, ICAR-Indian Institute of Maize Research, India

****Correspondence:***
Pradeep Sharma
Pradeep.Sharma@icar.gov.in

Specialty section:
This article was submitted to Plant Genomics, a section of the journal Frontiers in Genetics

Received: *09 November 2021*
Accepted: *15 March 2022*
Published: *12 April 2022*

Citation:
Singroha G, Kumar S, Gupta OP, Singh GP and Sharma P (2022) Uncovering the Epigenetic Marks Involved in Mediating Salt Stress Tolerance in Plants. Front. Genet. 13:811732. doi: 10.3389/fgene.2022.811732

1 INTRODUCTION

Unpredictable climatic conditions render plants suffer from an array of abiotic stress factors. Soil salinity is a key stressor impeding crop productivity and affects an area of more than one billion hectares all over the world and these numbers are constantly growing (FAO and ITPS, 2015).

At molecular level, plants respond to an environmental stress by implementing dynamic changes in gene expression and reprogramming the plant physiology (Lämke and Bäurle, 2017; Luo and He, 2020). In the last two decades, transcriptional responses have been explored to uncover the specific signaling pathways involved in salt stress responses and to distinguish the individual regulatory proteins and their targets. The chromatin architecture in eukaryotes is very dynamic and is modified in response to environmental stimulus. The transcriptional regulation of gene expression can be better apprehended by unveiling the underlying structural context. The regulation of gene expression by modulating chromatin architecture has been termed as epigenetics and is an essential mechanism

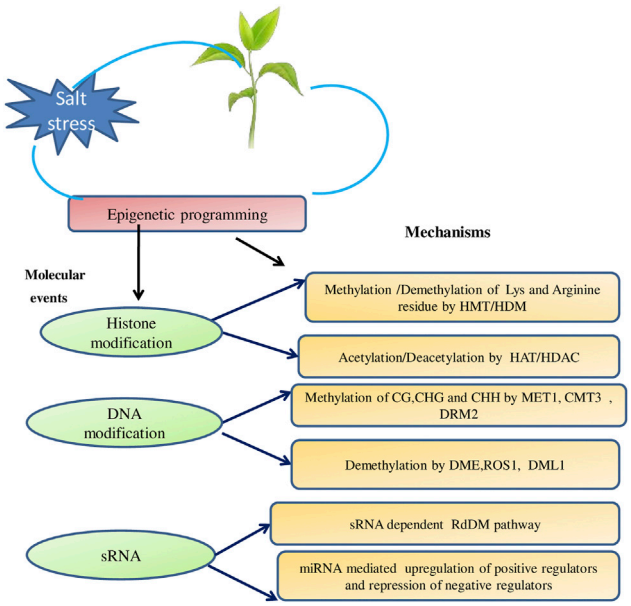


FIGURE 1 | Schematic representation of epigenetics re-programming in plants exposed to salinity stress at three level, i.e., DNA modifications, Histone modifications and small RNAs.

for biological phenomena, including developmental programming, expression of genes, genome stability and small RNA-mediated regulation, and so forth (Chang et al., 2020). Epigenetic changes are changes in the DNA backbone independent of changes in its sequence and are decisive for plant life cycle (Duan et al., 2018). Important Epigenetic components are histone modification, histone variants, DNA methylation, and some noncoding RNAs (ncRNA) (**Figure 1**). These modifications demonstrate an overall impact on chromatin organization and sway its availability to the transcriptional machinery and hence act as a benchmark in regulating gene expression (Crisp et al., 2016; Saroha et al., 2017; Duan et al., 2018; Singroha and Sharma, 2019).

Methylation of DNA is the most extensively investigated epigenetic modification and includes the insertion of a methyl group at 5′ position on cytosine bases (called 5-methylcytosine or 5mC) or 6′ position of the adenine bases (called N6-methyladenine or 6 mA) (Liang et al., 2018; Zhang et al., 2018). DNA methylation is associated with numerous processes vital for plant growth and acclimatization to stress (Zhang et al., 2018). Several authors demonstrated a perturbation in methylation patterns and thus altered gene expression under saline conditions (Li et al., 2014; Wang B et al., 2015; Konate et al., 2018).

In order to counter unfavorable environmental conditions, histone protein sustain some modifications at their N′ termini to modulate the gene expression for better survival. It has now been documented that histone acetylation and methylation are vital epigenetic marks in fine tuning gene expression under unfavorable conditions (Xie et al., 2015a). H3K4me3 and H3K27me3 are generally viewed as a pair of the opponent markers for enhancing or diminishing the expression of marked genes all the way through environmental changes (Zhang et al., 2009). Apart from histone modification and methylation of DNA, histone variants also impact chromatin dynamics. On account of differences in amino acid sequence and structure each histone protein is defined by several variants. Different histone variants display varying affinities with DNA and other histone protein, which imparts them the capacity to modify the state of chromatin compaction and attract regulatory protein complexes. These epigenetic changes together govern the accessibility of DNA to transcriptional machinery and consequently influence gene expression under diverse stress conditions. The modern approach to genetic improvement of crops for environmental stress resilience seek to enhance stress tolerance and involves comprehensive knowledge of its interconnections and flexibility in the expression of epigenetic regulation (Rodríguez López and Wilkinson, 2015). Therefore, epigenetic determinants have attracted plant breeder's interest since they are determinant of trans-generational phenotypic plasticity in plants under grueling environments. Hence, epigenetics play a very significant role in comprehending the complex mechanisms underpinning physical stress response and adaptability (Varotto et al., 2020). In this study, we have analyzed the current knowledge that connects the epigenetic and the transcriptional responses of plants under saline conditions, which might be essential for improving agricultural adaptability and reproducing climate smart crops.

2 DNA METHYLATION

Plant DNA methylation is referred to as N6-methyladenine (6 mA) or 5-methylcytosine (5 mC) (Zhang et al., 2018). However, in context of salinity 6 mA still remains enfolded and most of the reports acknowledge 5mC under salt stress. The 5mC is usually seen in all three sets of plant sequences: symmetrical CG and CHG together with asymmetrical CHH (where H = A, T or C) (Kumar et al., 2018). The methylation at different sequence contexts is catalyzed by sequence-specific methylases viz. CG methylation depends on MET1 (methyltransferase 1), CHG methylation requires DRM2 (domains rearranged methyltransferase 2) or CMT2 (chromomethylase 2) and CMT3 (chromomethylase 3) are vital for CHH methylation (Duan et al., 2018). The base excision pathway is one of the DNA repair pathways that can undo methylation state of a DNA and involves participation of DML2 (demeter-like 2), dme (demeter), ros1 (reprssor of transcriptional silencing 1) and five methylcytosine DNA glycosylase/DNA demethylase enzyme (Zhang et al., 2018; Liu and Lang, 2019).

Methylation of the promoter region has been generally associated with transcriptional repression while gene methylation activates transcription in *Arabidopsis thaliana*. Salt stress has been shown to affect methylation in different ways in different plant species and modify gene expression (Kinoshita and Seki, 2014; Voigt et al., 2015; Banerjee et al., 2017). Konate et al. (2018) observed increased DNA methylation

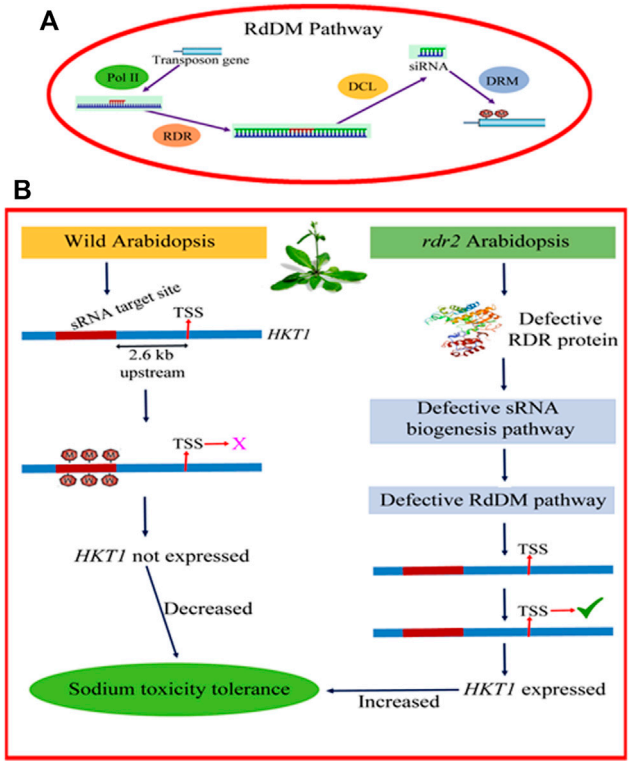


FIGURE 2 | (A) RDR-dependent RdDM pathway. This pathway provides a means to establish RNA–directed DNA methylation (RdDM) and eventually ensure stable transcriptional gene silencing (TGS) **(B)** Role of RDR-dependent RdDM pathway in regulating the methylation landscape of *HKT1* gene in Arabidopsis.

in *Hordeum vulgaris* leaves as compared to roots and claimed that salt-induced methylation is organ-specific. Chen et al. (2019) observed that 61.2% of CGs, 39.7% of CHG, and 3.2% of CHHs were methylated under salt stress in *Glycine max* roots which represent significantly lower methylation compared to control.

More often, salt-induced DNA methylation occurs inside or in close proximity to already identified stress-responsive genes (Karan et al., 2012; Wang B et al., 2015; Wang B et al., 2015). The expression of stress responsive genes is influenced by transposable element insertions in their upstream regions. Shahid (2020) reported increased methylation at CHH and CHG context in Miniature Inverted Repeat Transposable Elements in *OsHKT1;5* gene under salt stress. He observed role of methylation in regulation of *OsHKT1;5* gene (a major salt tolerance gene in rice that encode Na^+ transporter for exclusion of Na^+ from leaves and is important for Na^+/K^+ homeostasis under salt stress) and thus endowing salt tolerance (**Figure 2B**). High frequency of hypermethylation in the promoter located CpG sites has also been reported under salt stress conditions (Kumar et al., 2017; Ashapkin et al., 2020; Skorupa et al., 2021).

Owing to its heritable nature, any DNA methylation changes caused by environmental perturbations in plants have the prospect to be perpetuated and disseminated to future generations. This permits stress elicited methylation changes to proceed as a "memory" and help prime the plant or its progeny to counter more competently to the stress if re-exposed (Chaudhary et al., 2021).

3 HISTONE MODIFICATIONS

Histones are basic proteins consisting majorly of lysine and arginine residues that lay down the foundation of nucleosomal chromatin organization (Zhou et al., 2013). The N' termini of histone proteins, known as histone tails are the sites of covalent modifications such as acetylation, methylation, ubiquitination and phosphorylation. This covalent modification imparts different effects depending on the amino acid residue being modified and thus alters the genes activity (Banerjee et al., 2017). Indeed, studies in different plant species have demonstrated that histone modification is imperative to regulating gene expression under salt stress (Song et al., 2012; Li et al., 2014; Shen et al., 2014). Paul et al. (2017) reported differential regulation of *OsBZ8* gene expression due to significant differences in chromatin modification between *Oryza sativa* varieties IR64 and Nanabokra under salt stress. It has been demonstrated that tempering histone proteins provide an epigenetic molecular apparatus for priming plants to salt stress via the modulation of crucial salt responsive genes perpetuated throughout vegetative growth (Pikaard and Mittelsten Scheid, 2014).

3.1 Histone Acetylation

A negatively charged acetyl moiety on H3 and H4 histones serve to reduce the affinity between DNA and histone protein, enhancing DNA's accessibility to the transcriptional machinery (Onufriev and Schiessel, 2019). Acetylation of Lys residue 9 of histone H3 (H3K9ac) is largely investigated covalent modification and acts as new layers of supervision to cope with abiotic environmental stress through modulation of key regulatory factors (Zheng et al., 2016; Li et al., 2017; Ueda et al., 2017). Histone acetylation is frequently related with increased gene expression while deacetylation is associated with transcriptional repression (Zheng et al., 2016). Histone acetyl transferase (HATs) and histone deacetylases (HDAC) are the key enzymes that offer powerful transcriptional control mechanisms by catalyzing the addition and removal of an acetyl moiety respectively (Zhou et al., 2017; Kim et al., 2018).

Salt induced histone acetylation is linked with transcriptional activation of salt stress responsive genes reported in the case of *Nicotiana tabacum* (Sokol et al., 2007), *Zea mays* (Li et al., 2014) and *Saccharomyces cerevisiae* (Magraner-Pardo et al., 2014). Yolcu et al. (2016) demonstrated deposition of active histone marks such as H3K9ac and H3K4ac on the peroxidase gene resulting in its activation in *Beta vulgaris* and *B. maritime* (**Figure 3**). Increased expression of peroxidase gene has been linked with an activation of the ABA (abscisic acid) pathway and antioxidant enzymes, resulting in lower ROS (reactive oxygen species) accumulation and increased levels of osmotic metabolites therefore, augmenting salt tolerance (Su et al., 2020). Sako et al.

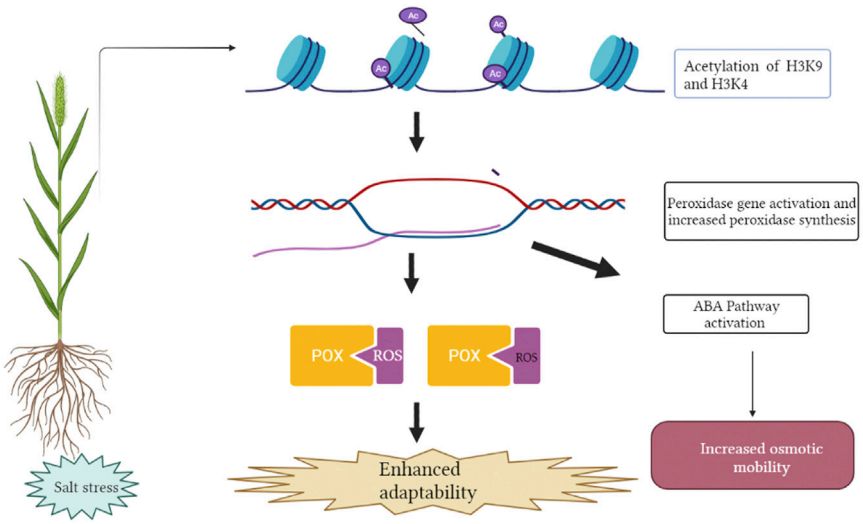


FIGURE 3 | Deposition of acetylation at H3K4 and H3K9 position leads to activation of salt responsive POX gene encoding peroxidase enzyme. Increased expression of peroxidase gene has been associated with activation of the ABA pathway (Absicsic acid) and antioxidant enzymes, resulting in lower ROS (Reactive Oxygen Species) accumulation and increased levels of osmotic metabolites. This figure was created using https://biorender.com.

(2016) reported that increased histone acetylation of *AtSOS1* and *AtSOS3* play an important role in salinity stress. Increased acetylation contributes to open a more relaxed chromatin confirmation ready for transcription. *TaHAG* (histone acetyltransferase) mediated H3 acetylation of polyploidy wheat genes involved in ROS production has been reported to up-regulate transcriptional changes of these genes in response to salt stress (Zheng et al., 2021). This gene in wheat and other crops can be manipulated as a potential target for salt tolerance improvement.

3.1.1 Role of HATs in Salinity Stress

The *Arabidopsis thaliana* genome contains four HAT (Histone acetylase transferase) gene families encoded by 12 HAT genes (Earley et al., 2007). Under salt stress conditions, the expression of cell wall related genes *ZmEXPANSIN B2* and ZmXYLOGLUCAN endotransglucosylase/hydrolase1) are up-regulated due to increased H3K9 acetylation at both the promoter and coding regions of genes. The increased acetylation of these genes is attributed to increased mRNA expression of two HAT genes (*ZmHATB and ZmGCN5*) under salt stress (Li et al., 2014). These observations have been further supported in *Arabidopsis*, where H3K9/K14 acetylation resulted in elevated expression levels of GCN5 under salt conditions and activated chitinase-like (CTL) protein involved in cell wall biosynthesis and salt tolerance (Zheng et al., 2019).

Although some HAT gene expression levels are shown to increase H4K5 acetylation during salt stress conditions, certain Histone deacetylases respond negatively to the salt stress resistance. Similarly *OsHDA1* was reported to negatively affect the transcriptional activation of *OsSOS1* in rice (Cheng et al., 2018). Zheng et al. (2021) claimed *TaHAG1* (histone acetyl transferase) to play decisive role in strengthening the salt tolerance in bread wheat. Further understanding the defined mechanisms by which HATs activities are modulated will offer new insight into the complex network regulating plant adaptation and tolerance to stress.

3.1.2 Role of Histone Deacetylases in Salinity Stress

Under favorable conditions, the repressive chromatin state of stress responsive genes is preserved by Histone deacetylases to keep gene transcripts at low levels. Histone deacetylases are involved in removing acetyl groups. Plants contain three families of Histone deacetylase proteins, i.e., i) Reduced potassium dependency 3 (RPD3)-like, ii) Silent Information Regulator 2 (SIRT) and iii) HD-tuins. The three Histone deacetylase families in the *A. thaliana* genome are encoded by 18 genes. Studies documented that upon exposure to abiotic stress, histone deacetylase genes display diversified responses and play a crucial role in how plants behave under such conditions (Asensi-Fabado et al., 2017; Kim et al., 2018).

Histone deacetylase over-expression in transgenic poplar plants reduced tolerance under salt stress (Ma et al., 2019a). Histone deacetylase9 constitutes a core histone deacetylase complex with PWR (POWERDRESS) and HOS15 (HIGH EXPRESSION OF OSMOTICALLY RESPONSIVE GENES), that binds to and directly represses many abiotic/biotic stress-responsive genes, including ethylene response factor (ERF) (ERF4/5/6/11), salt tolerance zinc finger (STZ), and kinase 2 (KIN2) genes, by modulating both histone acetylation (H3K27ac/H3K36ac/H3K56ac, H3.3K27/36ac andH4ac) and methylation (H3K9me2 and H3.1K36me2) (Mayer et al., 2019). Similarly *OsHDA1* was reported to negatively affect the transcriptional activation of *OsSOS1* in rice (Cheng et al., 2018). The OsHDA1 (histone deacetylase HDA1) is involved in the suppression of salt overly sensitive1 (*SOS1*) and late embryogenesis abundant protein1 (*LEA1*) genes, which are essential for salt tolerance in

rice, by decreasing H3 acetylation in the promoter regions of *LEA1* and *SOS1* genes (Cheng et al., 2018). In *Arabidopsis* class I (HDA19) family histone deacetylases are implicated in positive salinity responses and class II (HDA5/14/15/18) reduced potassium dependency3 (RPD3) histone deacetylases are involved in negative salinity responses (Ueda et al., 2017; Ueda et al., 2019). In *Hibiscus cannabinus,* HcHDA2, HcSRT2, HcHDA6, HcHDA8, HcHDA9, HcHDA19, and the levels of acetylation at H3K9ac, H3K27ac, and H4K5ac under salt stress conditions have been shown to be up-regulated (Wei et al., 2019). Similarly, HDA710/OsHDAC2, an HDA RPD3/HDA1 family member, contributes to controlling the rice salt stress genes by altering levels of H4 acetylation in their promoters. It regulates the acetylation at H4K5 and H4K16 under normal conditions. The accumulation of HDA710 transcripts under salt stress was considerably enhanced (Ullah et al., 2020). It is fascinating to break down specific function of the diverse HDACs in stress tolerance, genome-wide recognition of their target genes and investigation of alteration in histone acetylation at these genes under stress conditions. Moreover, how HDACs react to stress signaling to manage histone acetylation and expression of specific genes remains elusive.

3.2 Histone Methylation

Contrary to acetylation, histone methylation does not affect the electrostatic properties of histone proteins but it increases the hydrophobicity by changing intra or intermolecular interactions and may create novel binding sites for other proteins (Liu et al., 2010). Methyl group at Arg residue is added by Arg methyltransferases (PRMTs) while addition of methyl group at Lys residues is catalyzed by histone Lys methyltransferases (HKMTs). Two Arg methylation sites (H3R17 and H4R3) and five Lys methylation sites (H3K4, H3K9, H3K27, H3K36, and H4K20) have thus far been identified in plants (Liu et al., 2010). In *Glycine* max and *A. thaliana*, salt stress has been reported to increase methylation at fourth lysine of H3 (H3K4me3) and decrease histone H3 lysine 9 dimethylation and/or decreases histone H3 lysine 9 dimethylation (H3K9me2) associated with salt responsive genes (Bilichak et al., 2012; Song et al., 2012). Histone methylation in *Arabidosis* represent repressive (H4R3me2, H3K9me2/3, and H3K27me3) and active marks (H4R3me2, H3K4me3, and H3K36me2/3 (Liu et al., 2016). The presence or absence of methylation of Lys and/or Arg amino acids in histones alters their association with reader proteins, leading to modifications in chromatin structure that result in either transcriptional repression or activation (Teperino et al., 2010). Similarly, DNA methylation of H3 at 4th and 27th lysine in castor and rice crop plants has been demonstrated to regulate transcription of the critical salinity-response regulator (Karan et al., 2012; Han et al., 2020). Transcription of RSM1 (RADIALIS LIKE SANT-an MYB TF and key salt response regulator in salt signaling) has been reported to be guided by methylation at H3K4 and H3K27 in castor (Han et al., 2020). In the recent past it was found that the H3K4me0/1/2 code reader (GmPHD6) could specifically regulate the transcription of some salt-tolerance genes in *Glycine* max (Wei et al., 2017). Variation in methylation level at H3K4me3 and H3K27me3 has been reported to display differential expression level of salt responsive *OsBZ8* gene in rice varieties Nonabokra (salt tolerant) and IR64 (salt sensitive) (Paul et al., 2017). These observations evidently established important role of epigenetic marks H3K4me3 and H3K27me3 in regulating salt stress responsive genes and imparting salt tolerance. Furthermore, *JMJ15* gene (coding for H3K4 demethylase) over expression in *A. thaliana* under salt stress radically improved salt tolerance (Shen et al., 2014). The effects of histone methylation events vary depending on the site of the modification. For example, tri-methylation of the fourth lysine of H3 (H3K4me3) is an active mark for gene expression, and tri-methylation in the 27th lysine of H3 (H3K27me3) is a repressive mark of facultative heterochromatin (Doyle and Amasino, 2009). Although changes in histone modifications can be correlated with gene activity, the molecular mechanisms through which the chemical modifications influence chromosomal structure and the accessibility of transcription factors are still not fully understood. These relationships between the alteration of histone modifications and gene activity are highly conserved from yeast to human, and also in plants.

The histone methylation and acetylation have been extensively investigated in different plant species under salt stress conditions. Investigations deciphering other histone modifications may enrich our knowledge about other important epigenetic marks and their exploitation for breeding climate smart crops.

4 HISTONE VARIANTS

Of the various factors influencing chromatin dynamics and accessibility histone variants are also among the important ones that participate in modulating gene expression. Many species have been shown to encode numerous genes for core histone proteins, which are quite similar in amino acid sequence. Like histone proteins histone variants have also been shown to be differently expressed in *Oryza sativa* and *A. thaliana* (Talbert and Henikoff, 2021). In *A. thaliana*, 11 genes for H2B have been discovered, 13 for H2A, and 15 for H3 (Probst et al., 2020). The discovery that histone variant expression is tissue and developmental stage specific suggests that histone variations have particular functions in altering structural and functional properties of chromatin.

Histone variations that are replication-independent and replication-dependent can substitute for each other and are deliberately positioned within the genome. Each of the four histones (H2A, H2B, H3, and H1) have distinct variants. H2A is the most widely investigated histone and consists of H2A, H2A.Bbd, H2A.X, and H2A.Z variants (Bonisch and Hake, 2012). Similarly 14 variants of H4 (Siegel et al., 2009; Moosmann et al., 2011; Bonisch and Hake, 2012) and two different isoforms of H2A known as H2A.Z.1 and H2A.Z.2 displaying specific functions (differing in only three amino acids) have been reported (Coon et al., 2005; Eirín-López et al., 2009; Talbert et al., 2012). The expression of the H2A.Z variant of H2A histone has been diminished in *O. sativa* and *A. thaliana* during salt or other stress (Nguyen and Cheong, 2018; Zahraeifard et al., 2018). H2A.Z has been portrayed

TABLE 1 | Long non coding RNAs/miRNAs involved in imparting salt tolerance.

S.N.	lncRNA/miRNA	Plant Species	Characteristics	References
1	ThSAIR6	*Tamarix hispida*	Decreased the contents of H_2O_2 and enhanced activity of anti-oxidative enzymes	Xu et al. (2021)
2	AtR8lncRNA	*Arabidopsis thaliana*	Regulate seed germination in response to salt	Zhang et al. (2020)
3	LncRNA973	*Gossypium hirsutum*	Increased expression resulted into increased salt tolerance	Zhang et al. (2019)
4	Pal_00132209	*Populus alba*	Affect fucosyltransferase or NAC3 and regulates growth under salt stress	Ma et al. (2019b)
5	Pal _00184400	*Populus alba*	HKT1 and show differential expression in xylem	Ma et al. (2019b)
6	lnc_388, lnc_973, lnc_253	*Gossypium hirsutum*	Regulates tolerance to salt stress	Deng et al. (2018)
7	DRIR (Drought Induced long non coding RNA)	*Arabidopsis thaliana*	Regulates ABA mediated responses to both salt and drought	Qin et al. (2017)
8	*TCONS_00116877*	*Medicago truncatula*	Regulates oxidative stress under salt conditions	Wang et al. (2015)
9	*TCONS_00046739*	*Medicago truncatula*	Regulates cytochrome P450 under salt stress	Wang et al. (2015)
10	miR156, miR398	*Solanum lycopersicum*	Increased expression levels imparted salt tolerance	Cakir et al. (2021)
11	nta-miR156a_R + 3, farmiR159_L + 2_1ss22T, mes-MIR319e-p5_2ss12GC19 GA	*Ipomoea batatas*	tissue specific expression under salt stress	Yang et al. (2020)
12	miR26, miR05, miR20, miR31, miR11, miR28, miR15, miR14, miR32, miR09, miR22, miR33, miR19, miR24	*Pennisetum glaucum*	Shows altered expression under salinity	Shinde et al. (2020)
13	miR172, miR319, miR408, miR2590	*Medicago sativa*	Regulates gene associated with salt tolerance	Ma et al. (2019b)
14	TaemiR408	*Triticum aestivum*	Overexpresion resulted in enhanced salt tolerance	Bai et al. (2018)
15	miR164s, mir-36	*Zea mays*	up-regulated in leaves under salt treatment	Fu et al. (2017)
16	osa-miR1878, osa- miR2863c	*Oryza sativa*	Upregulated under salt stress	Goswami et al. (2017)
17	miR171b, miR167f	*Oryza sativa*	Promotes better adaptability to salt	Parmar et al. (2020)
18	sly-miR156e-5p, slymiRn23b, slymiRn50a	*Solanum pimpinellifolium*	Involved in stress related pathways	Zhao et al. (2017)
19	miR172	*Glycine max*	Improves salt tolerance	Pan et al. (2016)
20	miRNVL5	*Gossypium hirsutum*	Regulation of plant stress to salt	Gao et al. (2016)
21	miR-395	*Cucumis sativus*	Up-regulated and regulates ATP sulfurylase	Li et al. (2016)
22	miR156/157, miR158, miR166, miR168 and miR408	*Raphanus sativus*	Expression was upregulated significantly	Sun et al. (2015)
23	miR-160	*Gossypium raimondii*	Up-regulated under salt stress and control Auxin response factor (ARF)	Xie et al. (2015)

as a crucial thermosensor (Kumar and Wigge, 2010) during stress response. H2A.W predominantly found in heterochromatin is engaged in stress induced chromatin decondensation. In *A. thaliana* replacement of H3.3 has been shown to be correlated with transcriptional process and declining H3.3 brings down transcription of stress responsive genes (Wollmann et al., 2017). Accumulation of H3.3 avert H1 histone from acquiring its position at gene body to assist DNA methylation which further alienate deposition of H2A.Z (Zilberman et al., 2008; Wollman et al., 2017). This explains why H3.3 is indispensable for stress responsive gene expression. This aspect of chromatin modification is however not much explored yet and offer exciting possibilities to understand the role of histone variants at different growth and development stages in response to stress. Deposition of histone variations under stress gives the possible way to connect environmental cues to transcription downstream reactions. More investigations are required to define how it generates epigenetic memory clearly.

5 PLANT MICRORNAS AND LONG NON CODING RNAS: KEY EPIGENETIC REGULATORS

Plants adopt *de novo* DNA methylation and gene silencing (transcriptional) using 24- nucleotide small-interfering RNAs and long non-coding RNAs in the RNA-directed DNA methylation process (Kovalchuk, 2016). RNA dependent DNA methylation (RdDM) is the only system in plants that can introduce DNA methylation to cytosines irrespective of the sequence context (Magraner-Pardo et al., 2014) (**Figure 2A**). This pathway helps plants in surviving under adverse environmental conditions like salt stress (Fortes and Gallusci, 2017). Under saline conditions RdDM becomes down-regulated and elicit the expression of transcription factors central for salt stress tolerance (Xie et al., 2015a). The plant microRNAs (miRNA) are 20–24 nt, non-coding RNA species that have been portrayed as tiny yet potent regulators of gene expression in plants as well as animals. These miRNAs are either positively regulated by stress, where they enhance the repression of the genes

serving as negative regulators of stress tolerance or negatively regulated where the target is positive regulator of stress causing more accumulation of gene product (Sunkar et al., 2007; Singroha et al., 2021). The biogenesis of miRNAs has been reviewed by Singroha et al. (2021). Most of the miRNAs responsive to salt stress directly regulate transcription factors. miR164a/b/c/d/and miR1661m identified from *Zea mays* have been shown to target *MYB, NAC* and homeodomain-leucine zipper protein (HD-ZIP) transcription factors under salt stress (Ding et al., 2009). It has also been observed that miRNA exhibit species specific behavior in response to salt stress. For instance the expression of miR156 was induced under salt stress in *A. thaliana* while diminished in *Z. mays* (Liu et al., 2008; Ding et al., 2009). In the same way expression of miR396 was up-regulated in *A. thaliana* and *Z. mays* upon salt treatment but diminished in *O. sativa* (Liu et al., 2008; Ding et al., 2009).

MYB74 (a member of the R2R3-MYB gene family) is transcriptionally regulated mainly by RdDM pathway under salt stress in *A. thaliana*. 24-nt siRNAs (small interfering RNA) target a region approximately 500bp upstream of the transcription start site of *MYB74,* which is heavily methylated. Levels of DNA methylation in this region were significantly diminished in wild type plants under salt stress, whereas no changes were observed in RdDM mutants. These observations suggest that changes in the levels of the five 24-nt siRNAs regulate the *MYB74* transcription factor via RdDM under salt stress conditions (Xu et al., 2015). The salt-tolerant regulation of MYB transcription factors involves ABA signaling pathway and other signal transduction pathways in plants. Salt stress subjected plants exhibited significantly increased ABA content that can induce proline accumulation in plants, and enhance the activity of related protective enzyme and up-regulation of related stress responsive genes (Schmidt et al., 2013). The investigations made in this area have tried to extend our understanding of non-coding RNAs functional processes for salt stress in *A. thaliana* (Qin et al., 2017), *H. vulgare* (Karlik and Gozukirmizi, 2018), cotton (Zhang et al., 2019a), *Spirodela polirhiza* (Fu et al., 2020) and sorghum (Sun et al., 2020).

Under salt stress conditions *Z. mays* displayed down-regulation of miR-250, miR-205, miR-330 and miR-17 in leaves and roots (Fu et al., 2017). Down-regulation of these miRNAs enhanced the expression of their targets viz. *casein kinase II*, *GPX*, *P5CS*, *IF-1* and some other genes essential for better survival of the plant under saline conditions. This is how miRNAs regulate gene expression under stress conditions and help plants in their survival under harsh environmental conditions. Apart from 24 nt long miRNAs, the long non-coding RNAs abbreviated as lncRNAs have also been defined as riboregulators longer than 200 bp (Kapranov et al., 2007). They also regulate gene expression under stress conditions through transcriptional or post transcriptional silencing. Chen and associates (2019) identified 3030 long intergenic non-coding RNAs in *Glycine max* roots under salt stress conditions. For example, the long non-coding RNA NPC60 expression was escalated 100 times under salt stress condition. Similarly salt treatments enhanced levels of long non-coding RNA973 in cotton (Zhang et al., 2019b). The over expression of lncRNA973 displayed high salt tolerance, which modulates cotton salt genes expression. Ma et al. (2019a) demonstrated tissue, and species specific expression of long non-coding RNA in Poplar species under different salt stress conditions. A list of plant small and long non-coding RNAs expressed in response to salt stress is provided in **Table 1**.

6 CONCLUSION AND FUTURE OUTLOOK

Many findings have emphasized epigenetic regulations as powerful mechanisms for regulating the implications of salt stress on plants and provide an excellent foundation for development of salt-tolerant crop plants. In plants susceptible to salt stress, epigenetic controls are associated with the stringent control of gene expression. Epigenetic marks on stress-induced genes dynamically affect the accessibility of chromatin and the expression of those genes. The different regulatory mechanisms for abiotic stress responses might involve epigenetic alterations such as methylation, histone changes, chromatin remodelling, histone variants and lncRnAs.

The critical role of epigenetic modifications in regulating gene expression and their ability to transfer to the next generation makes them a unique adaptation tool for plants. The phenotypic plasticity caused by epigenetic variation, which in turn, is through changes in gene expression, will affect fitness and eventually natural selection in plants. Unlike classic DNA sequence mutations, epimutations can happen at much shorter times, and even though they are stable, they are primarily reversible, making them a perfect tool for a quick emergency response to unpredictable environmental stresses. It must also be highlighted that epigenetic changes are typically dependent on the underlying genetic variation, and these two factors must be addressed concurrently. Future study is required to better understand the epigenetic mechanisms behind chromatin changes and the resulting transcriptional regulation that impacts plant responses to environmental stresses. More study on the mechanism of hereditary stress memory is also required.

AUTHOR CONTRIBUTIONS

PS: conceptualization, methodology, data curation, formal analysis, tables generation, writing—original draft, writing—review and editing. GS: data curation, formal analysis, figures and tables generation, writing—original draft. OG: data curation, figures, writing—original draft. SK and GPS: writing—review and editing.

FUNDING

This work is supported by the Department of Science and Technology under the scheme DST-WOS-A (grant no DST/WOS-A/LS/41-2018) to GS and ICAR-LBSYOA scheme (Grant No. 1008018) to PS.

REFERENCES

Asensi-Fabado, M.-A., Amtmann, A., and Perrella, G. (2017). Plant Responses to Abiotic Stress: The Chromatin Context of Transcriptional Regulation. *Biochim. Biophys. Acta (Bba) - Gene Regul. Mech.* 1860, 106–122. doi:10.1016/j.bbagrm.2016.07.015

Ashapkin, V. V., Kutueva, L. I., and Aleksandrushkina, N. I. Epigenetic Mechanisms of Plant Adaptation to Biotic and Abiotic Stresses. *Int. J. Mol. Sci.* 2020;21:7457.doi:10.3390/ijms21207457

Bai, Q., Wang, X., and Chen, X. (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

Banerjee, A., Wani, S. H., and Roy, C. A. (2017). Epigenetic Control of Plant Cold Responses. *Front. Plant Sci.* 8, 1643. doi:10.3389/fpls.2017.01643

Bonisch, C., and Hake, S. B. (2012). Histone H2A Variants in Nucleosomes and Chromatin: More or Less Stable? *Nucleic Acids Res.* 40, 10719–10741. doi:10.1093/nar/gks865

Çakır, O., Arıkan, B., and Karpuz, B. (2021). Expression Analysis of miRNAs and Their Targets Related to Salt Stress in *Solanum lycopersicum* H-2274. *Biotechnol. Biotechnological Equipment* 35, 283–290.

Chang, Y. N., Zhu, C., Jiang, J., Zhang, H., Zhu, J. K., and Duan, C. G. (2020). Epigenetic Regulation in Plant Abiotic Stress Responses. *J. Integr. Plant Biol.* 62, 563–580. doi:10.1111/jipb.12901

Chen, R., Li, M., and Zhang, H. (2019). Continuous Salt Stress-Induced Long Non-coding RNAs and DNA Methylation Patterns in Soybean Roots. *BMC Genomics* 20, 730. doi:10.1186/s12864-019-6101-7

Cheng, X., Zhang, S., and Tao, W. (2018). INDETERMINATE SPIKELET1 Recruits Histone Deacetylase and a Transcriptional Repression Complex to Regulate rice Salt Tolerance. *Plant Physiol.* 178, 824–837. doi:10.1104/pp.18.00324

Coon, J. J., Ueberheide, B., and Syka, J. E. (2005). Protein Identification Using Sequential Ion/ion Reactions and Tandem Mass Spectrometry. *Proc. Natl. Acad. Sci. USA* 102, 9463–9468. doi:10.1073/pnas.0503189102

Crisp, P. A., Ganguly, D., Eichten, S. R., Borevitz, J. O., and Pogson, B. J. (2016). Reconsidering Plant Memory: Intersections between Stress Recovery, RNA Turnover, and Epigenetics. *Sci. Adv.* 2 (2), e1501340. doi:10.1126/sciadv.1501340

Deng, F., Zhang, X., and Wang, W. (2018). Identification of *Gossypium Hirsutum* Long Non-coding RNAs (lncRNAs) under Salt Stress. *BMC Plant Biol.* 18, 23. doi:10.1186/s12870-018-1238-0

Ding, D., Zhang, L., and Wang, H. (2009). Differential Expression of miRNAs in Response to Salt Stress in maize Roots. *Ann. Bot.* 103, 29–38. doi:10.1093/aob/mcn205

Doyle, M. R., and Amasino, R. M. (2009). A Single Amino Acid Change in the Enhancer of Zeste Ortholog CURLY LEAF Results in Vernalization-independent, Rapid Flowering in *Arabidopsis*. *Plant Physiol.* 151, 1688–1697. doi:10.1104/pp.109.145581

Duan, C. G., Zhu, J. K., and Cao, X. (2018). Retrospective and Perspective of Plant Epigenetics in China. *J. Genet. Genomics* 45, 621–638. doi:10.1016/j.jgg.2018.09.004

Earley, K. W., Shook, M. S., and Brower-Toland, B. (2007). Invitro Specificities of Arabidopsis Co-activator Histone Acetyltransferases: Implications for Histone Hyperacetylation in Gene Activation. *Plant J.* 52, 615–626. doi:10.1111/j.1365-313x.2007.03264.x

Eirín-López, J. M., González-Romero, R., and Dryhurst, D. (2009). The Evolutionary Differentiation of Two Histone H2A.Z Variants in Chordates (H2A.Z-1 and H2A.Z-2) Is Mediated by a Stepwise Mutation Process that Affects Three Amino Acid Residues. *BMC Evolutionary Biol.* 9, 31. doi:10.1186/1471-2148-9-31

FAO and ITPS (2015). *Status of the World's Soil Resources (SWSR) – Main Report*. Rome, Italy: Food and Agriculture Organization of the United Nations and Intergovernmental Technical Panel on Soils.

Fortes, A. M., and Gallusci, P. (2017). Plant Stress Responses and Phenotypic Plasticity in the Epigenomics Era: Perspectives on the Grapevine Scenario, a Model for Perennial Crop Plants. *Front. Plant Sci.* 8, 82. doi:10.3389/fpls.2017.00082

Fu, L., Ding, Z., Tan, D., Han, B., Sun, X., and Zhang, J. (2020). Genome-wide Discovery and Functional Prediction of Salt-Responsive lncRNAs in Duckweed. *BMC Genomics* 21, 212. doi:10.1186/s12864-020-6633-x

Fu, R., Zhang, M., and Zhao, Y. (2017). Identification of Salt Tolerance-Related microRNAs and Their Targets in Maize (Zea mays L.) Using High-Throughput Sequencing and Degradome Analysis. *Front. Plant Sci.* 8, 864. doi:10.3389/fpls.2017.00864

Gao, S., Yang, L., and Zeng, H. (2016). A Cotton miRNA Is Involved in Regulation of Plant Response to Salt Stress. *Sci. Rep.* 6, 19736. doi:10.1038/srep19736

Goswami, K., Tripathi, A., and Mishra, S. (2017). Comparative miRomics of Salt-Tolerant and Salt-Sensitive Rice. *J. Integr. Bioinformatics* 14, 20170002. doi:10.1515/jib-2017-0002

Han, B., Xu, W., and Ahmed, N. (2020). Changes and Associations of Genomic Transcription and Histone Methylation with Salt Stress in Castor Bean. *Plant Cel Physiol.* 61, 1120–1133. doi:10.1093/pcp/pcaa037

Kapranov, P., Cheng, J., and Dike, S. (2007). RNA Maps Reveal New RNA Classes and a Possible Function for Pervasive Transcription. *Science* 316, 1484–1488. doi:10.1126/science.1138341

Karan, R., DeLeon, T., and Biradar, H. (2012). Salt Stress Induced Variation in DNA Methylation Pattern and its Influence on Gene Expression in Contrasting Rice Genotypes. *PLoS ONE* 7, e40203. doi:10.1371/journal.pone.0040203

Karlik, E., and Gozukirmizi, N. (2018). Expression Analysis of lncRNA AK370814 Involved in the Barley Vitamin B6 Salvage Pathway under Salinity. *Mol. Biol. Rep.* 45, 1597–1609. doi:10.1007/s11033-018-4289-2

Kim, J. Y., Yang, W., and Forner, J. (2018). Epigenetic Reprogramming by Histone Acetyltransferase HAG1/AtGCN5 Is Required for Pluripotency Acquisition in *Arabidopsis*. *EMBO J.* 37, e98726. doi:10.15252/embj.201798726

Kinoshita, T., and Seki, M. (2014). Epigenetic Memory for Stress Response and Adaptation in Plants. *Plant Cel Physiol* 55, 1859–1863. doi:10.1093/pcp/pcu125

Konate, M., Wilkinson, M. J., and Mayne, B. T. (2018). Salt Stress Induces Non-CG Methylation in Coding Regions of Barley Seedlings (*Hordeum Vulgare*). *Epigenomes* 2, 12. doi:10.3390/epigenomes2020012

Kovalchuk, I. (2016). "Transgenerational Genome Instability in Plants," in *Genome Stability* (Amsterdam, Netherlands: Elsevier BV), 615–633. doi:10.1016/b978-0-12-803309-8.00036-7

Kumar, S., Beena, A. S., and Awana, M. (2017). Salt-Induced Tissue-specific Cytosine Methylation Downregulates Expression of HKT Genes in Contrasting Wheat (*Triticum aestivum* L.) Genotypes. *DNA Cel Biol* 36, 283–294. doi:10.1089/dna.2016.3505

Kumar, S., Chinnusamy, V., and Mohapatra, T. (2018). Epigenetics of Modified DNA Bases: 5- Methylcytosine and beyond. *Front. Genet.* 9, 640. doi:10.3389/fgene.2018.00640

Kumar, V., and Wigge, P. A. (2010). H2A.Z-Containing Nucleosomes Mediate the Thermosensory Response in *Arabidopsis*. *Cell* 140, 136–147. doi:10.1016/j.cell.2009.11.006

Lämke, J., and Bäurle, I. (2017). Epigenetic and Chromatin-Based Mechanisms in Environmental Stress Adaptation and Stress Memory in Plants. *Genome Biol.* 18, 124. doi:10.1186/s13059-017-1263-6

Li, F., Zheng, L. D., and Chen, X. (2017). Gcn5-mediated Rph1 Acetylation Regulates its Autophagic Degradation under DNA Damage Stress. *Nucleic Acids Res.* 45, 5183–5197. doi:10.1093/nar/gkx129

Li, H., Wang, Yu., and Wang, Z. (2016). Microarray and Genetic Analysis Reveals that Csa-miR159b Playsa Critical Role in Abscisic Acid-Mediated Heat Tolerance in Grafted Cucumber Plants. *Plant Cel Environ.* 39, 1790–1804. doi:10.1111/pce.12745

Li, H., Yan, S., and Zhao, L. (2014). Histone Acetylation Associated Up-Regulation of the Cell wall Related Genes Is Involved in Salt Stress Induced maize Root Swelling. *BMC Plant Biol.* 14, 105. doi:10.1186/1471-2229-14-105

Liang, Z., Shen, L., Cui, X., Bao, S., Geng, Y., Yu, G., et al. (2018). DNA N^6-Adenine Methylation in *Arabidopsis thaliana*. *Dev. Cel* 45, 406–416. doi:10.1016/j.devcel.2018.03.012

Liu, C., Lu, F., and Cui, X. (2010). Histone Methylation in Higher Plants. *Annu. Rev. Plant Biol.* 61, 395–420. doi:10.1146/annurev.arplant.043008.091939

Liu, C., Wang, C., Wang, G., Becker, C., Zaidem, M., and Weigel, D. (2016). Genome-wide Analysis of Chromatin Packing in *Arabidopsis thaliana* at Single-Gene Resolution. *Genome Res.* 26 (8), 1057–1068. Aug. doi:10.1101/gr.204032.116

Liu, H. H., Tian, X., and Li, Y. J. (2008). Microarray-based Analysis of Stress-Regulated microRNAs in *Arabidopsis thaliana*. *RNA* 14, 836–843. doi:10.1261/rna.895308

Liu, R., and Lang, Z. (2019). The Mechanism and Function of Active DNA Demethylation in Plants. *J. Integr. Plant Biol.* doi:10.1111/jipb.12879

Luo, X., and He, Y. (2020). Experiencing winter for spring Flowering: A Molecular Epigenetic Perspective on Vernalization. *J. Integr. Plant Biol.* 62, 104–117. doi:10.1111/jipb.12896

Ma, J., Bai, X., and Luo, W. (2019a). Genome-Wide Identification of Long Noncoding RNAs and Their Responses to Salt Stress in Two Closely Related Poplars. *Front. Genet.* 10, 777. doi:10.3389/fgene.2019.00777

Ma, J., Wang, Y., and Li, J. (2019b). Global Identification and Analysis of microRNAs Involved in Salt Stress Responses in Two Alfalfa (*Medicago Sativa* 'Millennium') Lines. *Can. J. Plant Sci.* 100, 445–455.

Magraner-Pardo, L., Pelechano, V., and Coloma, M. D. (2014). Dynamic Remodeling of Histone Modifications in Response to Osmotic Stress in. *SaccharomycesBMC Genom* 15, 247. doi:10.1186/1471-2164-15-247

Mayer, K. S., Chen, X., and Sanders, D. (2019). HDA9-PWR-HOS15 Is a CoreHistone Deacetylase Complex Regulating Transcription and Development. *Plant Physiol.* 180, 342–355. doi:10.1104/pp.18.01156

Moosmann, A., Campsteijn, C., and Jansen, P. W. (2011). Histone Variant Innovation in a Rapidly Evolving Chordate Lineage. *BMC Evol. Biol.* 11, 208. doi:10.1186/1471-2148-11-208

Nguyen, N. H., and Cheong, J. J. (2018). 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

Onufriev, A. V., and Schiessel, H. (2019). The Nucleosome: from Structure to Function through Physics. *Curr. Opin. Struct.* 56, 119–130. doi:10.1016/j.sbi.2018.11.003

Pan, W. J., Tao, J. J., and Cheng, T. (2016). Soybean miR172a Improves Salt Tolerance and Can Function as a Long-Distance Signal. *Mol. Plant* 9, 1337–1340. doi:10.1016/j.molp.2016.05.010

Parmar, S., Gharat, S. A., and Tagirasa, R. (2020). Identification and Expression Analysis of miRNAs and Elucidation of Their Role in Salt Tolerance in rice Varieties Susceptible and Tolerant to Salinity. *PLoS ONE* 15, e0230958. doi:10.1371/journal.pone.0230958

Paul, A., Dasgupta, P., and Roy, D. (2017). Comparative Analysis of Histone Modifications and DNA Methylation at OsBZ8 Locus under Salinity Stress in IR64 and Nonabokra rice Varieties. *Plant Mol. Biol.* 95, 63–88. doi:10.1007/s11103-017-0636-2

Pikaard, C. S., and Mittelsten Scheid, O. (2014). Epigenetic Regulation in Plants. *Cold Spring Harbor Perspect. Biol.* 6, a019315. doi:10.1101/cshperspect.a019315

Probst, A. V., Desvoyes, B., and Gutierrez, C. (2020). Similar yet Critically Different: The Distribution, Dynamics and Function of Histone Variants. *J. Exp. Bot.* 71, 5191–5204. doi:10.1093/jxb/eraa230

Qin, T., Zhao, H., and Cui, P. (2017). A Nucleus-Localized Long Non-coding RNA Enhances Drought and Salt Stress Tolerance. *Plant Physiol.* 175, 1321–1336. doi:10.1104/pp.17.00574

Rodríguez López, C. M., and Wilkinson, M. J. (2015). Epi-fingerprinting and Epi-Interventions for Improved Crop Production and Food Quality. *Front. Plant Sci.* 6, 397. doi:10.3389/fpls.2015.00397

Sako, K., Kim, J. M., and Matsui, A. (2016). Ky-2, a Histone Deacetylase Inhibitor, Enhances High-Salinity Stress Tolerance in *Arabidopsis thaliana*. *Plant Cel Physiol* 57, 776–783. doi:10.1093/pcp/pcv199

Saroha, M., Singroha, G., and Sharma, M. (2017). sRNA and Epigenetic Mediated Abiotic Stress Tolerance in Plants. *Ind. J. Plant Physiol.* 22, 458–469. doi:10.1007/s40502-017-0330-z

Schmidt, R., Schippers, J. H., Mieulet, D., Obata, T., Fernie, A. R., Guiderdoni, E., et al. (2013). MULTIPASS, a rice R2R3-type MYB Transcription Factor, Regulates Adaptive Growth by Integrating Multiple Hormonal Pathways. *Plant J.* 76, 258–273. doi:10.1111/tpj.12286

Shahid, S. (2020). A DNA Methylation Reader with an Affinity for Salt Stress. *Plant Cell* 32, 3380–3381. doi:10.1105/tpc.20.00800

Shen, Y., Conde, E., and Audonnet, L. (2014). Over-expression of Histone H3K4 Demethylase Gene JMJ15 Enhances Salt Tolerance in *Arabidopsis*. *Front. Plant Sci.* 5, 290. doi:10.3389/fpls.2014.00290

Shinde, H., Dudhatea, A., and Anandb, L. (2020). Small RNA Sequencing Reveals the Role of Pearl Millet miRNAs and Their Targets in Salinity Stress Responses. *S Afr. J. Bot.* 132, 395–402. doi:10.1016/j.sajb.2020.06.011

Siegel, T. N., Hekstra, D. R., and Kemp, L. E. (2009). Four Histone Variants Mark the Boundaries of Polycistronic Transcription Units in *Trypanosoma Brucei*. *Genes Dev.* 23, 1063–1076. doi:10.1101/gad.1790409

Singroha, G., and Sharma, P. (2019). "Epigenetic Modifications in Plants under Abiotic Stress," in *Epigenetics. Intechopen*. Editor R. Meccariello (London. doi:10.5772/intechopen.84455

Singroha, G., Sharma, P., and Sunkur, R. (2021). Current Status of microRNA-Mediated Regulation of Drought Stress Responses in Cereals. *Physiol. Plant* 172, 1808–1821. doi:10.1111/ppl.13451

Skorupa, M., Szczepanek, J., and Mazur, J. (2021). Salt Stress and Salt Shock Differently Affect DNA Methylation in Salt-Responsive Genes in Sugar Beet and its Wild, Halophytic Ancestor. *PLoS ONE* 16, e0251675. doi:10.1371/journal.pone.0251675

Sokol, A., Kwiatkowska, A., and Jerzmanowski, A. (2007). Up-regulation of Stress-Inducible Genes in Tobacco and *Arabidopsis* Cells in Response to Abiotic Stresses and ABA Treatment Correlates with Dynamic Changes in Histone H3 and H4 Modifications. *Planta* 227, 245–254. doi:10.1007/s00425-007-0612-1

Song, Y., Ji, D., and Li, S. (2012). The Dynamic Changes of DNA Methylation and Histone Modifications of Salt Responsive Transcription Factor Genes in Soybean. *PLoS One* 7, e41274. doi:10.1371/journal.pone.0041274

Su, P., Yan, J., and Li, W. (2020). A Member of Wheat Class III Peroxidase Gene Family, *TaPRX-2A,* Enhanced the Tolerance of Salt Stress. *BMC Plant Biol.* 20, 392. doi:10.1186/s12870-020-02602-1

Sun, X., Xu, L., and Wang, Y. (2015). Identification of Novel and Salt-Responsive miRNAs to Explore miRNA-Mediated Regulatory Network of Salt Stress Response in Radish (*Raphanus Sativus* L.). *BMC Genomics* 16, 197. doi:10.1186/s12864-015-1416-5

Sun, X., Zheng, H., and Li, J. (2020). Comparative Transcriptome Analysis Reveals New lncRNAs Responding to Salt Stress in Sweet Sorghum. *Front. Bioeng. Biotechnol.* 8, 331. doi:10.3389/fbioe.2020.00331

Sunkar, R., Chinnusamy, V., and Zhu, J. (2007). Small RNAs as Big Players in Plant Abiotic Stress Responses and Nutrient Deprivation. *Trends Plant Sci.* 12, 301–309. doi:10.1016/j.tplants.2007.05.001

Talbert, P. B., Ahmad, K., Almouzni, G., et al. (2012). A Unified Phylogeny-Based Nomenclature for Histone Variants. *Epigenet Chromatin* 25, 7. doi:10.1186/1756-8935-5-7

Talbert, P. B., and Henikoff, S. (2021). Histone variants a glanceJournal. *Cel Sci.* 134, jcs244749. doi:10.1242/jcs.244749

Teperino, R., Schoonjans, K., and Auwerx, J. (2010). Histone Methyl Transferases and Demethylases; Can They Link Metabolism and Transcription? *Cell Metab* 6 (12), 321–327. doi:10.1016/j.cmet.2010.09.004

Ueda, M., Matsui, A., and Tanaka, M. (2017). The Distinct Roles of Class I and II RPD3-like Histone Deacetylases in Salinity Stress Response. *Plant Physiol.* 175, 1760–1773. doi:10.1104/pp.17.01332

Ueda, M., Matsui, A., Watanabe, S., Kobayashi, M., Saito, K., Tanaka, M., et al. (2019). Transcriptome Analysis of the Hierarchical Response of Histone Deacetylase Proteins that Respond in an Antagonistic Manner to Salinity Stress. *Front. Plant Sci.* 10, 1323. doi:10.3389/fpls.2019.01323

Ullah, F., Xu, Q., and Zhao, Y. (2020). Histone Deacetylase HDA710 Controls Salt Tolerance by Regulating ABA Signaling in rice. *J. Integr. Plant Biol.* Epub. doi:10.1111/jipb.13042

Varotto, S., Tani, E., and Abraham, E. (2020). Epigenetics: Possible Applications in Climate-Smart Crop Breeding. *J. Exp. Bot.* 71, 5223–5236. doi:10.1093/jxb/eraa188

Voigt, A., Ribot, J., and Sabater, A. G. (2015). Identification of *Mest/Peg1* Gene Expression as a Predictive Biomarker of Adipose Tissue Expansion Sensitive to Dietary Anti-obesity Interventions. *Genes Nutr.* 10, 27. doi:10.1007/s12263-015-0477-z

Wang, B., Fu, R., and Zhang, M. (2015). Analysis of Methylation-Sensitive Amplified Polymorphism in Different Cotton Accessions under Salt Stress Based on Capillary Electrophoresis. *Genes Genom* 37, 713–724. doi:10.1007/s13258-015-0301-6

Wang, T. Z., Liu, M., and Zhao, M. G. (2015). Identification and Characterization of Long Non-coding RNAs Involved in Osmotic and Salt Stress in *Medicago Truncatula* Using Genome-wide High-Throughput Sequencing. *BMC Plant Biol.* 15, 131. doi:10.1186/s12870-015-0530-5

Wei, F., Tang, D., and Li, Z. (2019). Molecular Cloning and Subcellular Localization of Six *HDACs* and Their Roles in Response to Salt and Drought Stress in Kenaf (*Hibiscus Cannabinus* L.). *Biol. Res.* 52, 20. doi:10.1186/s40659-019-0227-6

Wei, W., Tao, J. J., and Chen, H. W. (2017). A Histone Code Reader and a Transcriptional Activator Interact to Regulate Genes for Salt Tolerance. *Plant Physiol.* 175, 1304–1320. doi:10.1104/pp.16.01764

Wollmann, H., Stroud, H., and Yelagandula, R. (2017). The Histone H3 Variant H3.3 Regulates Gene Body DNA Methylation in *Arabidopsis thaliana*. *Genome Biol.* 18, 94. doi:10.1186/s13059-017-1221-3

Xie, F., Wang, Q., and Sun, R. (2015). Deep Sequencing Reveals Important Roles of microRNAs in Response to Drought and Salinity Stress in Cotton. *J. Exp. Bot.* 66, 789–804. doi:10.1093/jxb/eru437

Xu, R., Wang, Y., Zheng, H., Lu, W., Wu, C., Huang, J., et al. (2015). Salt-induced Transcription Factor MYB74 Is Regulated by the RNA-Directed DNA Methylation Pathway in Arabidopsis. *J. Exp. Bot.* 66, 5997–6008. doi:10.1093/jxb/erv312

Xu, X., Lu, H., and Wang, Y. (2021). Salt Stress Tolerance Analysis of *SAIR*6 Long Non-coding RNA in *Tamarix Hispida*. *J. Beijing For. Univ.* 243, 36–43.

Yang, Z., Zhu, P., and Kang, H. (2020). High-throughput Deep Sequencing Reveals the Important Role that microRNAs Play in the Salt Response in Sweet Potato (*Ipomoea Batatas* L.). *BMC Genomics* 21, 164. doi:10.1186/s12864-020-6567-3

Yolcu, S., Ozdemir, F., and Güler, A. (2016). Histone Acetylation Influences the Transcriptional Activation of POX in *Beta Vulgaris* L. And *Beta Maritima* L. under Salt Stress. *Plant Physiol. Biochem.* 100, 37–46. doi:10.1016/j.plaphy.2015.12.019

Zahraeifard, S., Foroozan, M., Sepehri, A., and Rice, H2A. (2018). Z Negatively Regulates Genes Responsive to Nutrient Starvation but Promotes Expression of Key Housekeeping Genes. *J. Exp. Bot.* 14, 4907–4919. doi:10.1093/jxb/ery244

Zhang, H., Lang, Z., and Zhu, J. K. (2018). Dynamics and Function of DNA Methylation in Plants. *Nat. Rev. Mol. Cel Biol* 19, 489–506. doi:10.1038/s41580-018-0016-z

Zhang, N., Ziguang, L., and Shichen, S. (2020). Response of *AtR8* lncRNA to Salt Stress and its Regulation on Seed Germination in *Arabidopsis*. *Chin. J. Bot.* 55, 421–429.

Zhang, X., Bernatavichute, Y. V., and Cokus, S. (2009). Genome-wide Analysis of Mono-, Di- and Trimethylation of Histone H3 Lysine 4 in *Arabidopsis thaliana*. *Genome Biol.* 10, R62. doi:10.1186/gb-2009-10-6-r62

Zhang, X., Dongand Deng, F. (2019b). The Long Non-coding RNA lncRNA973 Is Involved in Cotton Response to Salt Stress. *BMC Plant Biol.* 19, 459. doi:10.1186/s12870-019-2088-0

Zhang, X., Dong, J., and Deng, F. (2019a). The Long Non-coding RNA lncRNA973 Is Involved in Cotton Response to Salt Stress. *BMC Plant Biol.* 19, 459. doi:10.1186/s12870-019-2088-0

Zhao, G., Yu, H. Y., and Liu, M. M. (2017). Identification of Salt Stress Responsive miRNAs from *Solanum lycopersicum* and *Solanum Pimpinellifolium*. *Plant Growth Regul.* 83, 129–140. doi:10.1007/s10725-017-0289-9

Zheng, M., Lin, J., and Liu, X., Histone Acetyltransferase TaHAG1 Acts as a Crucial Regulator to Strengthen Salt Tolerance of Hexaploid wheat, Plant Physiology, 2021;186: 1951–1969.doi:10.1093/plphys/kiab187

Zheng, M., Liu, X., and Lin, J. (2019). Histone Acetyltransferase GCN5 Contributes to Cell wall Integrity and Salt Stress Tolerance by Altering the Expression of Cellulose Synthesis Genes. *Plant J.* 97, 587–602. doi:10.1111/tpj.14144

Zheng, Y., Ding, Y., and SunXie, X. (2016). Histone Deacetylase HDA9 Negatively Regulates Salt and Drought Stress Responsiveness in *Arabidopsis*. *J. Exp. Bot.* 67, 1703–1713. doi:10.1093/jxb/erv562

Zhou, B. R., Feng, H., and Kato, H. (2013). Structural Model of the H1-Nucleosome Complex. *Proc. Natl. Acad. Sci.* 110, 19390–19395. doi:10.1073/pnas.1314905110

Zhou, S., Jiang, W., and Long, F. (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

Zilberman, D., Coleman-Derr, D., and Ballinger, T. (2008). Histone H2A.Z and DNA Methylation Are Mutually Antagonistic Chromatin marks. *Nature* 456, 125–129. doi:10.1038/nature07324

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.

Copyright © 2022 Singroha, Kumar, Gupta, Singh and Sharma. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Genetics

REVIEW
published: 19 April 2022
doi: 10.3389/fgene.2022.874648

Moving Beyond DNA Sequence to Improve Plant Stress Responses

*Faisal Saeed[1], Usman Khalid Chaudhry[1], Allah Bakhsh[2], Ali Raza[3], Yasir Saeed[4], Abhishek Bohra[5] and Rajeev K. Varshney[3,5,6]**

[1]Department of Agricultural Genetic Engineering, Faculty of Agricultural Sciences and Technologies, Nigde Omer Halisdemir University, Nigde, Turkey, [2]Centre of Excellence in Molecular Biology, University of the Punjab, Lahore, Pakistan, [3]Key Laboratory of Ministry of Education for Genetics, Breeding and Multiple Utilization of Crops, Oil Crops Research Institute, Center of Legume Crop Genetics and Systems Biology/College of Agriculture, Fujian Agriculture and Forestry University (FAFU), Fuzhou, China, [4]Department of Plant Pathology, Faculty of Agriculture, University of Agriculture, Faisalabad, Pakistan, [5]State Agricultural Biotechnology Centre, Centre for Crop and Food Innovation, Murdoch University, Murdoch, WA, Australia, [6]Center of Excellence in Genomics & Systems Biology, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, India

Plants offer a habitat for a range of interactions to occur among different stress factors. Epigenetics has become the most promising functional genomics tool, with huge potential for improving plant adaptation to biotic and abiotic stresses. Advances in plant molecular biology have dramatically changed our understanding of the molecular mechanisms that control these interactions, and plant epigenetics has attracted great interest in this context. Accumulating literature substantiates the crucial role of epigenetics in the diversity of plant responses that can be harnessed to accelerate the progress of crop improvement. However, harnessing epigenetics to its full potential will require a thorough understanding of the epigenetic modifications and assessing the functional relevance of these variants. The modern technologies of profiling and engineering plants at genome-wide scale provide new horizons to elucidate how epigenetic modifications occur in plants in response to stress conditions. This review summarizes recent progress on understanding the epigenetic regulation of plant stress responses, methods to detect genome-wide epigenetic modifications, and disentangling their contributions to plant phenotypes from other sources of variations. Key epigenetic mechanisms underlying stress memory are highlighted. Linking plant response with the patterns of epigenetic variations would help devise breeding strategies for improving crop performance under stressed scenarios.

Keywords: biotechnology, epigenetics, food security, abiotic stress, biotic stress, stress memory

OPEN ACCESS

Edited by:
Vinod K. K,
Indian Agricultural Research Institute (ICAR), India

Reviewed by:
Muthusamy Ramakrishnan, Nanjing Forestry University, China
Ming-Bo Wang, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia

****Correspondence:***
Rajeev K. Varshney
rajeev.varshney@murdoch.edu.au

Specialty section:
This article was submitted to Plant Genomics, a section of the journal Frontiers in Genetics

Received: *12 February 2022*
Accepted: *31 March 2022*
Published: *19 April 2022*

Citation:
Saeed F, Chaudhry UK, Bakhsh A, Raza A, Saeed Y, Bohra A and Varshney RK (2022) Moving Beyond DNA Sequence to Improve Plant Stress Responses. Front. Genet. 13:874648. doi: 10.3389/fgene.2022.874648

1 INTRODUCTION

Agriculture plays a vital role in feeding the rapidly growing world population. For fibre, fuel, and food, we usually depend on major crops such as cotton, maize, sugarcane, rice, barley, wheat, and soybean. An increase in world population day by day puts tremendous pressure on current food crop production systems (Chaudhry et al., 2021a; Junaid et al., 2021). Additionally, climate change results in several weather adversaries, and frequent disease and pest attacks threaten crop production worldwide (Raza et al., 2020). These stresses interfere with plant's physiological, biochemical, molecular, and cellular mechanisms, ultimately reducing overall growth, and production (Raza et al., 2020; Raza et al., 2021). With time, for better growth and cultivation, human beings carried out an artificial selection of thousands of plants to get desired traits in plants (Herron et al., 2020). In recent

years, new ways and tools have been discovered for the betterment of crops. For instance, affordable genetic systems for profiling of the plant genomes have led to the development of robust molecular diagnostics for rapid and precise selection of desirable crop plants (Gökçe and Chaudhry, 2020; Gökçe et al., 2021). In parallel, targeted genetic modification has been greatly benefitted by the availability of the whole genome sequence information in different crop species. For instance, genome editing techniques such as clustered regularly interspaced short palindromic repeats (CRISPR) technology using RNA-guided nucleases have facilitated the alteration of a variety of important plant phenotypes (Ma et al., 2016; Saeed et al., 2020; Dangol et al., 2021).

Feeding the growing world population would require harnessing the latest discoveries in plant epigenetics. A growing body of literature suggests that epigenetics contributes to many vital traits in different plant species. The term epigenetics refers to heritable changes in the phenotype, which are not due to a change in DNA sequence (Heard and Martienssen, 2014). In other words, epigenetics involves alterations in gene expressions that are stably transmitted from generation to generation. Plant epigenetics combines different research fields that help us understand how plants adjust their phenotypes other than modifying their DNA sequence under extreme stress conditions. The molecular processes encompassing epigenetics are DNA methylation and histone modification (Henikoff and Greally, 2016). The structure of chromatin is regulated by methylation of DNA and modification of histones, and these modifications remain crucial to the repression or activation of a gene (Ganai, 2020). The review article presents the recent advances in plant epigenetics, emphasizing plant stress response. The underlying mechanisms are discussed in the following sections.

2 MECHANISM

Rollin Hotchkiss in 1948 identified DNA methylation. After 30 years, Holliday and Pugh proposed that DNA methylation is an essential epigenetic hallmark (Holliday and Pugh, 1975). Methylation of DNA is among the key epigenetic mechanisms regulating various bioprocesses. In plants, DNA methylation is initiated by the RNA-directed DNA methylation (*RdDM*) pathway. In all sequence contexts, DNA methyltransferases DOMAINS REARRANGED METHYLTRANSFERASE 2 (*DRM2*) catalyzes the process of methylation (Zhang et al., 2018a). *RdDM* is further divided into canonical and non-canonical pathways. In the canonical pathway, RNA polymerase IV synthesizes single-stranded RNA (ssRNA) that are changed into double-stranded RNAs (dsRNAs) by involving RNA Dependent RNA polymerase 2 (*RDR2*). DICER LIKE 3 (*DCL3*) involves in the cutting off this dsRNA. This dsRNA is than converted into 24 bases of small interfering RNA (siRNAs) (Zhang et al., 2018a). The second part of this pathway depends on the transcription of non-coding RNAs (ncRNAs) by the involvement of Pol V (Liu et al., 2018). Canonical pathway initiated by Pol IV- dependent 24 nt siRNAs whereas non-canonical pathway initiated by pol II and small RNAs (sRNAs) are involved. These sRNAs are produced from dsRNAs. sRNAs consist of 21–24 nt but are cut by different DCL proteins. These sRNAs are involved in triggereing post transcriptional gene silencing (PTGS) (Martínez de Alba et al., 2013; Cuerda-Gil and Slotkin, 2016). Methylation of cytosine involves a change in cytosine to 5- methylcytosine (5-mC), and it takes place when a methyl group is transferred at 5′ positions of S-adenosyl methionine. Cytosine methylation, though dependent on plant species, ranges from 6 to 25% therefore, plants have high levels of methylcytosine (Steward et al., 2000). Methylation of DNA cytosine occurs in three sequence contexts in plants, i.e., CpG, CpHpG, and CpHpH (H known as for T, A, and C). After replication, DNA methylation of CpG and CpHpG can easily be copied because the methylation at the symmetrical CG and CHG sites can be maintained during DNA replication. The methylation at the non-symmetrical CHH sites is not maintained during replication and occurs *de novo* (Karlsson et al., 2011). Plants store this epigenetic memory at the vegetative phase under different stresses and transfer it to the next generation, which gets established during the development of germline cells. DNA is methylated at both the gene body and the promoter regions, and due to this methylation, it allows the gene to remain suppressed. Thus, lower methylation helps increase the expression of a gene (Finnegan et al., 1998). Several processes are involved in the epigenetic mechanism. These mechanisms are cytosine methylation, chromatin proteins, and post-translational modifications (Abdolhamid Angaji et al., 2010).

2.1 Active and Passive DNA Methylation and Role of Non-Coding RNAs

Different developmental, physiological, and stress stimuli are involved in regulating DNA methylation in plants. Histone and DNA methylation are inter-reliant procedures. In *Arabidopsis* mutant *met1*, CpG causes the loss of methylation of *H3K9* (Soppe et al., 2002; Tariq et al., 2003). But the loss of methylation of *H3K9* in kryptonite (KYP) did not affect the methylation of CpG site (Jasencakova et al., 2003). Consequently, it shows that methylation of CpHpG site is partly dependent on the activity of KYP due to loss of *H3K9* methylation (Jackson et al., 2002). The process of demethylation and methyltransferase both control the methylation of DNA. Demethylation follows two routes, i.e., the passive and the active. During the process of cell division, methylated DNA can vanish from the genome. If maintenance machinery present in dividing cells can be blocked. During the duplication of DNA inhibition of enzymatic activity, expression loss or elimination of DNA methyltransferase repair machinery leads toward extinction of 5-mC marks. This loss of 5-mC sites is known as passive DNA methylation (Feng et al., 2010).

Active DNA demethylation occurs by glycosylase activity by taking out the methylcytosines (Zhu et al., 2000). A single nucleotide gap is filled by demethylated cytosine with the help of the base excision repairing process (Agius et al., 2006). Many RNA molecules in the eukaryotic genome do not participate in protein production and are called non-coding RNAs (ncRNA)

(Mishra and Bohra, 2018). On a size basis, these ncRNAs are divided into two types. i.e., small ncRNA and long non-coding RNA (lncRNA). These groups were divided based on the size of the transcripts. Small ncRNAs possess less than 200 nucleotides (Bohra et al., 2021), whereas lncRNAs contain more than 200 nucleotides (Ghildiyal and Zamore, 2009; Quinn and Chang, 2016). Small and lncRNAs are important epigenetic players in regulating the plant stress response, growth, and development (El-Shami et al., 2007; Heo et al., 2013). Research has shown that lncRNAs serve as epigenetic regulators of gene expression at different stages (Karlik et al., 2019). LncRNAs work as *cis*-acting elements near the RNA synthesis sites (Zhao et al., 2020b). Further, *trans*-acting factors, they can also work away from synthesis sites (Suksamran et al., 2020). LncRNAs transcribed by polymerase II, III, IV, and V. They are further divided into five categories depending on their positions in genome near or away from protein-coding genes. The five categories are sense, antisense, bidirectional, intronic, and large intergenic lncRNA. Different lncRNAs are differentially expressed under various stresses and were suggested to play an important role (Urquiaga et al., 2021). These ncRNAs are involved in different epigenetic regulation mechanisms such as histone modification and DNA methylation (Ariel et al., 2014). Double-stranded RNAs synthesized by RNA Dependent RNA polymerases (RDRs) during this process, small interfering RNA (siRNA) arises (Song et al., 2019). siRNA are produced during transcription or transpositional reactivation of transposable elements (TE) during stress conditions (Hou et al., 2019b).

2.2 Histone Modification

Nucleosome architectures are altered in response to epigenetic changes; however, this alteration does not involve the DNA sequence (Zhao et al., 2021). Gene expression lowers or increases due to histone modification or DNA methylation (Singh and Prasad, 2021). During developmental processes or under any stresses, epigenetic changes in chromatin structure are regular and extremely active (Bhadouriya et al., 2021). Acetylation/deacetylation, SUMOylation, ubiquitin of histone proteins, phosphorylation/dephosphorylation are key processes involved in histone modification. These histone proteins are also altered chemically; they change their different physical or chemical properties (Boulanger et al., 2021). In the regulation of gene expression, the histones also release their subunits from the octamer core. These known modifications help increase DNA accessibility and speed up the selection process of binding proteins, which participate in DNA replication, transcription, or DNA repair (Pang et al., 2020). In the DNA methylation process, which occurs in the eukaryotic genome, at 50 positions of nitrogenous cytosine base, a methyl group attaches (-CH3) and forms a 5-mC (Pandey et al., 2017). These methylation processes can be asymmetrical and symmetrical; commonly CHH methylation process is known as asymmetrical, and on the other side, CHG and CG represent symmetrical methylation (Parent et al., 2021). As we discussed earlier *RdDM* is also common in plants (Singroha and Sharma, 2019). Cytosine methylation regulates gene expression by controlling the interaction of nucleic acid with transcription factors and chromatin proteins (Casati and Gomez, 2021). Patterns of DNA methylation are constant and particular to the exact cell type. These patterns are heritable and remain the same throughout life (Singh and Prasad, 2021).

2.3 Epigenetic Memory in Plants

Plant memorizes the epigenetic changes, and it helps them to adapt under biotic and abiotic stresses (Kinoshita and Seki, 2014; Crisp et al., 2016; He and Li, 2018). For instance, in *Arabidopsis thaliana*, two important factors that memorize during stresses are modification of histone and *HSFA2* (Heat shock factors). When a plant faces heat shock, the level of *H3K4* (H3 lysine K4 methylation) methylation remains high at least for 2 days. This process is also linked with transcriptional memory. The expression of heat stress response and transcriptional heat shock memory is dependent on the accumulation of *H3K4* methylation and *HSFA2*. *REF6*, known as RELATIVE OF EARLY FLOWERING 6, exhibits a positive response and transfers as long-term memory of epigenetic changes in *A. thaliana* (Liu et al., 2019a). *LSD-1* (Lysine-specific histone demethylase-1) in wheat recorded upregulation during heat stress as compared to normal plants. It is linked with modification of histone in the generation of transgenerational thermotolerance by heat priming. These changes induced by heat shock, transgenerational epigenetic memory, or changes in phenotype can be carried out at least two to three generations (Suter and Widmer, 2013; Zhonga et al., 2013). Priming of organismal stress response explains the events by which transient stimulus alters plant for future exposure to stress (Conrath et al., 2015). The term priming basically referred to immunity against pathogens but was later applied to abiotic stress. Priming is a reversible event because it only changed the phenotypical appearance of plant and does not change genetic makeup (Hilker et al., 2016). There are still many questions related to epigenetic memory. The specificity of stability of DNA and choramtin and their existence during mitosis and upkeep of memory. The mechanisms directly linked to chromatin changes which is further linked to transcriptional responses when plant faces stress are still not clear (Asensi-Fabado et al., 2017). A plant that once faces any harmful or stress conditions can recover from that stress, and epigenetic memory helps its future survival under stress conditions. As illustrated in **Figure 1**, the plant activates the epigenetic stress memory against future stresses, and the plant will remain protected.

3 DYNAMICS DURING BIOTIC STRESSES

On exposure to biotic stress, the defense machinery evokes the immune system, such as basal defense machinery and pathogen-associated molecular patterns (Muthamilarasan and Prasad, 2013). DNA methylation changes in plants as a defense response against biotic stress. The role of DNA methylation has been reported in *A. thaliana* (Dowen et al., 2012). In *A. thaliana*, the *met1* and *ddc* mutants could not produce infectious symptoms by *Pseudomonas syringae* pv. *tomato*

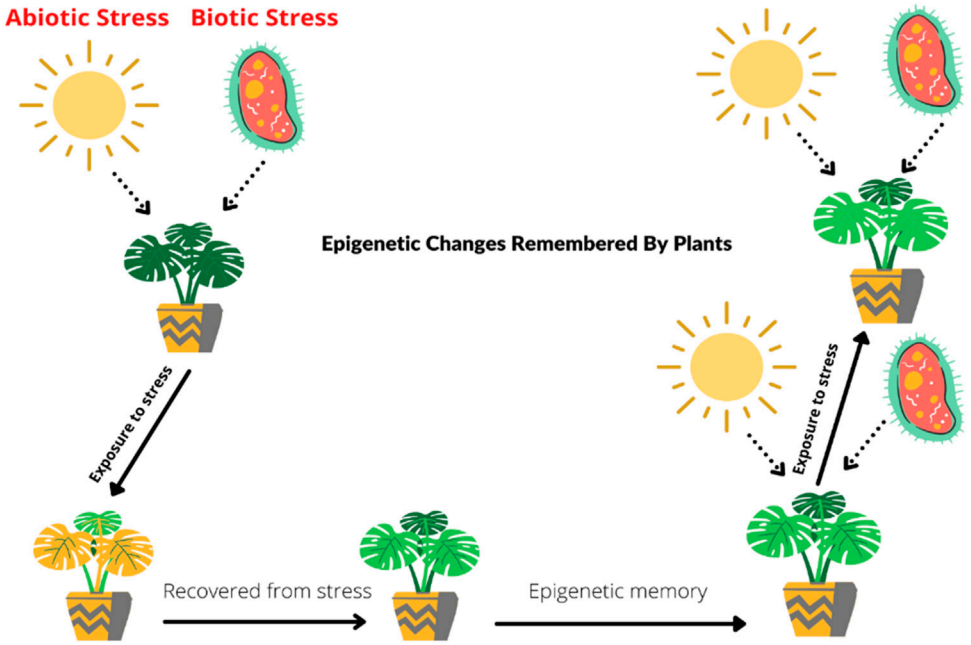


FIGURE 1 | Mechanisms underlying epigenetic memory in plants during stress. Plants' epigenetic memory helps protect them from different stresses. Whenever a plant faces stress regardless of its biotic or abiotic nature, it starts recovery against stress, and the plant epigenetic stress memory stores that information. Due to this stored memory, stress does not affect the plant on subsequent exposures.

DC3000 due to the elimination of methylation capability of cytosine. Another gene named *ELP2* initiates DNA methylation in *Arabidopsis,* and pathogen-altered methylation of DNA takes place (Dowen et al., 2012; Yu et al., 2013). Akimoto and colleagues artificially reduced the DNA methylation at the promoter region of *R* gene (*Xa 21G*) of rice (Akimoto et al., 2007). During pathogen infection, the influence on the expression of defense-related genes was shown, caused by hypomethylation of DNA. When plants are infected with *Xanthomonas oryzae pv. oryzae*, rice mutant lines with constitutive silencing of *Xa21G* gene showed more resistance as compared to wild type (Akimoto et al., 2007). With advancements in molecular biology, scientists found new pathways and events in virus-infected plants. Plants generally employ siRNA-mediated DNA methylation against viruses or biotic stresses; in this mechanism, the plant methylated its genomic parts (Emran et al., 2012).

3.1 Biotic Stress-Related to the Epigenetic Configuration

Elucidation of the epigenetic alterations under biotic stress helps understand the plant-pathogen interaction (Zogli and Libault, 2017). In the analysis of plant-microbe relation, harmful *pst* introduced in Arabidopsis and DNA methylation changes occurred in all sequences of context. On the other side, when a non-harmful strain (i.e., bacterial strain) was introduced, only changes in CHG and CG methylation were detected (Dowen et al., 2012). Changes in methylation were common proximal to genes related to defense, and their activation correlated with transcription, so they have a role in reaction to pathogens (Yu et al., 2013).

Limited information is available on plant epigenetic effects created by fungal pathogens or oomycetes (Crespo-Salvador et al., 2018). *Arabidopsis* roots infected by cyst showed a huge change in DNA methylation and small RNA. In general, at the stage of infection, dynamic shifts take place (Joseph et al., 2021). However, DNA methylation change can be associated with several regions; transcriptional and epigenetic changes are related to each other and affect genes responsible for defense (Hewezi et al., 2017). In non-model crop species, understanding of epigenetic modifications is often limited, including patterns of DNA methylation (Herrera et al., 2016). In *Brassica rapa*, changes in DNA methylation were related to differences in floral morphology and less attraction of pollinators (Kellenberger et al., 2016). Detailed knowledge and high-resolution analysis will facilitate understanding plant-microbe interaction and underlying epigenetics changes (Richards et al., 2017). Plant-to-plant interactions facilitated by microbiota related to roots or allelochemicals obtained from plants may have some effect on chromatin arrangement (Venturelli et al., 2015).

3.2 Plant Epigenetic Influences on Biotic

Plants can be affected by neighbouring plants, microbes, and herbivores. Due to this, epigenetic changes affect plant phenotypes and plant's interactions with other ogranisms (Latzel et al., 2012). Advances in plant epigenetics have deepened our understanding of the plant response against biotic stresses (Marfil et al., 2009). *AGO4* mutants of *A. thaliana* lacking methylation are susceptible to *pst* (Dowen

et al., 2012). Similarly, overexpression of histone lysine demethylase enhanced resistance against blight in rice (Li et al., 2013). Relation between epigenetics and plant interaction with fungal pathogens is evident from enhanced susceptibility *A. thaliana* mutants (*dml1 dml2 ros1*) (Le et al., 2014). Post-translation histone modifications are also included in defense against pathogens (Xia et al., 2013). Evidence of inclusion of epialleles in biotic relations was discovered from mutants of *Arabidopsis* (Johannes et al., 2009; Reinders et al., 2009). Furthermore, these lines were also used to identify epigenetic QTLs that established a connection between epigenetic modifications and phenotypic variability (Latzel et al., 2012).

4 PLANT EPIGENETIC CHANGES AND REGULATION DURING ABIOTIC STRESSES

Climate change is crucial concerning, the adaptation of the crops to the changing climate scenarios as well as the growth of future crops to ensure food security. Environmental changes trigger drought, salt, heat, and cold stresses (Raza, 2020; Raza et al., 2020). Any fluctuation in temperature negatively affects plants growth and development, resulting in poor yield (Raza, 2021; Raza et al., 2021). Over the past few decades, several studies have explained the mechanisms of abiotic stresses, but reports on epigenetic regulation are still limited (Raza et al., 2021). Cold/chilling stress influences plant metabolic enzyme activities, responsible for gene expression (Raza et al., 2020; He et al., 2021; Raza et al., 2022). Cold stress is another important abiotic stress that retards plant growth and yield. To mitigate the risks associated with cold stress, plants have evolved signaling system that stimulates the expression of cold-stress-related genes. Plant response to cold stress is well-characterized, and research highlights the profound role of C- REPEAT BINDING FACTOR (*CBF*)-COLD RESPONSIVE (COR) pathways. Cold stress is reported to stimulate transcription factors (*TFs*) expression, which includes *CBF* family proteins. The *TFs* bind to the promoter region of downstream *COR* genes that activate its gene expression (Zhu, 2016). A recent study in *A. thaliana* described that chromatin remodeler PICKLE (*PKL*) is involved in *CBF*-dependent cold stress response (Yang et al., 2019). Histone methylation and histone modifications play a significant role against cold stress. For example, acetylation of histone is found enriched in several cold-responsive genes (Park et al., 2018). It is regulated dynamically by histone acetyltransferases (*HATs*) and histone deacetylases (*HDACs*) (Li et al., 2021). In recent research, *A. thaliana* plants were exposed to cold stress, and overexpression of histone deacetylase 2D (*HD2D*) exhibited lower lipid peroxidation with decreased accumulation of malondialdehyde contents that eliminated the oxidative burst (Han et al., 2016). Furthermore, plants under the influence of cold stress showed induction of histone acetylation in the promoter region of *COR47* and *COR15A* (Pavangadkar et al., 2010). The *H3K4me3* and *H3K27me3* have been reported to be involved in response to vernalization. The study elucidated their regional regulation and contributions to epigenetic memory. The vernalized *Brachypodium distachyon* induced epigenetic changes that regulate multiple genes to coordinate biological processes (Huan et al., 2018).

A global rise in temperature has attracted the attention of plant scientists to improve crop's adaptation to future scenarios (Jha et al., 2014; Haider et al., 2022). In response to heat stress, Heat Shock Transcription Factor A1s (*HSFA1s*) are the main *TFs* controlled by phosphorylation or dephosphorylation and protein-protein interactions. A temperature higher than required for the normal growth functioning of a plant instantly disrupts its photosynthetic machinery with the absorption of increased light. It damages the photosystem II, thylakoid protein phosphorylation. Heat stress stimulates hyper-phosphorylation likewise activates heat shock *TFs*. Acetylation is the epigenetic modification that alters the *H2A* and *H3* histones, two important players associated with heat stress response in plants. For instance, actin-related protein 6 (*ARP6*) in *Arabidopsis* is reported to regulate gene expression. It encodes the *SWR1* complex that is necessary for the insertion of *H2A.Z* histone in nucleosomes as a replacement for *H2A* histone (Nie and Wang, 2021). It is also reported as an indispensable event for temperature sensing. Moreover, acetylation of *H3K56* is related to accumulating RNA polymerase II and activates *TFs* with exposure to heat stress (Haider et al., 2021). Research has demonstrated that the *RdDM* pathway and histone dynamics are involved in the responses against heat stress (Lämke et al., 2016; Yang et al., 2018). The heat shock proteins (*HSPs*) are primarily involved in conferring tolerance against heat stress in plants, regulating folding and unfolding of proteins (Singh et al., 2016). The heat stress stimulates the persistent expression of *H3K4me3* and *H3K9Ac* on *HSP18*, *APX2*, and *HSP70* genes (Lämke et al., 2016). In *A. thaliana*, the member of the F-Box family protein such as suppressor of *DRM1 DRM2 CMT3* (*SDC*) proteins facilitates the degradation of the protein. Heat stress for a prolonged period induces transcriptional expression of a subgroup of genes, which ultimately assist the plant in recovering from heat stress (Popova et al., 2013; Sanchez and Paszkowski, 2014). The *SDC* gene targets the *RdDM* pathway that can be silenced epigenetically under normal growth conditions. However, its activation under heat stress suggests a transcriptional response, which overcomes the silencing effect of *RdDM* at some loci.

Like heat, salt stress is another key challenge to global agriculture (Jha et al., 2014). The plant faces salt stress with the higher accumulation of salt contents, mainly increased sodium ions (Na^+) content that cause ionic toxicity. The growth and development of plants are impaired following secondary oxidative stress (Chaudhry et al., 2021b; Hafeez et al., 2021). The participation of Histone acetyltransferase (*HAT*) in regulating salt tolerance has been elucidated in *A. thaliana* (Zheng et al., 2019). Higher uptake and accumulation of Na^+ in the GENERAL CONTROL NONDEREPRESSIBLE 5 (*GCN5*) mutant as compared to wild-type plants impaired the growth of the mutant in response to salt stress. Additionally, *GCN5* can bind with cell wall synthesis genes such as CHITINASE-LIKE 1 (*CTL1*), and *MYB54*. Notably, lower *H3K9ac* and *H3K14ac* concentrations in mutant due to salt

stress suggested that the *GCN5* is a conserved epigenetic regulator (Zheng et al., 2019). A recent study on *GCN5* in wheat has shown that target genes responsible for producing ROS species such as H_2O_2 (Zheng et al., 2021). The calcium ions (Ca^{2+}) CALCINEURIN B-LIKE PROTEIN (*CBL*) CBL INTERACTING PROTEIN KINASE (*CIPK*) component performs an essential role in regulating cellular ionic homeostasis (Zhu, 2016). Higher Na^+, lower K^+, excessiveness of Mg^{2+}, and higher pH levels stimulate cytosolic Ca^{2+} signaling for the activation of SALT OVERLY SENSITIVE3 (*SOS3*)-*SOS2*, *CBL2/3-CIPK3/9/23/26*, *CBL1/9-CIPK23*, and *SCaBP1-CIPK11/14* that causes phosphorylation and regulation of the activity of H^+ ATPase, Mg^{2+} transporter, *Arabidopsis* K^+ TRANSPORTER (*AKT1*, K^+ channel), and *SOS1* (Na^+/H^+ antiporter) (Zhu, 2016). HIGH-AFFINITY K^+ CHANNEL 1 (*HKT1*), which facilitates Na^+ influx in plants, is vital transporter for coordinating with the *SOS* pathway to confer salt tolerance (Rus et al., 2001). In *A. thaliana* wild-type plants, a small RNA target region was identified at approximately 2.6 kb upstream of *HKT1* that was reported to be highly methylated (Baek et al., 2011). It was reported that a lower DNA methylation level in *RdDM* mutant *rdr2* led to an enhanced expression of *HKT1*, thus highlighting the role of *RdDM*-mediated regulation of gene expression (Miryeganeh, 2021). Another study in wheat revealed that salt stress-induced cytosine methylation caused suppression of *TaHKT2* expression in roots and shoots of both tolerant and sensitive genotypes (Kumar et al., 2017). The *TFs* induced by salt stress include *MYB74* of the *R2R3-MYB* family. The *MYB74* promoter is extremely methylated due to the *RdDM* pathway, and in salt stress, 24-nt siRNA levels and DNA methylation were almost imperceptible at *MYB74* is accompanied by the upregulated expression of *MYB74* (Xu et al., 2015).

Histone dynamics have been associated with drought stress response in plants (To and Kim, 2014). ABA-mediated signaling, playing an important role in drought stress in plants, is influenced by epigenetic modulation caused by either DNA methylation or histone acetylation. For example, analysis of *ABA*-deficient mutant maize (vp10) revealed differential methylation of several stress-responsive genes and *TE*. The key enzyme involved in the synthesis of *ABA* is NINE CIS-EPOXYCAROTENOID DIOXYGENASE 3 (*NCED3*) (Nambara and Marion-poll, 2003). Plant acclimatization to drought stress improved following deposition of *H3K4me* in the *NCED3* gene that caused higher gene expression (Ding et al., 2011). Moreover, the elevated expression level has been noted in the genes *RAP2.4*, *RD29A*, *RD29B*, and *RD22* in response to drought stress (Takahashi et al., 2000). The increased levels of *H3K4me3* and *H3K9Ac* in the promoter regions of *RAP2.4*, *RD22*, *RD29A*, and *RD29B* also contributed to the activation of genes expression. It was suggested that histone marks in response to drought stress also varied with the intensity of stress. As *H3K4me3* and *H3K9Ac* levels were higher with the exposure to severe drought in contrast to mild stress conditions (Kim et al., 2012). In *A. thaliana*, lower deposition of *H3K27me3* in the gene body region of drought-associated *TFs* resulted in resistance to drought (Sebastian Ramirez-prado et al., 2019). The *H3K27me3* reader protein is LIKE HETEROCHROMATIN PROTEIN 1 in the *PRC1* complex (Mozgova and Hennig, 2015). Moreover, drought-stressed plants had modified DNA methylation levels that ultimately altered expression levels of several drought-responsive genes (Liang et al., 2014). The miniature inverted-repeat transposable element (*MITE*) inserts in the promoter region of the *NAC* gene suppress the expression with the deposition of *RdDM* and *H3K9me3* (Mao et al., 2015). Likewise, zinc finger gene *ZmMYB087* is related to the metabolism group *MYB* transcription factor to regulate the biosynthesis of the secondary cell wall. CW-type zinc finger protein participates in the methylation of histone *H3* that is essential for epigenetic memory (Sallam and Moussa, 2021).

Optimal nutrient supply is essential for plant growth and development, but the excessive nutrients in the soil impair it by causing nutrient stress (Salim and Raza, 2020). Higher nitrogen in the soil down-regulated the expression of root nitrogen transporter, *NRT2.1*. The gene repression necessitates the involvement of the HIGH NITROGEN INSENSITIVE 9 (*HNI9*) in depositing *H3K27me3* on the *NRT2.1* gene (Widiez et al., 2011). Iron homeostasis in *Arabidopsis* was negatively regulated by *PRMT5*-mediated *H4R3* symmetric dimethylation (*H4R3sme2*) (Fan et al., 2013). The *PRMT5* linked with the *bHLH* genes, i.e., *AtbHLH38* and *AtbHLH100*, for the symmetrical demethylation of *H4R3* with no change in its gene expression (*PRMT5*) (Fan et al., 2013). Histone acetyltransferase *GCN5* is involved in regulating iron homeostasis by FERRIC REDUCTASE DEFECTIVE 3 (*FRD3*) (Xing et al., 2015). The *GCN5* can directly bind to the promoter region of iron-associated genes, which includes *FRD3*, to modulate the acetylation levels of *H3K6* and *H3K14* (Xing et al., 2015). The *H3K4me3* acetylation and histone variant *H2A.Z* have a key role in response to phosphorus-deficient soil conditions. The protein PHD ALFIN-LIKE 6 (*AL6*) binds to the *H3K4me3* mark that influences the maturation of transcript and stability of vital genes necessary for elongation of root hairs (Chandrika et al., 2013). Additionally, histone modifications in response to deficient phosphorus showed vast remodeling of DNA methylation (Secco et al., 2015). Gene expression levels of DNA methylase were induced on exposure to limiting phosphorus conditions (Yong-Villalobos et al., 2015). **Table 1** enlists various crops where epigenetic mechanisms controlling response to various stresses have been elucidated.

5 DEVELOPING EXPERIMENTAL POPULATIONS TO DISENTANGLE EPIGENETIC EFFECTS FROM DNA SEQUENCE VARIATION

Development of experimental populations based on the two genotypes that have little DNA sequence polymorphism but show extensive variation in their methylation patterns could help greatly to overcome the confounding effect of DNA methylation and DNA sequence variation, exemplified by epigenetic recombinant inbred lines (epiRILs) in *Arabidopsis*

TABLE 1 | Stress-related epigenetic mechanisms for improved crop development under stress conditions.

Crop	Mechanism	Reference
Drought stress		
Rice	DNA methylation at a specific site	Wang et al. (2011)
Barley	Excessive accumulation of *H3* and loss of *H3K9me2*	Temel et al. (2017)
Maize	Enrichment of *H3K4me3* and *H3K36me3*	Xu et al. (2017)
Maize	Modified dynamics of *H3K4me3* and *H3K9ac*	Forestan et al. (2018)
Soybean	Upregulated isomiRNAs	Sosa-Valencia et al. (2017)
Pea	Cytosine hypermethylation	Labra et al. (2002)
Cotton	Histone modification	Chen et al. (2019)
Salt stress		
Wheat	Increased cytosine methylation of *HKT* genes	Kumar et al. (2017)
Rice	Differentially methylated regions of DNA	Ferreira et al. (2019)
Rice	Demethylation in the promoter of *OsMYB91* gene and modification of histone	Zhu et al. (2015)
Temperature stress		
Soybean	Cytosine hypomethylation	Hossain et al. (2017)
Wheat	Higher histone demethylation of several genes	Wang et al. (2016)
Maize	Modification of *H3K4me2* and *H3K9ac*	Hou et al. (2019a)
Maize	Higher acetylation of histone and reduction of *H3K9me3*	Wang et al. (2014)
Maize	Decreased acetylation of histone	Hu et al. (2011)
Maize	Higher accumulation of *H3K9ac*	Hu et al. (2012)
Mustard	Non-coding RNA mediated regulation	Bhatia et al. (2020)
Rice	Methylation of promoter region	Guo et al. (2019)
Biotic stress		
Potato	BABA primed histone modification against *Phythophtra infestans*	Meller et al. (2018)
Tomato	Methylation in cytosine residue to improve resistance against Tomato spotted wilt virus	Werghi et al. (2021)
Olive	Methylation to improve resistance against *Verticilium dahliae*	Crespo-Salvador et al. (2020)
Tomato	Improved resistance against pathogen creating mutants of Histone domain	Bvindi et al. (2022)

based on Columbia (with wild type DDM1 allele) and Col-ddm1 mutant (Johannes et al., 2009). Similarly, *Arabidopsis* epiRILs showed distinct phenotypic variations with altered resistance stress (Lloyd and Lister, 2022). Such experimental populations laid a foundation to study epigenetic contributions to novel phenotypic variations. The combined studies of epiRILs and natural accessions suggested that epigenetic diversity is an important component of functional biodiversity. Additionally, comprehensive analysis of genes responsible for epigenetic machinery in epiRILs led to the mechanisms engaged in heritable DNA methylation and the resulting impact on plant phenotype (Johannes et al., 2009). DNA methylation is inherited in a stable Mendelian fashion in some genomic regions (Colomé-Tatche et al., 2012). The epiRIL population of *ddm1*, showed stable inheritance of differentially methylated regions for several generations without disrupting the DNA sequence and established a role for epigenetic quantitative trait loci (epiQTL) in phenotypic manifestation (Zhang et al., 2013).

The epiRILs of Arabidopsis differ for DNA methylation but show very minute changes in their DNA sequences. The development of epiRILs is similar to the creation of classic RILs, which includes the crossing of two genetically divergent parents and subsequently, inbred lines are established. DNA methylation exhibited developmental phenotypic changes that suggested the suppression of phenotypic plasticity, and it can be assessed with the construction of epiRILs (Bossdorf et al., 2010). The F2 progenies can be screened for the homozygosity of alleles to confirm the function of DNA methylation machinery and subsequently maintain the epigenetic chimeric chromosome developed by recombination at the F1 meiosis stage. The epiRILs harbor phenotypic variations for plant morphological characteristics, growth rate, and abiotic stress responses. Studies on epiRILs showed that epigenetic variations contributed to the functional diversity that had a similar impact on populations as noticed in genetic diversity. Similarly, higher epigenetic changes resulted in improved plant phenotypes translating to higher productivity and resilience in populations (Latzel et al., 2012). The heritable phenotypic variations among epiRILs in response to drought stress revealed phenotypic plasticity with the significant variation in root: shoot ratio that has the potential for developing stress-resilient plants (Zhang et al., 2013). The crop improvement is based on the stable transmission of epialleles in inheritance. With the introduction of methods for creating epiRILs in crop plants, the generation of epimutagenesis and engineered epigenetic changes would contribute to bridging the gaps for harnessing epigenetic variations for trait improvement to confer stress tolerance.

6 DIFFERENT METHODS OF DETECTION OF EPIGENETIC CHANGES IN THE GENOME

In recent years, several methods have been discovered to profile large-scale epigenetic modifications. With rapid progress in the field of biological sciences now, these methods are accurate,

precise and affordable. Sodium bisulphite seqeucning and methylated immunoprecipitation sequencing (MeDIP-seq) in combination with latest sequencing technologies like single-molecule real-time sequencing (SMRT-seq) and Nanopore sequencing (Liang et al., 2019) are used genome-wide methylation profiling.

6.1 Sodium Bisulfite Method

DNA methylation can be detected by sodium bisulfite, which converts demethylated cytosines into uracil (Kushwaha et al., 2016). However, it did not affect the 5-mC (Papanicolau-Sengos and Aldape, 2022). The sodium bisulfite method of detection of epigenetic changes is a standard technology for the detection of 5-mC due to its capacity to provide huge information about methylated DNA segments (Zhao et al., 2020a). Different methods after conversion are also used for further analysis, such as PCR analysis, sequencing of the methylated genome is widely used. It can provide information about even a single methylated nucleotide (Matzke and Mosher, 2014). Frommer and colleagues pioneered DNA methylation analysis based on bisulfite genomic sequencing. In this method, firstly, DNA is denatured by using alkali and treated with bisulfite. After that, in the second part, the region of interest was amplified by PCR using bisulfite-specific primers (Frommer et al., 1992). On the other sides, this sequencing method has several disadvantages, and we cannot differentiate between 5-mC and 5-hydroxymethyl cytosine. Unsuccessful reaction of bisulfite may cause a failure to covert entire demethylated cytosines to uracil, thus leading to false-positive results (Fraga and Esteller, 2002). When template DNA is treated with bisulfite, it is difficult to design primers for multiplex PCR reactions (Callinan and Feinberg, 2006). To overcome these problems, a modified method of bisulfite sequencing, reduced representation bisulfite sequencing (RRBS), was proposed (Meissner et al., 2005). This method gained traction among researchers due to its low cost and coverage of all regions. This method only required specific genomic sequences (Hahn et al., 2014). This can help detect different levels of methylation and can analyze the specific genes under stress and control conditions. RRBS technique provides useful data for further methylation measurements. Examination of wild tobacco plants under salt and low temperature analyzed by bisulfite sequencing revealed demethylation at GC sites in coding regions, and demethylation was found in promoter regions (Rehman and Tanti, 2020). In 2019, Liu and colleagues demonstrated two methods for the detection of 5-mC and 5-hydroxymethyl cytosine. One is ten-eleven translocation (TET)-assisted pyridine borane sequencing (TAPS) (Liu et al., 2019b) and its modified version long-read TET-assisted pyridine borane sequencing (lrTAPS) (Liu et al., 2020). With the help of pyridine borane, TAPS uses TET oxidation of 5-methyl cytosine and 5-hydroxymethyl cytosine to produce 5-carboxylcytosine that ultimately reduces to dihydrouracil (DHU). After that, the PCR reaction converts DHU to thymine and at last recognizes 5-mC and 5-hydroxymethyl cytosine (Liu et al., 2019b).

6.2 Methylation Detection Based on Antibody

Methylated DNA immunoprecipitation (MeDIP) is one of the sensitive and widely used methods for methylome studies of plants. It is a purification method of gDNA fragments that contain methylated sites (Zhang et al., 2006). In this method, the fluorescent dye uses chemical derivatives of normal and modified deoxyribonucleotides to covalently bind to 6-methylated adenine (6 mA). Further, this 6 mA was quantified by using Capillary Electrophoresis with Laser-induced fluorescence (CE-LIF). The quest for alternatives like light sources instead of expensive lasers would impart cost-effectiveness to this method (Nguyen and Kang, 2019).

6.3 Other Methods of Methylation Detection

Bisulfite sequencing can only detect changes at C, but it cannot be distinguished between 5mC and 5hmC, and also it needs a reference genome (Flusberg et al., 2010). There are several methods of methylation detection, such as HPLC-MS/MS that can detect even the low to the lowest amount of 6-methyl adenine in plant genome (Huang et al., 2015). Another method, dot blot assay, is used with specific antibodies applied, but this approach has a limit of detection (Liang et al., 2019). These two methods are widely used for the detection of 6-methyl adenine, but these approaches are unable to reveal the location of methylated sites.

6.4 Different Next-Generation Sequencing Tools

The nucleotide sequencing technologies have evolved rapidly over the past 20 years (Varshney et al., 2007; Bohra et al., 2020; Varshney et al., 2021). First-generation sequencing technologies invented by Sanger and Maxam Gilbert revolutionized the molecular biology field (Sanger et al., 1977). In the mid-90s first Sanger sequencing platform (ABI 370) was made commercially available (Watts and MacBeath, 2003). The Sanger sequencing is the preferred technique for plant molecular biologists. It deals with the DNA genome of plants to the sequence. The ABI 370 xl DNA sequencer has maximum accuracy of up to 99.99%. It can generate reads as small as 1.9 kb to as long as 84 kb, and in 3-h run, it can create up to 300–400 bp reads (Liu et al., 2014). The major drawback associated with this sequencing technology is its time and resource-intensive nature. These bottlencks were overcome by the introduction of the second-generation of sequencing or next generation seqeucning (NGS) in 2005. This NGS technologies not only cut down the cost of sequencing but also produce million of reads in less time (Kchouk et al., 2017). Further development of the third-generation sequencing methods overcame many problems related to the second-generation sequencing methods, including sample preparation, product amplification, and time. The third-generation sequencing technique SMRT-seq (Single-molecule real-time sequencing) offers exact sequences and measures the nucleotide energy rate during sequencing. This sequencing method detects DNA changes at a single base. The NGS-based methologies have helped greatly to study

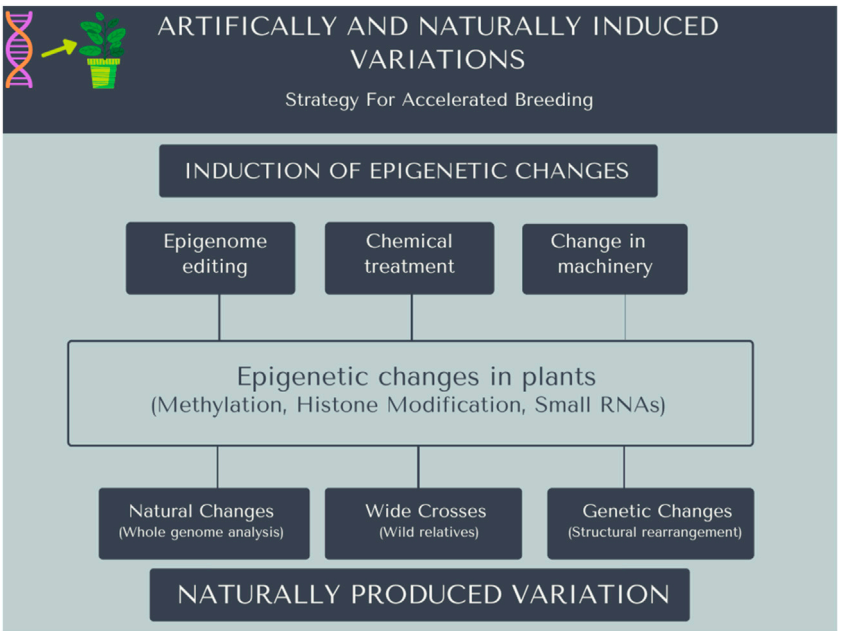


FIGURE 2 | Artificially or naturally induced mutations in the plant genome can be helpful for accelerated breeding. Epigenetic changes can be induced in the plant genome using different methods, i.e., naturally or artificially. Artificially it can be induced using editing techniques, chemically treated plants to create mutations, and alter the machinery using different molecular approaches. Wide crosses of plants and naturally occurring changes in the genome in nature can be useful for accelerated breeding.

genome-wide methylation patterns, as exemplified in *Arabidopsis thaliana* (Liang et al., 2018), rice (Zhang et al., 2018b), and fig (*Ficus carcia*) (Usai et al., 2021). The recently developed PacBio sequencing method requires only 5 h from sample to produce reads, and also it reduces costs. The only disadvantage of this system is a higher rate of error (14%) (Chin et al., 2016; Kchouk et al., 2017). The *in situ* sequencing (ISS) represents a fourth-generation platform that directly sequences the nucleic acids with a higher accuracy (Mignardi and Nilsson, 2014; Ke et al., 2016).

7 COMBINING EPIGENETICS AND GENE EDITING TO IMPROVE STRESS RESPONSE

Biotic and abiotic stresses deteriorate the yield and quality of crop plants. Conventional breeding approaches are increasingly constrained to rapidly develop plant varieties having adaptation to changing climatic conditions. It is the need of the hour to utilize different approachesfor accelerated crop improvement (Watson et al., 2018). The development of stress-resilient crops requires plant breeding approaches to exploit variations beyond DNA sequence (Fiaz et al., 2019). Plant stress memoryrelated genes or factors can be used by genome editing techniques for a better understanding of the regulatory mechanisms underlying stress response. Plant response to various stresses (biotic and abiotic) is known to involve several genes (Raza et al., 2021). Epigenetic changes can be induced in plants, and these can be used as a helpful strategy for the improvement of crops and can accelerate the breeding process, as shown in **Figure 2**. Nevertheless, these epigenetic changes can be induced artificially by editing of epigenome or chemically treated to create mutation (**Figure 2**). The CRISPR/Cas9 protein is successfully utilized as a dCas9 to modify epigenetic changes. The dCas9 protein is attached with the epigenetic modifier to targeted modifications that result in altered gene expression (Adli, 2018).

After the modification in Cas9 protein and the emergence of CRISPR/dCas9 technology, scientists start regulatory and reporter genes to investigate the ability of dCas9 for epigenome editing (Zezulin and Musunuru, 2018). It was reported that the plant genome was modified by methylation and demethylation at target DNA and resultantly developed late-flowering phenotypes (Adli, 2018). It has been reported that the fusion of repressor domains (such as KRAB/SID) with dCas9 led a significant improvement of transcriptional repression (Gilbert et al., 2013). Furthermore, the fusion of active transcriptional domain *VP16/VP64* activates the expressional level of the gene of interest that permits the screening of stress-tolerant genotypes (Miglani et al., 2020). Nowadays, synthetic regulation of transcriptional modifications has been successfully considered for traits improvement by the CRISPR technique (Piatek et al., 2015). This technique has been used for precise and spatial modification by avoiding undesirable pleiotropic effects. For instance, the promoter of *OsRAV2* gene, a *TF* responsible for salt stress, was genetically manipulated by CRISPR/Cas technique. The impaired growth of mutant lines by salt stress confirmed the important role of *GT1* during normal plant growth and development (Duan et al., 2016). Conventional plant

breeding has served humanity and this planet for a century and has always helped produce high-yield crops and fulfill the requirement of food for humans (Hickey et al., 2019). To this end, emerging plant breeding technologies such as epigenome editing can help to achieve food security targets by protecting plants from biotic or abiotic stresses.

8 CONCLUSION AND PROSPECTS

Understanding the interactions between epigenetics and plant stress response has huge potential to develop modern crops adapted to future climatic conditions. Unlike abiotic stress, epigenetic regulation of biotic stress tolerance of plants is more complex, and it remains more challenging to carry out experiments for different species *in vitro*. In the future, biotechnologists, ecologists, and molecular biologists may collaborate to find out the mechanism involved between epigenetics and stress response. Methylome profiling has remained a challenge owing to the cost and technological considerations. However, recent advances in high-throughput assays have helped relieve this bottleneck. Interdisciplinary research efforts may help to sequence the genomes of diverse accessions and also play a vital role in establishing genomic tools to find out the epigenetic changes to stresses (Schrey et al., 2013). The growing information on whole genomes and gene content would inspire future researchers to decipher the epigenetics-mediated response of plants to a variety of biotic stresses. The foremost objective will be to survey epigenetic variation and quantify its effects on phenotypes in different crops (Kawakatsu et al., 2016). Furthermore, the candidate loci for the target traits can be identified, and epigenetic profiling may be precisely targeted to the candidate regions (Aller et al., 2018). A key challenge to decipher the role of epigenetic regulation is the application of stress treatments in controlled conditions. In the natural environment, plants often deal with multiple stresses at one time. With the rapid advances in genome-wide methylation profiling and gene editing techniques, we envisage a better understanding of the epigenetic changes controlling plant stress response, which will be crucial to precisely manipulate the epigenetic regulatory mechanism for improved crop performance.

AUTHOR CONTRIBUTIONS

FS and ABa conceived the idea. FS, UC, AR, ABa, and YS contributed to writing and literature search. FS, AR, ABo, and RKV proofread and edited the manuscript. All authors have proofread and approved the final version of the manuscript.

FUNDING

RKV is thankful to Bill & Melinda Gates Foundation, United States (The Tropical Legumes III project, OPP1114827), and Australia-India Strategic Research Fund from the Department of Biotechnology, Government of India, for supporting this work in part.

REFERENCES

Abdolhamid Angaji, S., Hedayati, S. S., Poor, R. H., Poor, S. S., Shiravi, S., and Madani, S. (2010). Application of RNA Interference in Plants. *Plant Omics* 3, 77–84.

Adli, M. (2018). The CRISPR Tool Kit for Genome Editing and beyond. *Nat. Commun.* 9. doi:10.1038/s41467-018-04252-2

Agius, F., Kapoor, A., and Zhu, J.-K. (2006). Role of the Arabidopsis DNA Glycosylase/lyase ROS1 in Active DNA Demethylation. *Proc. Natl. Acad. Sci. U.S.A.* 103, 11796–11801. doi:10.1073/pnas.0603563103

Akimoto, K., Katakami, H., Kim, H.-J., Ogawa, E., Sano, C. M., Wada, Y., et al. (2007). Epigenetic Inheritance in rice Plants. *Ann. Bot.* 100, 205–217. doi:10.1093/aob/mcm110

Aller, E. S. T., Jagd, L. M., Kliebenstein, D. J., and Burow, M. (2018). Comparison of the Relative Potential for Epigenetic and Genetic Variation to Contribute to Trait Stability. *G3 Genes, Genomes, Genet.* 8, 1733–1746. doi:10.1534/g3.118.200127

Ariel, F., Jegu, T., Latrasse, D., Romero-Barrios, N., Christ, A., Benhamed, M., et al. (2014). Noncoding Transcription by Alternative Rna Polymerases Dynamically Regulates an Auxin-Driven Chromatin Loop. *Mol. Cel* 55, 383–396. doi:10.1016/j.molcel.2014.06.011

Asensi-Fabado, M.-A., Amtmann, A., and Perrella, G. (2017). Plant Responses to Abiotic Stress: The Chromatin Context of Transcriptional Regulation. *Biochim. Biophys. Acta (Bba) - Gene Regul. Mech.* 1860, 106–122. doi:10.1016/j.bbagrm.2016.07.015

Baek, D., Jiang, J., Chung, J.-S., Wang, B., Chen, J., Xin, Z., et al. (2011). Regulated AtHKT1 Gene Expression by a Distal Enhancer Element and DNA Methylation in the Promoter Plays an Important Role in Salt Tolerance. *m* 52, 149–161. doi:10.1093/pcp/pcq182

Bhadouriya, S. L., Mehrotra, S., Basantani, M. K., Loake, G. J., and Mehrotra, R. (2020). Role of Chromatin Architecture in Plant Stress Responses: An Update. *Front. Plant Sci.* 11, 603380. doi:10.3389/FPLS.2020.603380/FULL

Bhatia, G., Singh, A., Verma, D., Sharma, S., and Singh, K. (2020). Genome-wide Investigation of Regulatory Roles of lncRNAs in Response to Heat and Drought Stress in Brassica Juncea (Indian Mustard). *Environ. Exp. Bot.* 171, 103922. doi:10.1016/j.envexpbot.2019.103922

Bohra, A., Gandham, P., Rathore, A., Thakur, V., Saxena, R. K., Naik, S. J. S., et al. (2021). Identification of microRNAs and Their Gene Targets in Cytoplasmic Male Sterile and fertile Maintainer Lines of Pigeonpea. *Planta* 253, 1–15. doi:10.1007/s00425-021-03568-6

Bohra, A., Jha, U. C., Godwin, I., and Varshney, R. K. (2020). Genomic Interventions for Sustainable Agriculture. *Plant Biotechnol. J.* 18, 2388–2405.

Bossdorf, O., Arcuri, D., Richards, C. L., and Pigliucci, M. (2010). Experimental Alteration of DNA Methylation Affects the Phenotypic Plasticity of Ecologically Relevant Traits in *Arabidopsis thaliana*. *Evol. Ecol.* 24, 541–553. doi:10.1007/s10682-010-9372-7

Boulanger, M., Chakraborty, M., Tempé, D., Piechaczyk, M., and Bossis, G. (2021). SUMO and Transcriptional Regulation: The Lessons of Large-Scale Proteomic, Modifomic and Genomic Studies. *Moleculesmdpi.com* 26, 828. doi:10.3390/molecules26040828

Bvindi, C., Lee, S., Tang, L., Mickelbart, M. V., Li, Y., and Mengiste, T. (2022). Main Manuscript for Improved Pathogen and Stress Tolerance in Tomato Mutants of SET Domain Histone 3 Lysine Methyltransferases. *bioRxiv* 2022. 02.19.481141. doi:10.1101/2022.02.19.481141

Callinan, P. A., and Feinberg, A. P. (2006). The Emerging Science of Epigenomics. *Hum. Mol. Genet.* 15, R95, R101 Spec No. doi:10.1093/hmg/ddl095

Casati, P., and Gomez, M. S. (2021). Chromatin Dynamics during DNA Damage and Repair in Plants: New Roles for Old Players. *J. Exp. Bot.* 72, 4119–4131. doi:10.1093/jxb/eraa551

Chandrika, N. N. P., Sundaravelpandian, K., Yu, S-M., and Schmidt, W. (2013). ALFIN-LIKE 6 Is Involved in Root Hair Elongation during Phosphate Deficiency in Arabidopsis, Wiley Online Libr. *New Phytol.* 198, 709–720. 2013, undefined. doi:10.1111/nph.12194

Chaudhry, U. K., Gökçe, Z. N. Ö., and Gökçe, A. F. (2021a). Drought and Salt Stress Effects on Biochemical Changes and Gene Expression of Photosystem II and Catalase Genes in Selected Onion Cultivars. *Biologia (Bratisl).* 76, 3107–3121. doi:10.1007/S11756-021-00827-5

Chaudhry, U. K., Junaid, M. D., and Gökçe, A. F. (2021b). "Influence of Environmental Adversities on Physiological Changes in Plants," in *Developing Climate-Resilient Crops* (Boca Raton, FL: CRC Press), 85–110. doi:10.1201/9781003109037-5-5

Chen, H., Feng, H., Zhang, X., Zhang, C., Wang, T., and Dong, J. (2019). An Arabidopsis E3 Ligase HUB2 Increases Histone H2B Monoubiquitination and Enhances Drought Tolerance in Transgenic Cotton. *Plant Biotechnol. J.* 17, 556–568. doi:10.1111/pbi.12998

Chin, C. S., Peluso, P., Sedlazeck, F. J., Nattestad, M., Concepcion, G. T., Clum, A., et al. (2016). Phased Diploid Genome Assembly with Single-Molecule Real-Time Sequencing. *Nat. Methods* 13, 1050–1054. doi:10.1038/nmeth.4035

Colomé-Tatché, M., Cortijo, S., Wardenaar, R., Morgado, L., Lahouz, B., Sarazin, A., et al. (2012). Features of the Arabidopsis Recombination Landscape Resulting from the Combined Loss of Sequence Variation and DNA Methylation. *Proc. Natl. Acad. Sci. U. S. A.* 109, 16240–16245. doi:10.1073/pnas.1212955109

Conrath, U., Beckers, G. J. M., Langenbach, C. J. G., and Jaskiewicz, M. R. (2015). Priming for Enhanced Defense. *Annu. Rev. Phytopathol.* 53, 97–119. doi:10.1146/annurev-phyto-080614-120132

Crespo-Salvador, Ó., Escamilla-Aguilar, M., López-Cruz, J., López-Rodas, G., and González-Bosch, C. (2018). Determination of Histone Epigenetic marks in Arabidopsis and Tomato Genes in the Early Response to Botrytis Cinerea. *Plant Cel Rep* 37, 153–166. doi:10.1007/S00299-017-2218-9

Crespo-Salvador, Ó., Sánchez-Giménez, L., López-Galiano, M. J., Fernández-Crespo, E., Schalschi, L., García-Robles, I., et al. (2020). The Histone marks Signature in Exonic and Intronic Regions Is Relevant in Early Response of Tomato Genes to Botrytis Cinerea and in miRNA Regulation. *Plants* 9. doi:10.3390/plants9030300

Crisp, P. A., Ganguly, D., Eichten, S. R., Borevitz, J. O., and Pogson, B. J. (2016). Reconsidering Plant Memory: Intersections between Stress Recovery, RNA Turnover, and Epigenetics. *Sci. Adv.* 2. doi:10.1126/sciadv.1501340

Cuerda-Gil, D., and Slotkin, R. (2016). Non-canonical RNA-Directed DNA Methylation. nature.com. Available at: https://idp.nature.com/authorize/casa?redirect_uri=https://www.nature.com/articles/nplants2016163&casa_token=Odyo1slr4hcAAAAA:Z0_pz-xDixXs-Ia1AlsPnEBxI-LNNuf5I9_DjNoj3WMfEJNXESQh4yrZGMPpTmSMpHJAw9JZ2TKVeIEtXg.doi:10.1038/nplants.2016.163

Dangol, S. Das., Hashmi, M. H., Saeed, F., Yel, I., Öztürk, A., and Bakhsh, A. (2021). Utilizing RNA-Based Approaches to Understand Plant-Insect Interactions. Available at: https://books.google.com/books?hl=en&lr=&id=V8IoEAAAQBAJ&oi=fnd&pg=PA392&dq=SD+dangol+et+al+2021&ots=wu5CIsvrvO&sig=g4vB7Sq3mGe4fg-Lfnu3Fgdj9ks (Accessed October 12, 2021).doi:10.1007/978-3-030-64994-4_17

Ding, Y., Avramova, Z., and Journal, M. F.-T. P. (2011). The Arabidopsis Trithorax-like Factor ATX1 Functions in Dehydration Stress Responses via ABA-dependent and ABA-independent Pathways, Wiley Online Libr. *Plant J.* 66, 735–744. 2011, undefined. doi:10.1111/j.1365-313X.2011.04534.x

Dowen, R. H., Pelizzola, M., Schmitz, R. J., Lister, R., Dowen, J. M., Nery, J. R., et al. (2012). Widespread Dynamic DNA Methylation in Response to Biotic Stress. *Proc. Natl. Acad. Sci. U. S. A.* 109. doi:10.1073/pnas.1209329109

Duan, Y., Li, J., Qin, R., Xu, R., Li, H., Yang, Y.-C., et al. (2016). Identification of a Regulatory Element Responsible for Salt Induction of rice OsRAV2 through *Ex Situ* and *In Situ* Promoter Analysis. *Plant J.* 90, 49–62. Springer. doi:10.1007/s11103-015-0393-z

El-Shami, M., Pontier, D., Lahmy, S., Braun, L., Picart, C., Vega, D., et al. (2007). Reiterated WG/GW Motifs Form Functionally and Evolutionarily Conserved ARGONAUTE-Binding Platforms in RNAi-Related Components. *Genes Dev.* 21, 2539–2544. doi:10.1101/gad.451207

Emran, A. M., Tabei, Y., Kobayashi, K., Yamaoka, N., and Nishiguchi, M. (2012). Molecular Analysis of Transgenic Melon Plants Showing Virus Resistance Conferred by Direct Repeat of Movement Gene of Cucumber green Mottle Mosaic Virus. *Plant Cel Rep* 31, 1371–1377. doi:10.1007/S00299-012-1237-9

Fan, H., Zhang, Z., Wang, N., Cui, Y., Sun, H., Liu, Y., et al. (2013). SKB1/PRMT5-mediated Histone H4R3 Dimethylation of Ib Subgroup bHLH Genes Negatively Regulates Iron Homeostasis in *Arabidopsis thaliana*, Wiley Online Libr. *Plant J.* 77, 209–221. doi:10.1111/tpj.12380

Feng, S., Jacobsen, S. E., and Reik, W. (2010). Epigenetic Reprogramming in Plant and Animal Development. *Science* 330 (80), 622–627. doi:10.1126/science.1190614

Ferreira, L. J., Donoghue, M. T. A., Barros, P., Saibo, N. J., Santos, A. P., and Oliveira, M. M. (2019). Uncovering Differentially Methylated Regions (DMRs) in a Salt-Tolerant Rice Variety under Stress: One Step towards New Regulatory Regions for Enhanced Salt Tolerance. *Epigenomes* 3, 4. doi:10.3390/epigenomes3010004

Fiaz, S., Ahmad, S., Noor, M., Wang, X., Younas, A., Riaz, A., et al. (2019). Applications of the CRISPR/Cas9 System for rice Grain Quality Improvement: Perspectives and Opportunities. *Int. J. Mol. Sci.* 20 (4), 888. 2019, undefined. doi:10.3390/ijms20040888

Finnegan, E. J., Genger, R. K., Kovac, K., Peacock, W. J., and Dennis, E. S. (1998). DNA Methylation and the Promotion of Flowering by Vernalization. *Natl. Acad. Sci.* 95, 5824–5829. Available at: https://www.pnas.org/content/95/10/5824.short (Accessed October 6, 2021). doi:10.1073/pnas.95.10.5824

Flusberg, B. A., Webster, D. R., Lee, J. H., Travers, K. J., Olivares, E. C., Clark, T. A., et al. (2010). Direct Detection of DNA Methylation during Single-Molecule, Real-Time Sequencing. *Nat. Methods* 7, 461–465. doi:10.1038/nmeth.1459

Forestan, C., Farinati, S., Rouster, J., and Genetics, H. L. (2018). Control of Maize Vegetative and Reproductive Development, Fertility, and rRNAs Silencing by HISTONE DEACETYLASE 108. *Genetics* 208 (4), 1443–1466. 2018, undefined. doi:10.1534/genetics.117.300625

Fraga, M. F., and Esteller, M. (2002). DNA Methylation: A Profile of Methods and Applications. *Biotechniques* 33, 632–649. doi:10.2144/02333rv01

Frommer, M., McDonald, L. E., Millar, D. S., Collis, C. M., Watt, F., Grigg, G. W., et al. (1992). A Genomic Sequencing Protocol that Yields a Positive Display of 5- Methylcytosine Residues in Individual DNA Strands. *Proc. Natl. Acad. Sci. U. S. A.* 89, 1827–1831. doi:10.1073/pnas.89.5.1827

Ganai, S. A. (2020). "Summa of Erasers of Histone Acetylation with Special Emphasis on Classical Histone Deacetylases (HDACs)," in *Histone Deacetylase Inhibitors in Combinatorial Anticancer Therapy* (Springer Singapore), 67–74. doi:10.1007/978-981-15-8179-3_3

Ghildiyal, M., and Zamore, P. D. (2009). Small Silencing RNAs: An Expanding Universe. *Nat. Rev. Genet.* 10, 94–108. doi:10.1038/nrg2504

Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., et al. (2013). XCRISPR-mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell* 154, 442. doi:10.1016/j.cell.2013.06.044

Gökçe, A. F., Chaudhry, U. K., and Junaid, M. D. (2021). "Mapping QTLs for Abiotic Stress," in *Developing Climate-Resilient Crops* (Boca Raton, FL: CRC Press), 175–201. doi:10.1201/9781003109037-9-9

Gökçe, A. F., and Chaudhry, U. K. (2020). "Use of QTL in Developing Stress Tolerance in Agronomic Crops," in *Agronomic Crops* (Springer), 527–556. doi:10.1007/978-981-15-0025-1_25

Guo, H., Wu, T., Li, S., He, Q., Yang, Z., Zhang, W., et al. (2019). The Methylation Patterns and Transcriptional Responses to Chilling Stress at the Seedling Stage in rice. *Int. J. Mol. Sci.* 20, 5089. doi:10.3390/ijms20205089

Hafeez, M. B., Raza, A., Zahra, N., Shaukat, K., Akram, M. Z., Iqbal, S., et al. (2021). Gene Regulation in Halophytes in Conferring Salt Tolerance. *Handbook of Bioremediation*, 341–370. doi:10.1016/b978-0-12-819382-2.00022-3

Hahn, M. A., Li, A. X., Wu, X., and Pfeifer, G. P. (2014). "Single Base Resolution Analysis of 5-methylcytosine and 5-hydroxymethylcytosine by RRBS and TAB-RRBS," in *Cancer Epigenetics: Risk Assessment, Diagnosis, Treatment, and Prognosis* (New York: Springer), 273–287. doi:10.1007/978-1-4939-1804-1_14

Haider, S., Iqbal, J., Naseer, S., Yaseen, T., Shaukat, M., Bibi, H., et al. (2021). Molecular Mechanisms of Plant Tolerance to Heat Stress: Current Landscape and Future Perspectives. *Plant Cel Rep* 40, 2247–2271. doi:10.1007/s00299-021-02696-3

Haider, S., Raza, A., Iqbal, J., Shaukat, M., and Mahmood, T. (2022). Analyzing the Regulatory Role of Heat Shock Transcription Factors in Plant Heat Stress

Tolerance: a Brief Appraisal. *Mol. Biol. Rep.* 1–15. doi:10.1007/s11033-022-07190-x

Han, Z., Yu, H., Zhao, Z., Hunter, D., Luo, X., Duan, J., et al. (2016). AtHD2D Gene Plays a Role in Plant Growth, Development, and Response to Abiotic Stresses in Arabidopsis Thaliana. *Front. Plant Sci.* 7. doi:10.3389/fpls.2016.00310

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

He, Y., and Li, Z. (2018). Epigenetic Environmental Memories in Plants: Establishment, Maintenance, and Reprogramming. *Trends Genet.* 34, 856–866. doi:10.1016/j.tig.2018.07.006

Heard, E., and Martienssen, R. A. (2014). Transgenerational Epigenetic Inheritance: Myths and Mechanisms. *Cell* 157, 95–109. doi:10.1016/j.cell.2014.02.045

Henikoff, S., and Greally, J. M. (2016). Epigenetics, Cellular Memory and Gene Regulation. *Curr. Biol.* 26, R644–R648. doi:10.1016/j.cub.2016.06.011

Heo, J. B., Lee, Y. S., and Sung, S. (2013). Epigenetic Regulation by Long Noncoding RNAs in Plants. *Chromosom. Res.* 21, 685–693. doi:10.1007/s10577-013-9392-6

Herrera, C. M., Medrano, M., and Bazaga, P. (2016). Comparative Spatial Genetics and Epigenetics of Plant Populations: Heuristic Value and a Proof of Concept. *Mol. Ecol.* 25, 1653–1664. doi:10.1111/mec.13576

Herron, S. A., Rubin, M. J., Ciotir, C., Crews, T. E., Van Tassel, D. L., and Miller, A. J. (2020). Comparative Analysis of Early Life Stage Traits in Annual and Perennial Phaseolus Crops and Their Wild Relatives. *Front. Plant Sci.* 11. doi:10.3389/FPLS.2020.00034/FULL

Hewezi, T., Lane, T., Piya, S., Rambani, A., Rice, J. H., and Staton, M. (2017). Cyst Nematode Parasitism Induces Dynamic Changes in the Root Epigenome. *Plant Physiol.* 174, 405–420. doi:10.1104/pp.16.01948

Hickey, L. T., Hafeez, N. 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

Hilker, M., Schwachtje, J., Baier, M., Balazadeh, S., Bäurle, I., Geiselhardt, S., et al. (2016). Priming and Memory of Stress Responses in Organisms Lacking a Nervous System. *Biol. Rev.* 91, 1118–1133. doi:10.1111/brv.12215

Holliday, R., and Pugh, J. E. (1975). DNA Modification Mechanisms and Gene Activity during Development. *Sci.*187 (80), 226–232. doi:10.1126/science.111109810.1126/science.187.4173.226

Hossain, M., Kawakatsu, T., Kim, K., Zhang, N., Nguyen, C. T., Khan, S. M., et al. 2017, undefined (2017). Divergent Cytosine DNA Methylation Patterns in Single-cell, Soybean Root Hairs. Wiley Online Libr. 214, 808–819. doi:10.1111/nph.14421

Hou, H., Zhao, L., Zheng, X., Gautam, M., Yue, M., Hou, J., et al. (2019a). Dynamic Changes in Histone Modification Are Associated with Upregulation of Hsf and rRNA Genes during Heat Stress in maize Seedlings. *Protoplasma* 256, 1245–1256. doi:10.1007/S00709-019-01364-4

Hou, J., Lu, D., Mason, A. S., Li, B., Xiao, M., An, S., et al. (2019b). Non-coding RNAs and Transposable Elements in Plant Genomes: Emergence, Regulatory Mechanisms and Roles in Plant Development and Stress Responses. *Planta* 250, 23–40. doi:10.1007/s00425-019-03166-7

Hu, X., Lu, M., Li, C., Liu, T., Wang, W., Wu, J., et al. (2011). Differential Expression of Proteins in maize Roots in Response to Abscisic Acid and Drought. *Acta Physiol. Plant* 33, 2437–2446. doi:10.1007/S11738-011-0784-Y

Hu, X., Wu, X., Li, C., Lu, M., Liu, T., Wang, Y., et al. (2012). Abscisic Acid Refines the Synthesis of Chloroplast Proteins in Maize (Zea mays) in Response to Drought and Light. *PLoS One* 7. doi:10.1371/JOURNAL.PONE.0049500

Huan, Q., Mao, Z., Chong, K., and Phytologist, J. Z.-N. (2018). Global Analysis of H3K4me3/H3K27me3 in Brachypodium Distachyon Reveals VRN3 as Critical Epigenetic Regulation point in Vernalization and Provides Insights into Epigenetic Memory, Wiley Online Libr. *New Phytol.* 219, 1373–1387. 2018, undefined. doi:10.1111/nph.15288

Huang, W., Xiong, J., Yang, Y., Liu, S. M., Yuan, B. F., and Feng, Y. Q. (2015). Determination of DNA Adenine Methylation in Genomes of Mammals and Plants by Liquid Chromatography/mass Spectrometry. *RSC Adv.* 5, 64046–64054. doi:10.1039/c5ra05307b

Jackson, J. P., Lindroth, A. M., Cao, X., and Jacobsen, S. E. (2002). Control of CpNpG DNA Methylation by the KRYPTONITE Histone H3 Methyltransferase. *Nature* 416, 556–560. doi:10.1038/nature731

Jasencakova, Z., Soppe, W. J. J., Meister, A., Gernand, D., Turner, B. M., and Schubert, I. (2003). Histone Modifications in Arabidopsis - High Methylation of H3 Lysine 9 Is Dispensable for Constitutive Heterochromatin. *Plant J.* 33, 471–480. doi:10.1046/J.1365-313X.2003.01638.X

Jha, U. C., Bohra, A., and Singh, N. P. (2014). Heat Stress in Crop Plants: Its Nature, Impacts and Integrated Breeding Strategies to Improve Heat Tolerance. *Plant Breed* 133, 679–701. doi:10.1111/pbr.12217

Johannes, F., Porcher, E., Teixeira, F. K., Saliba-Colombani, V., Simon, M., Agier, N., et al. (2009). Assessing the Impact of Transgenerational Epigenetic Variation on Complex Traits. *Plos Genet.* 5. doi:10.1371/journal.pgen.1000530

Joseph, T. J., Chandhini, S., Das, S., Mysore, S. K., and Shah, M. J. (2021). Methylation Status of Arabidopsis DNA Repair Gene Promoters during Agrobacterium Infection Reveals Epigenetic Changes in Three Generations. *Plant Mol. Biol. Rep.* doi:10.1007/S11105-021-01287-6

Junaid, M. D., Chaudhry, U. K., and Gökçe, A. F. (2021). Climate Change and Plant Growth – South Asian Perspective. *Clim. Change Plants*, 37–53. doi:10.1201/9781003108931-3-3

Karlik, E., Ari, S., and Gozukirmizi, N. (2019). LncRNAs: Genetic and Epigenetic Effects in Plants. *Biotechnol. Biotechnol. Equip.* 33, 429–439. doi:10.1080/13102818.2019.1581085

Karlsson, M., Weber, W., and Fussenegger, M. (2011). De Novo design and Construction of an Inducible Gene Expression System in Mammalian Cells. *Methods Enzymol.* 497, 239–253. doi:10.1016/B978-0-12-385075-100011-110.1016/b978-0-12-385075-1.00011-1

Kawakatsu, T., Huang, S., Jupe, F., Sasaki, E., and Cell, R. S. (2016). Epigenomic Diversity in a Global Collection of *Arabidopsis thaliana* Accessions Elsevier Enhanced Reader, 2016, U. *Elsevier*, 492–505. Available at: https://www.sciencedirect.com/science/article/pii/S0092867416308522 (Accessed October 6, 2021).

Kchouk, M., Gibrat, J. F., and Elloumi, M. (2017). Generations of Sequencing Technologies: From First to Next Generation. *Biol. Med.* 09. doi:10.4172/0974-8369.1000395

Ke, R., Mignardi, M., Hauling, T., and Nilsson, M. (2016). Fourth Generation of Next-Generation Sequencing Technologies: Promise and Consequences. *Hum. Mutat.* 37, 1363–1367. doi:10.1002/humu.23051

Kellenberger, R. T., Schlüter, P. M., and Schiestl, F. P. (2016). Herbivore-Induced DNA Demethylation Changes floral Signalling and Attractiveness to Pollinators in brassica Rapa. *PLoS One* 11. doi:10.1371/JOURNAL.PONE.0166646

Kim, J. M., To, T. K., and Seki, M. (2012). An Epigenetic Integrator: New Insights into Genome Regulation, Environmental Stress Responses and Developmental Controls by HISTONE DEACETYLASE 6. *m* 53, 794–800. doi:10.1093/pcp/pcs004

Kinoshita, T., and Seki, M. (2014). Epigenetic Memory for Stress Response and Adaptation in Plants. *Plant Cel Physiol* 55, 1859–1863. doi:10.1093/pcp/pcu125

Kumar, S., Beena, A. S., Awana, M., and Singh, A. (2017). Physiological, Biochemical, Epigenetic and Molecular Analyses of Wheat (*Triticum aestivum*) Genotypes with Contrasting Salt Tolerance. *Front. Plant Sci.* 8. doi:10.3389/FPLS.2017.01151/FULL

Kushwaha, M., Rostain, W., Prakash, S., Duncan, J. N., and Jaramillo, A. (2016). Using RNA as Molecular Code for Programming Cellular Function. *ACS Synth. Biol.* 5, 795–809. doi:10.1021/acssynbio.5b00297

Labra, M., Ghiani, A., Citterio, S., Sgorbati, S., Sala, F., Vannini, C., et al. (2002). Analysis of Cytosine Methylation Pattern in Response to Water Deficit in Pea Root Tips. *Plant Biol.* 4, 694–699. doi:10.1055/S-2002-37398

Lämke, J., Brzezinka, K., Altmann, S., and Bäurle, I. (2016). A Hit-and-run Heat Shock Factor Governs Sustained Histone Methylation and Transcriptional Stress Memory. *EMBO J.* 35, 162–175. doi:10.15252/EMBJ.201592593

Latzel, V., Zhang, Y., Karlsson Moritz, K., Fischer, M., and Bossdorf, O. (2012). Epigenetic Variation in Plant Responses to Defence Hormones. *Ann. Bot.* 110, 1423–1428. doi:10.1093/aob/mcs088

Le, T. N., Schumann, U., Smith, N. A., Tiwari, S., Khang Au, P. C., Zhu, Q. H., et al. (2014). DNA Demethylases Target Promoter Transposable Elements to Positively Regulate Stress Responsive Genes in Arabidopsis. *Genome Biol.* 15. doi:10.1186/s13059-014-0458-3

Li, S., He, X., Gao, Y., Zhou, C., Chiang, V. L., and Li, W. (2021). Histone Acetylation Changes in Plant Response to Drought Stress. *Genes (Basel).* 12, 1409. doi:10.3390/genes12091409

Li, T., Chen, X., Zhong, X., Zhao, Y., Liu, X., Zhou, S., et al. (2013). Jumonji C Domain Protein JMJ705-Mediated Removal of Histone H3 Lysine 27 Trimethylation Is Involved in Defense-Related Gene Activation in rice. *Plant Cell* 25, 4725–4736. doi:10.1105/tpc.113.118802

Liang, D., Zhang, Z., Wu, H., Huang, C., Shuai, P., Ye, C. Y., et al. (2014). Single-base-resolution Methylomes of Populus Trichocarpa Reveal the Association between DNA Methylation and Drought Stress. *BMC Genet.* 15. doi:10.1186/1471-2156-15-S1-S9

Liang, Z., Riaz, A., Chachar, S., Ding, Y., Du, H., and Gu, X. (2019). Epigenetic Modifications of mRNA and DNA in Plants. *Mol. Plant*. doi:10.1016/j.molp.2019.12.007

Liang, Z., Shen, L., Cui, X., Bao, S., Geng, Y., Yu, G., et al. (2018). DNA N6-Adenine Methylation in *Arabidopsis thaliana*. *Dev. Cel* 45, 406–416. e3. doi:10.1016/j.devcel.2018.03.012

Liu, J., Feng, L., Gu, X., Deng, X., Qiu, Q., Li, Q., et al. (2019a). An H3K27me3 Demethylase-HSFA2 Regulatory Loop Orchestrates Transgenerational Thermomemory in Arabidopsis. *Cell Res* 29, 379–390. doi:10.1038/s41422-019-0145-8

Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., et al. (2014). Comparison of Next-Generation Sequencing Systems. *J. Biomed. Biotechnol.* 2012, 251364. doi:10.1155/2012/251364

Liu, W., Duttke, S. H., Hetzel, J., Groth, M., Feng, S., Gallego-Bartolome, J., et al. (2018). RNA-directed DNA Methylation Involves Co-transcriptional Small-RNA-Guided Slicing of Polymerase V Transcripts in Arabidopsis. *Nat. Plants* 4, 181–188. doi:10.1038/s41477-017-0100-y

Liu, Y., Cheng, J., Siejka-Zielińska, P., Weldon, C., Roberts, H., Lopopolo, M., et al. (2020). Accurate Targeted Long-Read DNA Methylation and Hydroxymethylation Sequencing with TAPS. *Genome Biol.* 21. doi:10.1186/s13059-020-01969-6

Liu, Y., Siejka-Zielińska, P., Velikova, G., Bi, Y., Yuan, F., Tomkova, M., et al. (2019b). Bisulfite-free Direct Detection of 5-methylcytosine and 5-hydroxymethylcytosine at Base Resolution. *Nat. Biotechnol.* 37, 424–429. doi:10.1038/s41587-019-0041-2

Lloyd, J. P. B., and Lister, R. (2022). Epigenome Plasticity in Plants. *Nat. Rev. Genet.* 23, 55–68. doi:10.1038/s41576-021-00407-y

Ma, X., Zhu, Q., Chen, Y., and Liu, Y. G. (2016). CRISPR/Cas9 Platforms for Genome Editing in Plants: Developments and Applications. *Mol. Plant* 9, 961–974. doi:10.1016/j.molp.2016.04.009

Mao, H., Wang, H., Liu, S., Li, Z., Yang, X., Yan, J., et al. (2015). A Transposable Element in a NAC Gene Is Associated with Drought Tolerance in maize Seedlings. *Nat. Commun.* 6. doi:10.1038/ncomms9326

Marfil, C. F., Camadro, E. L., and Masuelli, R. W. (2009). Phenotypic Instability and Epigenetic Variability in a Diploid Potato of Hybrid Origin, Solanum Ruiz-Lealii. *BMC Plant Biol.* 9. doi:10.1186/1471-2229-9-21

Martínez de Alba, A. E., Elvira-Matelot, E., and Vaucheret, H. (2013). Gene Silencing in Plants: A Diversity of Pathways. *Biochim. Biophys. Acta - Gene Regul. Mech.* 1829, 1300–1308. doi:10.1016/j.bbagrm.2013.10.005

Matzke, M. A., and Mosher, R. A. (2014). RNA-directed DNA Methylation: An Epigenetic Pathway of Increasing Complexity. *Nat. Rev. Genet.* 15, 570. doi:10.1038/nrg3794

Meissner, A., Gnirke, A., Bell, G. W., Ramsahoye, B., Lander, E. S., and Jaenisch, R. (2005). Reduced Representation Bisulfite Sequencing for Comparative High-Resolution DNA Methylation Analysis. *Nucleic Acids Res.* 33, 5868–5877. doi:10.1093/nar/gki901

Meller, B., Kuźnicki, D., Arasimowicz-Jelonek, M., Deckert, J., and Floryszak-Wieczorek, J. (2018). Baba-primed Histone Modifications in Potato for Intergenerational Resistance to Phytophthora Infestans. *Front. Plant Sci.* 9. doi:10.3389/fpls.2018.01228

Miglani, G., Kaur, A., and Kaur, L. (2019). Plant Gene Expression Control Using Genome-And Epigenome-Editing Technologies. *Taylor Fr* 34, 1–63. 2020, undefined. doi:10.1080/15427528.2019.1678541

Mignardi, M., and Nilsson, M. (2014). Fourth-generation Sequencing in the Cell and the Clinic. *Genome Med.* 6. doi:10.1186/gm548

Miryeganeh, M. (2021). Plants' Epigenetic Mechanisms and Abiotic Stress. *Genes (Basel)* 12. doi:10.3390/genes12081106

Mishra, A., and Bohra, A. (2018). Non-coding RNAs and Plant Male Sterility: Current Knowledge and Future Prospects. *Plant Cel Rep* 37, 177–191. doi:10.1007/s00299-018-2248-y

Mozgova, I., and Hennig, L. (2015). The Polycomb Group Protein Regulatory Network. *Annu. Rev. Plant Biol.* 66, 269–296. doi:10.1146/ANNUREV-ARPLANT-043014-115627

Muthamilarasan, M., and Prasad, M. (2013). Plant Innate Immunity: An Updated Insight into Defense Mechanism. *J. Biosci.* 38, 433–449. doi:10.1007/s12038-013-9302-2

Nambara, E., and Marion-Poll, A. (2003). ABA Action and Interactions in Seeds. *Trends Plant Sci.* 8, 213–217. doi:10.1016/S1360-1385(03)00060-8

Nguyen, B. T., and Kang, M. J. (2019). Application of Capillary Electrophoresis with Laser-Induced Fluorescence to Immunoassays and Enzyme Assays. *Molecules* 24. doi:10.3390/molecules24101977

Nie, W.-F., and Wang, J. (20212021). Actin-Related Protein 4 Interacts with PIE1 and Regulates Gene Expression in Arabidopsis. *Genes* 12 (520 12), 520. doi:10.3390/GENES12040520

Pandey, G., Yadav, C. B., Sahu, P. P., Muthamilarasan, M., and Prasad, M. (2017). Salinity Induced Differential Methylation Patterns in Contrasting Cultivars of Foxtail Millet (Setaria Italica L.). *Plant Cel Rep* 36, 759–772. doi:10.1007/s00299-016-2093-9

Pang, M., Sun, X., Ausió, J., and Ausió, J. (2020). Histone H4 Variant, H4G, Drives Ribosomal RNA Transcription and Breast Cancer Cell Proliferation by Loosening Nucleolar Chromatin Structure, Wiley Online Libr. *J. Cel Physiol* 235, 235. 2020, undefined. doi:10.1002/jcp.29770

Papanicolau-Sengos, A., and Aldape, K. (2022). DNA Methylation Profiling: An Emerging Paradigm for Cancer Diagnosis. *Annu. Rev. Pathol. Mech. Dis.* 17. doi:10.1146/annurev-pathol-042220-022304

Parent, J., Cahn, J., Herridge, R. P., Grimanelli, D., and Robert, A. M. (2021). Small RNAs Guide Histone Methylation in Arabidopsis Embryos. *genesdev.cshlp.org* 35 (11-12), 841–846. 2021, undefined. doi:10.1101/gad.343871.120

Park, J., Lim, C. J., Shen, M., Park, H. J., Cha, J. Y., Iniesto, E., et al. (2018). Epigenetic Switch from Repressive to Permissive Chromatin in Response to Cold Stress. *Proc. Natl. Acad. Sci. U. S. A.* 115. doi:10.1073/pnas.1721241115

Pavangadkar, K., Thomashow, M. F., and Triezenberg, S. J. (2010). Histone Dynamics and Roles of Histone Acetyltransferases during Cold-Induced Gene Regulation in Arabidopsis. *Plant Mol. Biol.* 74, 183–200. doi:10.1007/S11103-010-9665-9

Piatek, A., Ali, Z., Baazim, H., Al-Shareef, S., and Aouida, M. (2015). RNA-guided Transcriptional Regulation in Planta via Synthetic dCas9-based Transcription Factors. *Wiley Online Libr.* 13, 578–589. biotechnology, and 2015, undefined. doi:10.1111/pbi.12284

Popova, O. V., Dinh, H. Q., Aufsatz, W., and Jonak, C. (2013). The RdDM Pathway Is Required for Basal Heat Tolerance in Arabidopsis. *Mol. Plant* 6, 396–410. doi:10.1093/mp/sst023

Quinn, J. J., and Chang, H. Y. (2016). Unique Features of Long Non-coding RNA Biogenesis and Function. *Nat. Rev. Genet.* 17, 47–62. doi:10.1038/nrg.2015.10

Raza, A., Ashraf, F., Zou, X., Zhang, X., and Tosif, H. (2020). "Plant Adaptation and Tolerance to Environmental Stresses: Mechanisms and Perspectives," in *Plant Ecophysiology and Adaptation under Climate Change: Mechanisms and Perspectives I: General Consequences and Plant Responses* (Springer Singapore), 117–145. doi:10.1007/978-981-15-2156-0_5

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. (2020). Metabolomics: a Systems Biology Approach for Enhancing Heat Stress Tolerance in Plants. *Plant Cel Rep*. doi:10.1007/s00299-020-02635-8

Raza, A., Su, W., Jia, Z., Luo, D., Zhang, Y., Gao, A., et al. (2022). Mechanistic Insights into Trehalose-Mediated Cold Stress Tolerance in Rapeseed (Brassica Napus L.) Seedlings. *Front. Plant Sci.* 13, 857980. doi:10.3389/fpls.2022.857980

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

Rehman, M., and Tanti, B. (2020). Understanding Epigenetic Modifications in Response to Abiotic Stresses in Plants. *Biocatal. Agric. Biotechnol.* 27. doi:10.1016/j.bcab.2020.101673

Reinders, J., Wulff, B. B. H., Mirouze, M., Marí-Ordóñez, A., Dapp, M., Rozhon, W., et al. (2009). Compromised Stability of DNA Methylation and Transposon Immobilization in Mosaic Arabidopsis Epigenomes. *Genes Dev.* 23, 939–950. doi:10.1101/gad.524609

Richards, C., Alonso, C., Becker, C., Bossdorf, O., Bucher, E., Colomé-Tatché, M., et al. (2017). Ecological Plant Epigenetics: Evidence from Model and Non-model Species, and the Way Forward, Wiley Online Libr. *Ecol. Lett.* 20, 1576–1590. 2017, undefined. doi:10.1111/ele.12858

Rus, A., Yokoi, S., Sharkhuu, A., Reddy, M., Lee, B. H., Matsumoto, T. K., et al. (2001). AtHKT1 Is a Salt Tolerance Determinant that Controls Na+ Entry into Plant Roots. *Proc. Natl. Acad. Sci. U. S. A.* 98, 14150–14155. doi:10.1073/pnas.241501798

Saeed, F., Hashmi, M. H., Hossain, M. J., Ali, M. A., and Bakhsh, A. (2020). *Transgenic Technology Based Value Addition in Plant Biotechnology*, 123–156. doi:10.1016/b978-0-12-818632-9.00006-xTransgenic Technologies for Efficient Insect Pest Management in Crop Plants

Salim, N., and Raza, A. (2020). Nutrient Use Efficiency (NUE) for Sustainable Wheat Production: a Review. *Taylor Fr* 43, 297–315. doi:10.1080/01904167.2019.1676907

Sallam, N., and Moussa, M. (2021). DNA Methylation Changes Stimulated by Drought Stress in ABA-Deficient maize Mutant Vp10. *Plant Physiol. Biochem.* 160, 218–224. doi:10.1016/j.plaphy.2021.01.024

Sanchez, D. H., and Paszkowski, J. (2014). Heat-Induced Release of Epigenetic Silencing Reveals the Concealed Role of an Imprinted Plant Gene. *Plos Genet.* 10. doi:10.1371/JOURNAL.PGEN.1004806

Sanger, F., Nicklen, S., and Coulson, A. (1977). DNA Sequencing with Chain-Terminating. *Proc. Natl. Acad. Sci. USA* 74, 5463–5467. Available at: https://www.pnas.org/content/74/12/5463.short.doi:10.1073/pnas.74.12.5463

Schrey, A. W., Alvarez, M., Foust, C. M., Kilvitis, H. J., Lee, J. D., Liebl, A. L., et al. (2013). Ecological Epigenetics: Beyond MS-AFLP. *Integr. Comp. Biol.*, 340–350. doi:10.1093/icb/ict012

Sebastian Ramirez-Prado, J., Latrasse, D., Yaneth Rodriguez-Granados, N., Huang, Y., Manza-Mianza, D., Brik-Chaouche, R., et al. (2019). The Polycomb Protein LHP1 Regulates *Arabidopsis thaliana* Stress Responses through the Repression of the MYC2-dependent branch of Immunity. *Wiley Online Libr.* 100, 1118–1131. doi:10.1111/tpj.14502

Secco, D., Wang, C., Shou, H., Schultz, M. D., Chiarenza, S., Nussaume, L., et al. (2015). Stress Induced Gene Expression Drives Transient DNA Methylation Changes at Adjacent Repetitive Elements. *Elife* 4. doi:10.7554/eLife.09343.001

Singh, R. K., Jaishankar, J., Muthamilarasan, M., Shweta, S., Dangi, A., and Prasad, M. (20162016). Genome-wide Analysis of Heat Shock Proteins in C4 Model, Foxtail Millet Identifies Potential Candidates for Crop Improvement under Abiotic Stress. *Sci. Rep.* 61 (6), 1–14. doi:10.1038/srep32641

Singh, R. K., and Prasad, M. (2021). Delineating the Epigenetic Regulation of Heat and Drought Response in Plants. *Crit. Rev. Biotechnol.* doi:10.1080/07388551.2021.1946004

Singroha, G., and Sharma, P. (2019). Epigenetic Modifications in Plants under Abiotic Stress. *Epigenetics*. doi:10.5772/intechopen.84455

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

Soppe, W. J. J., Jasencakova, Z., Houben, A., Kakutani, T., Meister, A., Huang, M. S., et al. (2002). DNA Methylation Controls Histone H3 Lysine 9 Methylation and Heterochromatin Assembly in Arabidopsis. *EMBO J.* 21, 6549–6559. doi:10.1093/emboj/cdf657

Sosa-Valencia, G., Romero-Pérez, P. S., Miguel Palomar, V., Covarrubias, A. A., and Reyes, J. L. (2017). Insights into the Function of the phasiRNA-Triggering miR1514 in Response to Stress in Legumes. *Plant Signal. Behav.* 12. doi:10.1080/15592324.2017.1284724

Steward, N., Kusano, T., and Sano, H. (2000). Expression of ZmMET1, a Gene Encoding a DNA Methyltransferase from maize, Is Associated Not Only with DNA Replication in Actively Proliferating Cells, but Also with Altered DNA Methylation Status in Cold-Stressed Quiescent Cells. *Nucleic Acids Res.* 28, 3250–3259. doi:10.1093/nar/28.17.3250

Suksamran, R., Saithong, T., Thammarongtham, C., and Kalapanulak, S. (2020). Genomic and Transcriptomic Analysis Identified Novel Putative Cassava lncRNAs Involved in Cold and Drought Stress. *Genes (Basel)* 11. doi:10.3390/genes11040366

Suter, L., and Widmer, A. (2013). Phenotypic Effects of Salt and Heat Stress over Three Generations in *Arabidopsis thaliana*. *PLoS One* 8. doi:10.1371/JOURNAL.PONE.0080819

Takahashi, S., Katagiri, T., Yamaguchi-Shinozaki, K., and Shinozaki, K. (2000). An Arabidopsis Gene Encoding a Ca2+ -binding Protein Is Induced by Abscisic Acid during Dehydration. *m* 41, 898–903. doi:10.1093/pcp/pcd010

Tariq, M., Saze, H., Probst, A. V., Lichota, J., Habu, Y., and Paszkowski, J. (2003). Erasure of CpG Methylation in Arabidopsis Alters Patterns of Histone H3 Methylation in Heterochromatin. *Proc. Natl. Acad. Sci.* 100, 8823–8827. doi:10.1073/PNAS.1432939100

Temel, A., Janack, B., and Agronomy, K. H. (2017). Drought Stress-Related Physiological Changes and Histone Modifications in Barley Primary Leaves at HSP17 Gene, mdpi.com. *Agronomy* 7 (2), 43. 2017, undefined. doi:10.3390/agronomy7020043

To, T. K., and Kim, J. M. (2014). Epigenetic Regulation of Gene Responsiveness in Arabidopsis. *Front. Plant Sci.* 4. doi:10.3389/FPLS.2013.00548/FULL

Urquiaga, M. C. de. O., Thiebaut, F., Hemerly, A. S., and Ferreira, P. C. G. (2021). From Trash to Luxury: The Potential Role of Plant LncRNA in DNA Methylation during Abiotic Stress. *Front. Plant Sci.* 11. doi:10.3389/fpls.2020.603246

Usai, G., Vangelisti, A., Simoni, S., Giordani, T., Natali, L., Cavallini, A., et al. (2021). DNA Modification Patterns within the Transposable Elements of the Fig (Ficus Carica l.) Genome. *Plants* 10, 1–13. doi:10.3390/plants10030451

Varshney, R. K., Bohra, A., Yu, J., Graner, A., Zhang, Q., and Sorrells, M. E. (2021). Designing Future Crops: Genomics-Assisted Breeding Comes of Age. *Trends Plant Sci.* 26, 631–649. doi:10.1016/j.tplants.2021.03.010

Varshney, R. K., Marcel, T. C., Ramsay, L., Russell, J., Röder, M. S., Stein, N., et al. (2007). A High Density Barley Microsatellite Consensus Map with 775 SSR Loci. *Theor. Appl. Genet.* 114, 1091–1103. doi:10.1007/s00122-007-0503-7

Venturelli, S., Belz, R. G., Kämper, A., Berger, A., von Horn, K., Wegner, A., et al. (2015). Plants Release Precursors of Histone Deacetylase Inhibitors to Suppress Growth of Competitors. *Plant Cell* 27, 3175–3189. doi:10.1105/tpc.15.00585

Wang, P., Zhao, L., Hou, H., Zhang, H., Huang, Y., Wang, Y., et al. (2014). Epigenetic Changes Are Associated with Programmed Cell Death Induced by Heat Stress in Seedling Leaves of Zea Mays. *m* 56, 965–976. doi:10.1093/pcp/pcv023

Wang, W. S., Pan, Y. J., Zhao, X. Q., Dwivedi, D., Zhu, L. H., Ali, J., et al. (2011). Drought-induced Site-specific DNA Methylation and its Association with Drought Tolerance in rice (Oryza Sativa L.). *J. Exp. Bot.* 62, 1951–1960. doi:10.1093/jxb/erq391

Wang, Y., Li, H., Sun, Q., and Yao, Y. (2016). Characterization of Small RNAs Derived from tRNAs, rRNAs and snoRNAs and Their Response to Heat Stress in Wheat Seedlings. *PLoS One* 11. doi:10.1371/JOURNAL.PONE.0150933

Watson, A., Ghosh, S., Williams, M., and plants, W. C.-N. (20182017). undefinedSpeed Breeding Is a Powerful Tool to Accelerate Crop Research and Breeding. *nature.com*. doi:10.1101/161182

Watts, D., and MacBeath, J. R. E. (2003). "Automated Fluorescent DNA Sequencing," in *DNA Sequencing Protocols* (Totowa, NJ: Humana Press), 153–170. doi:10.1385/1-59259-113-2:153

Werghi, S., Herrero, F. A., Fakhfakh, H., and Gorsane, F. (2021). Auxin Drives Tomato Spotted Wilt Virus (TSWV) Resistance through Epigenetic Regulation of Auxin Response Factor ARF8 Expression in Tomato. *Gene*. doi:10.1016/j.gene.2021.145905

Widiez, T., El Kafafi, E. S., Girin, T., Berr, A., Ruffel, S., Krouk, G., et al. (2011). High Nitrogen Insensitive 9 (HNI9)-Mediated Systemic Repression of Root NO 3- Uptake Is Associated with Changes in Histone Methylation. *Proc. Natl. Acad. Sci. U. S. A.* 108, 13329–13334. doi:10.1073/pnas.1017863108

Xia, S., Cheng, Y., Huang, S., Win, J., Soards, A., and Jinn, T. (2013). Regulation of Transcription of Nucleotide-Binding Leucine-Rich Repeat-Encoding Genes SNC1 and RPP4 via. *Artic. Plant Physiol.* 162 (3), 1694–1705. doi:10.1104/pp.113.214551

Xing, J., Wang, T., Liu, Z., Xu, J., Yao, Y., Hu, Z., et al. (2015). General Control Nonrepressed Protein5-Mediated Histone Acetylation of Ferric Reductase Defective3 Contributes to Iron Homeostasis in Arabidopsis. *Plant Physiol.* 168, 1309–1320. doi:10.1104/pp.15.00397

Xu, J., Qi, W., Freeling, M., Xuecai, Z., Xu, Y., Mao, Y., et al. (2017). Natural Antisense Transcripts Are Significantly Involved in Regulation of Drought Stress in maize. *Nucleic Acids Res.* 45, 5126–5141. doi:10.1093/nar/gkx085

Xu, W., Dubos, C., and Lepiniec, L. (2015). Transcriptional Control of Flavonoid Biosynthesis by MYB-bHLH-WDR Complexes. *Trends Plant Sci.* 20, 176–185. doi:10.1016/j.tplants.2014.12.001

Yang, C., Shen, W., Chen, H., Chu, L., Xu, Y., Zhou, X., et al. (2018). Characterization and Subcellular Localization of Histone Deacetylases and Their Roles in Response to Abiotic Stresses in Soybean. *BMC Plant Biol.* 18. doi:10.1186/S12870-018-1454-7

Yang, R., Hong, Y., Ren, Z., Tang, K., Zhang, H., Zhu, J. K., et al. (2019). A Role for PICKLE in the Regulation of Cold and Salt Stress Tolerance in Arabidopsis. *Front. Plant Sci.* 10. doi:10.3389/fpls.2019.00900

Yong-Villalobos, L., González-Morales, S. I., Wrobel, K., Gutiérrez-Alanis, D., Cervantes-Peréz, S. A., Hayano-Kanashiro, C., et al. (2015). Methylome Analysis Reveals an Important Role for Epigenetic Changes in the Regulation of the Arabidopsis Response to Phosphate Starvation. *Proc. Natl. Acad. Sci. U. S. A.* 112, E7293–E7302. doi:10.1073/pnas.1522301112

Yu, A., Lepère, G., Jay, F., Wang, J., Bapaume, L., Wang, Y., et al. (2013). Dynamics and Biological Relevance of DNA Demethylation in Arabidopsis Antibacterial Defense. *Proc. Natl. Acad. Sci. U. S. A.* 110, 2389–2394. doi:10.1073/pnas.1211757110

Zezulin, A., and Musunuru, K. (2018). Turning up the Heat with Therapeutic Epigenome Editing. *Cell Stem Cell* 22, 10–11. doi:10.1016/j.stem.2017.12.013

Zhang, H., Lang, Z., and Zhu, J. K. (2018a). Dynamics and Function of DNA Methylation in Plants. *Nat. Rev. Mol. Cel Biol.* 19, 489–506. doi:10.1038/s41580-018-0016-z

Zhang, Q., Liang, Z., Cui, X., Ji, C., Li, Y., Zhang, P., et al. (2018b). N 6 -Methyladenine DNA Methylation in Japonica and Indica Rice Genomes and its Association with Gene Expression, Plant Development, and Stress Responses. *Mol. Plant* 11, 1492–1508. doi:10.1016/j.molp.2018.11.005

Zhang, X., Yazaki, J., Sundaresan, A., Cokus, S., Chan, S. W. L., Chen, H., et al. (2006). Genome-wide High-Resolution Mapping and Functional Analysis of DNA Methylation in Arabidopsis. *Cell* 126, 1189–1201. doi:10.1016/j.cell.2006.08.003

Zhang, Y. Y., Fischer, M., Colot, V., and Bossdorf, O. (2013). Epigenetic Variation Creates Potential for Evolution of Plant Phenotypic Plasticity. *New Phytol.* 197, 314–322. doi:10.1111/nph.12010

Zhao, B., Xi, Y., Kim, J., and Advances, S. S.-S. (20212021). undefinedChromatin Architectural Proteins Regulate Flowering Time by Precluding Gene Looping. advances.sciencemag.org. Available at: https://advances.sciencemag.org/content/7/24/eabg3097.abstract (Accessed October 6, 2021).730973108doi:10.1126/sciadv.abg3097

Zhao, L. Y., Song, J., Liu, Y., Song, C. X., and Yi, C. (2020a). Mapping the Epigenetic Modifications of DNA and RNA. *Protein Cell* 11, 792–808. doi:10.1007/s13238-020-00733-7

Zhao, M., Wang, T., Sun, T., Yu, X., Tian, R., and Zhang, W. H. (2020b). Identification of Tissue-specific and Cold-Responsive lncRNAs in Medicago Truncatula by High-Throughput RNA Sequencing. *BMC Plant Biol.* 20. doi:10.1186/s12870-020-2301-1

Zheng, M., Lin, J., Liu, X., Chu, W., Li, J., Gao, Y., et al. (2021). Histone Acetyltransferase TaHAG1 Acts as a Crucial Regulator to Strengthen Salt Tolerance of Hexaploid Wheat. *Plant Physiol.* 186, 1951–1969. doi:10.1093/plphys/kiab187

Zheng, M., Liu, X., Lin, J., Liu, X., Wang, Z., Xin, M., et al. (2019). Histone Acetyltransferase GCN5 Contributes to Cell wall Integrity and Salt Stress Tolerance by Altering the Expression of Cellulose Synthesis Genes. *Plant J.* 97, 587–602. doi:10.1111/tpj.14144

Zhonga, S. H., Liu, J. Z., Jin, H., Lin, L., Li, Q., Chen, Y., et al. (2013). Warm Temperatures Induce Transgenerational Epigenetic Release of RNA Silencing by Inhibiting siRNA Biogenesis in Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.* 110, 9171–9176. doi:10.1073/pnas.1219655110

Zhu, B., Zheng, Y., Angliker, H., Schwarz, S., Thiry, S., Siegmann, M., et al. (2000). 5-Methylcytosine DNA Glycosylase Activity Is Also Present in the Human MBD4 (G/T Mismatch Glycosylase) and in a Related Avian Sequence. *Nucleic Acids Res.* 28, 4157–4165. doi:10.1093/nar/28.21.4157

Zhu, J. K. (2016). Abiotic Stress Signaling and Responses in Plants. *Cell* 167, 313–324. doi:10.1016/j.cell.2016.08.029

Zhu, N., Cheng, S., Liu, X., Du, H., Dai, M., Zhou, D. X., et al. (2015). The R2R3-type MYB Gene OsMYB91 Has a Function in Coordinating Plant Growth and Salt Stress Tolerance in rice. *Plant Sci.* 236, 146–156. doi:10.1016/j.plantsci.2015.03.023

Zogli, P., and Libault, M. (2017). Plant Response to Biotic Stress: Is There a Common Epigenetic Response during Plant-Pathogenic and Symbiotic Interactions? *Plant Sci.* 263, 89–93. doi:10.1016/j.plantsci.2017.07.008

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.

Copyright © 2022 Saeed, Chaudhry, Bakhsh, Raza, Saeed, Bohra and Varshney. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Genetics

REVIEW
published: 29 June 2022
doi: 10.3389/fgene.2022.834366

Indian Wheat Genomics Initiative for Harnessing the Potential of Wheat Germplasm Resources for Breeding Disease-Resistant, Nutrient-Dense, and Climate-Resilient Cultivars

Sundeep Kumar[1], Sherry R. Jacob[1], Reyazul Rouf Mir[2], V. K. Vikas[3], Pawan Kulwal[4], Tilak Chandra[1], Satinder Kaur[5], Uttam Kumar[6], Suneel Kumar[1], Shailendra Sharma[7], Ravinder Singh[8], Sai Prasad[9], Anju Mahendru Singh[10], Amit Kumar Singh[1], Jyoti Kumari[1], M. S. Saharan[11], Subhash Chander Bhardwaj[12], Manoj Prasad[13], Sanjay Kalia[14] and Kuldeep Singh[1†]*

[1]Indian Council of Agricultural Research-National Bureau of Plant Genetic Resources, New Delhi, India, [2]Division of Genetics and Plant Breeding, Sher-e-Kashmir University of Agricultural Sciences and Technology of Kashmir (SKUAST-Kashmir), Jammu and Kashmir, India, [3]ICAR-Indian Agricultural Research Institute, New Delhi, India, [4]State Level Biotechnology Centre, Mahatma Phule Krishi Vidyapeeth, Rahuri, India, [5]School of Agricultural Biotechnology, Punjab Agricultural University, Ludhiana, India, [6]Borlaug Institute for South Asia, Ludhiana, India, [7]Department of Genetics and Plant Breeding, Chaudhary Charan Singh University, Meerut, Uttar Pradesh, [8]School of Biotechnology, Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu (SKUAST-Jammu), Jammu and Kashmir, India, [9]Indian Agriculture Research Institute Regional Research Station, Indore, India, [10]Division of Genetics, Indian Agricultural Research Institute, New Delhi, India, [11]Division of Plant Pathology, Indian Agricultural Research Institute, New Delhi, India, [12]CAR-Indian Institute of Wheat and Barley Research, Shimla, India, [13]Laboratory of Plant Virology, National Institute of Plant Genome Research, New Delhi, India, [14]Department of Biotechnology, Ministry of Science and Technology, New Delhi, India

Wheat is one of the major staple cereal food crops in India. However, most of the wheat-growing areas experience several biotic and abiotic stresses, resulting in poor quality grains and reduced yield. To ensure food security for the growing population in India, there is a compelling need to explore the untapped genetic diversity available in gene banks for the development of stress-resistant/tolerant cultivars. The improvement of any crop lies in exploring and harnessing the genetic diversity available in its genetic resources in the form of cultivated varieties, landraces, wild relatives, and related genera. A huge collection of wheat genetic resources is conserved in various gene banks across the globe. Molecular and phenotypic characterization followed by documentation of conserved genetic resources is a prerequisite for germplasm utilization in crop improvement. The National Genebank of India has an extensive and diverse collection of wheat germplasm, comprising Indian wheat landraces, primitive cultivars, breeding lines, and collection from other countries. The conserved germplasm can contribute immensely to the development of wheat cultivars with high levels of biotic and abiotic stress tolerance. Breeding wheat varieties that can give high yields under different stress environments has not made much headway due to high genotypes and environmental interaction, non-availability of truly resistant/tolerant germplasm, and non-availability of reliable markers linked with the QTL having a significant impact on resistance/tolerance. The development of new breeding technologies like genomic selection (GS), which takes into account the

OPEN ACCESS

Edited by:
Dragan Perovic,
Julius Kühn-Institut, Germany

Reviewed by:
Babar Hussain,
University of Central Punjab, Pakistan
Hakan Ozkan,
Çukurova University, Turkey
Prashant Vikram,
International Center for Biosaline Agriculture (ICBA), United Arab Emirates
Yuanfeng Hao,
Institute of Crop Sciences (CAAS), China

***Correspondence:**
Sundeep Kumar
Sundeep.Kumar@icar.gov.in

†Present address:
Kuldeep Singh,
Genetic Resource Division, ICRISAT, Patancheru, Hyderabad, India

Specialty section:
This article was submitted to Plant Genomics, a section of the journal Frontiers in Genetics

Received: 13 December 2021
Accepted: 09 May 2022
Published: 29 June 2022

Citation:
Kumar S, Jacob SR, Mir RR, Vikas VK, Kulwal P, Chandra T, Kaur S, Kumar U, Kumar S, Sharma S, Singh R, Prasad S, Singh AM, Singh AK, Kumari J, Saharan MS, Bhardwaj SC, Prasad M, Kalia S and Singh K (2022) Indian Wheat Genomics Initiative for Harnessing the Potential of Wheat Germplasm Resources for Breeding Disease-Resistant, Nutrient-Dense, and Climate-Resilient Cultivars. Front. Genet. 13:834366. doi: 10.3389/fgene.2022.834366

G × E interaction, will facilitate crop improvement through enhanced climate resilience, by combining biotic and abiotic stress resistance/tolerance and maximizing yield potential. In this review article, we have summarized different constraints being faced by Indian wheat-breeding programs, challenges in addressing biotic and abiotic stresses, and improving quality and nutrition. Efforts have been made to highlight the wealth of Indian wheat genetic resources available in our National Genebank and their evaluation for the identification of trait-specific germplasm. Promising genotypes to develop varieties of important targeted traits and the development of different genomics resources have also been highlighted.

Keywords: wheat, Indian wheat genomics initiative, genetic resources, genomics selection, gene bank, abiotic stress, biotic stress

1 INTRODUCTION

Wheat, a climate-sensitive crop, is grown on 31.76 million ha in India (ICAR-IIWBR, 2021). Most of the wheat-growing area faces several biotic and abiotic stresses and soil nutrient scarcity, resulting in poor quality grains and, finally, reduced yield (Pillay and Kumar, 2018; Grote et al., 2021). In the coming years, we will face a slew of challenges in ensuring food security for India's millions of people (Dev and Sharma, 2010; Pandey et al., 2021). It is the need of the hour to explore the idle genetic diversity leading to the development of resilient and better-performing cultivars under challenging situations (Phogat et al., 2021). Breeding wheat varieties that can give high yields under different stress environments has not made much headway due to large genotype × environment interaction, non-availability of truly resistant germplasm, and non-availability of reliable markers linked with the QTL, having a significant impact on resistance to biotic and abiotic stresses and quality traits (Vishwakarma et al., 2015; Beres et al., 2020; Khakda et al., 2020; Xiong et al., 2021). Despite consistent efforts by breeders to utilize genetic resources in breeding programs which is reflected in the form of discussions, review articles, meetings, and presentations made on harnessing the advantage of genetic diversity present in conserved germplasm, there are only a few success stories till date where specific traits have been introgressed from traditional landraces or germplasm to elite breeding lines (Mascher et al., 2019; Sharma et al., 2021; Zeibig et al., 2021). One of several other reasons is the lack of concerted efforts to create a single platform for key people working on the harvesting of diversity at the national and international levels. Markedly, the dwarfing genes, which led to the Green Revolution, were introduced into wheat and rice breeding lines from East Asian landraces (Hedden, 2003). In the case of barley, the deployment of the *mlo* alleles, commonly found in Ethiopian barley landraces (Jorgensen, 1992), led to broad-spectrum resistance to powdery mildew. In 1967, Krull and Borlaug stated, "the problem at present is less, a lack of genetic variation, but rather of efficiency in identifying and incorporating it" (Pistorius, 1997). Currently, genebank managers, plant breeders, and geneticists have reached a consensus that there is an urgent need for the systematic evaluation of the evolutionary potential of large seed collections stored in cold rooms (Mascher et al., 2019; Singh et al., 2020; Zeibig et al., 2021). Collection and conservation of these genetic resources are the primary prerequisites for their access and use in crop improvement programs. Over the past several decades, concerted efforts have been made toward extensive collection and *ex situ* conservation of wheat germplasm accessions belonging to each gene pool category. Reportedly, around 800,000 wheat accessions are conserved across 80 different germplasm collections (Singh et al., 2007). These collections are proportionately represented in major global gene banks, the data for which can be accessed on the Genesys PGR platform. The Genesys database has information on 464,784 wheat accessions held in 40 global collections. The largest collection is that of CIMMYT, Mexico (150,178), followed by USDA-ARS (66,246), ICARDA (48,149), the Australian Grains Genebank (42,626), and the NI Vavilov Institute (35,314) ("https://www.genesys-pgr.org" accessed on 25 Jan 2022). Out of the total entries of wheat in the Genesys system, 7,535 accessions are of Indian origin, which are held in the Australian Grains Genebank (1,533), USDA-ARS (1,320), CIMMYT (1,315), John Innes Institute, United Kingdom (1,174), NI Vavilov Institute (1,154), and ICARDA (407) (Jacob et al., 2015; Gauchan and Joshi, 2019). A major proportion of these (1,879) are traditional cultivars or landraces, which constitute the high-priority genetic wealth in any crop (Azeez et al., 2018; Marone et al., 2021). Of the 7,535 accessions, 2,197 are declared as part of the multilateral system (MLS) of ITPGRFA and, hence, freely available for distribution to all signatories of the treaty from their respective holding institutes (Jacob et al., 2015; Vernooy, 2019). Even before the treaty regime, these Indian resources have made a significant contribution to global wheat programs, and the most significant examples are those of NP4 and Hard Red Calcutta (Singh S. K. et al., 2015). The NP4 and Hard Red Calcutta are prevalent in the pedigrees of several modern wheat varieties grown across the world. The National Genebank in India is the second-largest genebank in the world and has the fifth largest collection of wheat genetic resources, including several unique landraces and exotic ones (Tyagi, 2016; Phogat et al., 2021). Based on the *ex situ* germplasm collection size conserved in long-term storage, the Indian National Genebank has the second-largest collection in the world (a total of 459,885 accessions conserved in NGB, India, as on 31 March 2022), next only to the USDA Genebank (Tyagi et al., 2015). In the case of wheat, its *ex situ* collection in the National Genebank (34,000 accessions as on 31 March 2022) is the fifth-largest collection in the world (CIMMYT Genebank ranks first)

(Jacob et al., 2015; Adhikari L. et al., 2022). These indigenous germplasm accessions are a valuable repository of economically important traits. The main goal of this review was to chart out a strategy for accelerating the use of this germplasm in the wheat-breeding program to address the various challenges in wheat production that in turn would minimize yield losses and maximize farmers' income.

Most of the wheat-breeding programs across the world have relied more on limited sets of diverse genotypes. This has resulted in the narrowing of the genetic base of cultivated wheat and became a dominant production constraint. Therefore, an expansion in the genetic base of genotypes should be considered in the wheat-breeding program. This can be carried out in various ways: 1) use of plant genetic resources, including wild relatives and landraces; 2) germplasm-assisted breeding using advanced genomic tools; and 3) development of transgenic and use of modern techniques like gene editing. Several gene banks across the globe house a large number of diverse germplasm accessions of wheat. These germplasm accessions harbor many important genes not only for various biotic and abiotic stresses but also for nutritional qualities and yield traits. To harness the true potential of the vast wealth of accessions stored in the Indian National Genebank, the Department of Biotechnology (DBT), Govt. of India, has supported a mega project to dissect the available genetic resources for new trait discovery using genomics and phenomics approaches and their integration for improving climate resilience, productivity, and nutritional quality. In addition, it will also help in the identification of the novel QTL and the markers linked with these QTL for these traits. The presented reviews emphasized the dominant stresses limiting wheat production in India and its impact on global food security. The possible path of a second green revolution using preserved genetic resources in the Indian Genebank with prospects to make a necessary plan for the exigencies that may be arising due to various biotic and abiotic stresses, nutrient utilization, and sustainability was also discussed.

2 THE CHALLENGES: IMPROVING WHEAT BIOTIC AND ABIOTIC STRESS RESILIENCE, PRODUCTIVITY, QUALITY, NUTRITION, AND SUSTAINABILITY

Wheat is one of the key cereal crops not only in India but also in the world and is the primary grain consumed by humans around the world. It is a food source for around 35% of the world population, a major cereal crop, and the main contributor to the agricultural economy of India (Nagarajan, 2005; Joshi et al., 2007a) and needs to be systematically worked upon for sustenance and improvement. With a world population that is estimated to increase to nearly 10 billion by 2050, the demand for wheat would also increase at an annual rate of about 1.7% (Alexandratos and Bruinsma, 2012). On the other hand, wheat yield is growing at about 1% annually, hence, is not keeping pace with the increasing demand (Hatfield and Beres, 2019). Wheat production is being hampered by newly evolved, more aggressive pests and diseases, limited water resources, limited arable land, and rapidly changing climatic conditions (Beres et al., 2020). Wheat plays a substantial role in global food security and provides nutrition to a major part of the population in developing countries. Although with the breeding efforts made over the decades, several countries, including India, have attained self-sufficiency in wheat production, it is high time to think and plan for the future. This is important because the population is expected to grow at a higher rate than the dwindling land area for cultivation every year, pathogens are ever-evolving, and abiotic environments are constantly changing, all mingled with sudden outbreaks (Hussain, 2015; Hasanuzzaman et al., 2019). The first challenge to wheat productivity has been biotic stress, which is caused by an infinite, ever-evolving pathogen, and mining appropriate Indian wheat germplasm against such incidents is required.

2.1 Biotic Stresses

The production and productivity of wheat crop is hampered by various diseases including, rusts (leaf rust, stem rust, and stripe rust), powdery mildew, spot blotch, Karnal bunt, and *Fusarium* head blight (Singh and Rajaram, 2002; Hussain 2015; Vikas et al., 2020; Roy et al., 2021). Although, continuous efforts have been made to develop disease-resistant wheat varieties for several devastating diseases, it is also true that knocking down of the resistance genes against these diseases happens simultaneously (Ahirwar et al., 2018; Kumar et al., 2018). This is where the emphasis is needed because resistance breeding programs have frequently relied on single major genes, and there is large-scale cultivation of genotypes with almost identical resistance (Vikas et al., 2020; Roy et al., 2021). Moreover, due to directional selection, the genetic base has become narrow, leading to a monoculture (Bourke et al., 2021). This is going to be a very serious threat to wheat production in the coming decades. There is a need for the identification and stacking of multiple resistance genes for a particular disease in a single genotype so that the duration of resistance can be increased. Similarly, the combination of resistance genes for multiple diseases can prove very effective in tackling any kind of disease epidemic. There have been frequent breakdowns of deployed resistance against major diseases such as rusts, powdery mildew, and spot blotch. This suggests a need for continuous effort to search for novel sources of durable resistance against the emerging virulent races of wheat rusts in available wheat germplasm.

2.1.1 Rusts

Rusts are still a significant biotic stress in wheat. Various researchers describe a capitulate loss of 10–100% in wheat due to rust diseases, which depends on the genotype of the cultivar, whether resistant or susceptible, inceptive infection time, rate of pathogenesis, duration of the disease, virulence factor, and the environment (Singh G. et al., 2017; Bhardwaj et al., 2019). Rust losses can vary from one year to the next and from region to region (Sawhney, 1995). The first stem rust epidemic was reported in 1786 in Madhya Pradesh, a major wheat-growing state in India (Nagarajan and Joshi, 1985), while we have

experienced the continuous incidence of stripe rust in the northern part of India (Gupta and Kant, 2012; Vaibhav et al., 2017). *Puccinia graminis* f. sp. *tritici*, a causative agent of stem rust, resulted in up to 100% yield loss (Leonard and Szabo, 2005). The instantaneous emergence of a novel race of stem rust in Africa called *Ug99* spread to the Middle East, Iran, and other countries, making it a serious concern for global wheat productivity (Singh et al., 2008; Singh R. P. et al., 2015). This race was compatible enough to break down the *Sr31* gene, which has been widely used by breeders against stem rust to have a sufficient level of resistance for over two decades (Pretorius et al., 2000). This pathogen has been rapidly evolving since 1999, resulting in thirteen diverse variants under one lineage (RustTracker.org, 2019). In India, stem rust threatens approximately seven million ha of wheat-growing area (Bhardwaj et al., 2019). Similarly, stripe rust caused by *P. striiformis* Eriks. is dominating the northern provinces of India (Bhardwaj et al., 2019) and takes a heavy toll by reducing annual yields by about 30–50%. The wheat variety PBW343, which was a ruling variety in the North Western Plains Zone of India, has succumbed to stripe rust (Singh R. P. et al., 2017; Bhardwaj et al., 2019). The gross capital loss expected due to leaf rust (Caused by *P. tririciana*) pathogen varies and may be up to 60% under severe conditions (McIntosh, 1998).

Breeding for disease resistance is the most economic and imperishable component of integrated crop disease management (Vasistha et al., 2017; Kumar S. et al., 2015). Approximately, 83, 80, and 61 stripes, leaf, and stem rust resistance genes, respectively, have been curated and cataloged in wheat (McIntosh, 2020). However, the prompt evolution of novel virulent races makes most of the resistance genes ineffective. Unfortunately, the majority of the Indian wheat cultivars lack resistance to stripe rust and their tolerance has been fleeting even though they were evaluated as possessing an adequate level of resistance before being released to farmers. This posed a need for a durable and sustainable solution. In wheat, genetic resistance to rust pathogens can be categorized as follows: 1) all-stage resistance and race-specific or seedling resistance conferred by major genes (Chen, 2013); 2) race-specific adult plant resistance (APR); and 3) partial resistance and slow-rusting or non-race-specific adult plant resistance conferred by minor genes (Johnson and Law, 1973; Das et al., 1992). If resistance genes are used alone, there is a danger of the outbreak of a disease, but if several genes are combined into a single genotype (gene pyramiding), the duration of efficient resistance can be increased. The combination of minor genes with major disease resistance genes has been found to be effective and can attain durable resistance.

2.1.2 Karnal Bunt

Tilletia indica is the causative agent of Karnal bunt, a disease with the greatest impact on the grain and food industry. The disease not only causes yield loss but also adversely affects grain quality due to infested kernels (Fuentes-Dávila and Rajaram, 1994). Grain infested with Karnal bunt attracts quarantine regulations that restrict infested seeds' transboundary movement. It was first reported in Karnal, India (Mitra, 1931), and was characterized as a minor disease till 1968. The disease was further ascertained in innumerable other regions throughout Northern and Central India. Later on, the disease was observed in several other countries, such as Nepal, Afghanistan, Iran, Iraq, Pakistan, Mexico, South Africa, and the United States (Rush et al., 2005). The pathogen infects wheat at the heading stage before seed formation; hence, the symptom is manifested only when the grains are matured in the ear heads. Traditional use of genetic resistance could be the best solution to manage disease severity. Although a huge collection of resistance sources was retrieved from diverse adapted zones, very few of them have been studied for detailed genetic analyses and used in the breeding program. Development of genetic markers, mapping of resistance genes, and characterization of new resistance loci can help to develop improved cultivars using germplasm (Brar et al., 2018).

2.1.3 *Fusarium* Head Blight (FHB)

FHB, or head scab, is caused by different *Fusarium* species where *F. graminearum* and *F. culmorum* are considered dangerous due to the contamination of the grains by mycotoxins like deoxynivalenol (DON), nivalenol (NIV), and zearalenone (ZON). Yield losses occur due to shriveled grain, low test weight, and failure of seed formation. Mycotoxin accumulation is a major concern from an international trade perspective. Mycotoxins, especially DON and its acetylated forms (3-ADON and 15-ADON), make grain unsuited for food or feed upon accumulation (Brar et al., 2019). Although the import or export of FHB-infested wheat grain across international boundaries is allowed by defining a certain threshold, many beverage and food industries have self-imposed regulations (McMullen et al., 2012). Like with other diseases, the adoption of resistant cultivars is the most effective and convenient way to control this disease (Steiner et al., 2017). The complex genetics of FHB resistance makes it difficult to dissect desired resistance because it is under multigene control and associated with genotype × environment interactions. The classic example includes *Fhb1* derived from a Chinese variety, Sumai 3 that provides resistance against FHB was popularized by various breeding programs (Lv et al., 2014). However, *Fhb2* from Sumai 3 (Lu et al., 2010) and *Fhb7* (Guo et al., 2015; Wang et al., 2020) from *Thinopyrum ponticum* have also been used in the resistant breeding program. Meanwhile, pyramiding resistance genes into susceptible cultivars remains a formidable challenge because major sources of resistance genes (such as Sumai 3 and Wangshuibai) are associated with undesirable agronomic traits (Dvorjak, 2014; Li et al., 2016). In the Indian context, it causes notable yield loss if rain coincides with anthesis, which is prevalent in Punjab, Himachal Pradesh, Uttarakhand, and hilly areas of Tamil Nadu. There is a dire need for incorporating resistance against FHB in Indian cultivars, keeping in view the importance of increasing exports of Indian wheat. Germplasm could be the ideal source of resistance for sustainable approaches against such cataclysmic diseases.

2.1.4 Spot Blotch

Spot blotch (SB), a destructive leaf disease of wheat caused by *Cochliobolus sativus* (anamorph: *Bipolaris sorokiniana*), is

considered an economically important disease prevalent worldwide. This disease could result in as high as 70% yield losses under severe epidemic conditions (Ayana et al., 2018). The disease-favoring climate is more prevalent in South Asian and American countries, where warm and humid conditions persist throughout the wheat cropping season (Saari, 1998; Joshi et al., 2007b; Gupta P. K. et al., 2018). From the Indian perspective, the eastern parts are the main epidemic zones, from where it is spread into the cooler traditional rice–wheat areas like the North West Plain Zone (NWPZ) (Villareal et al., 1995; Joshi et al., 2007a). The resistance level in high-yielding wheat genotypes is unsatisfactory and needs to be improved remarkably, mainly in the humid regions of South Asia (Sharma and Duveiller, 2006; Joshi et al., 2007b). Complex quantitative inheritance of SB resistance in wheat has slowed the progress in breeding for SB resistance (Kumar S. et al., 2015). Crucial findings using both bi-parental mapping populations and association mapping panels have accessed assorted SB resistance QTL on all chromosomes except 1D, 3B, 3D, 4A, 4B, 4D, 5D, and 6A (Lu et al., 2016; Gupta P. K. et al., 2018). Nevertheless, only three prime QTLs were assigned, *Sb1* on 7D (Lillemo et al., 2013), *Sb2* on 5B (Kumar S. et al., 2015), and *Sb3* on 3B (Lu et al., 2016). Accessing novel resistance genes by exploiting wheat germplasm could be vital against such a ruinous disease.

2.1.5 Powdery Mildew

Another cataclysmic disease caused by the biotrophic fungus *Blumeria graminis* (DC) E.U. Speer f. sp. *tritici* Em. Marchal (Syn. *Erysiphe graminis* DC f. sp. *tritici,* Em. Marchal) is powdery mildew (PM) of wheat, a foliar disease of universal occurrence resulting in dreadful yield loss (Mwale et al., 2017). Its severity usually climaxes in areas with high precipitation and a maritime-like climate (Bennett, 1984). However, it has gained importance in other regions due to the application of a higher dose of nitrogenous fertilizer and the cultivation of modern semi-dwarf wheat genotypes (Wang et al., 2005; Morgounov et al., 2012). As far as yield losses are concerned, they range from 15 to 40% depending upon the varieties and climatic conditions. Earlier, this disease was confined to the North Hill Zone (NHZ) of India, but now it is also spreading toward the North-Western Plains Zone (NWPZ) of India due to climate change, which has led to the development of new races. The gene with the highest level of resistance was studied on the wheat-rye translocation (1B/1R) fragment that originated from the cultivar Veery. This cultivar was substantially used to develop many PM-resistant cultivars around the globe (Friebe and Heun, 1989), including India (Vikas et al., 2020). Also, 68 loci providing resistance against wheat PM have been mapped to various wheat chromosomes (McIntosh, 2020). Moreover, the finding of several PM resistance genes, research should be carried out on finding novel resources to unravel genes, alleles, and SNPs because of the breakdown of truly race-specific resistance genes due to the emergence of new pathotypes or races (Hsam et al., 2003). Hence, it is necessary to explore adaptive wheat germplasm to identify refreshed resistance genes and use them against rapidly evolving pathogens.

2.1.6 Other Pathogen/Pests

The yield losses due to insect pests have increased in the post-Green Revolution era (Dhaliwal et al., 2010). Unlike biotic stress resistance, the resistance gene had a very minor contribution in protecting wheat against insect pests because of the high impact of environmental conditions like temperature and light on the survival and behavior of the insects (Alford et al., 2014). Thus, a more concerted effort and methodology are required to identify and recruit the most effective insect pest resistance genes. Amongst the different pests, aphids, wheat weevil, wheat midge, termites, Hessian fly, armyworm, and cereal cyst nematode (CCN) are important arthropods feeding on wheat. CCN is becoming a serious threat to wheat production in several states of India (Singh and Kaur, 2015; Smiley et al., 2017). There is a higher perception of the CCN threat due to the fact that the identified genes confer a limited level of resistance to specific CCN pathotypes. Since the molecular mechanism of known genes is not known, it is essential to identify new genes and understand their interactions and functions in conferring resistance to CCN. **Figure 1** basically explains the challenges faced by biotic and abiotic stresses in wheat production for the Indian wheat-breeding program which in turn accelerates or develops genebank genomic selection models using a combination of genebank genomics, genetic diversity, population structural analysis, and genomic sequences with phenomics, precise trait data, high throughput trait data, and speed breeding which is effective for speeding up the use of wheat germplasm lines coupled with breeding to enhance the process of variety development club with the genomic selection.

2.2 Abiotic Stress

Abiotic stresses are equally important which limit wheat production worldwide. Among abiotic stresses, salt, drought, and terminal heat stress are the three utmost constraints for successful wheat production in most parts of India. Climate change has been shown to have a high impact on wheat yield due to rising temperatures and water scarcity in India and other wheat-growing regions of the world. Wheat grain filling is suppressed at a temperature above 30°C because of reduced starch synthase activity (Jenner, 1994). Short-term extreme increases in temperature of 5–10°C can have quite catastrophic effects on yield, as an increase in the ethylene signal after heat spikes has been shown to lead directly to grain abortion (Hays et al., 2007). Moreover, it is estimated that a 1°C increase in temperature can result in a 10% decrease in wheat productivity in low-altitude countries (Lobell et al., 2011). Heat stress at the terminal stage of the wheat crop is a crucial abiotic stress that restricts plant growth and the accumulation of starch, which in turn causes yield unpredictability in many wheat-growing parts of the world (Gupta et al., 2012). Different reports have predicted that the average global temperatures will increase in the coming years (Malhi et al., 2021). With an estimated rise in global temperatures of up to 1.5°C by the year 2030 and a 1.8–4°C rise by the end of the century, the challenges facing wheat production are enormous and need to be tackled immediately. In addition to this, the pattern of diurnal and nocturnal

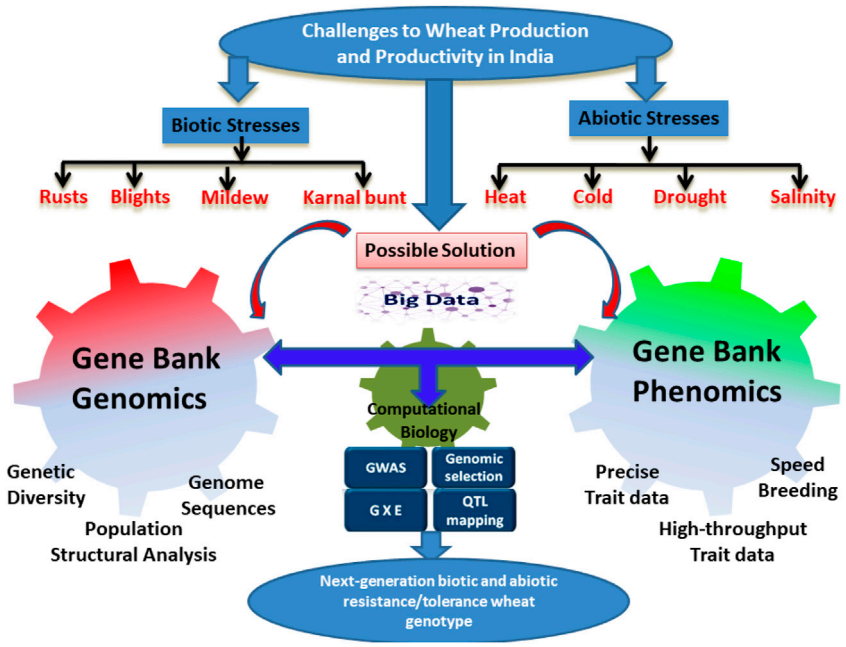


FIGURE 1 | Challenges faced by Indian wheat-breeding programs and the possible solution. The figures show different biotic and abiotic stresses limiting wheat production and productivity in India. The figures also explain how genomic interventions and gene bank phenomics could be used in development of next-generation wheat varieties with enhanced biotic and abiotic stresses.

temperatures has also started changing, which is resulting in warmer nights (Gupta et al., 2012). Heat stress, particularly at the terminal stage of wheat, is the major limiting factor for plant productivity and is a major cause of yield instability in many parts of the world. Changes in temperature patterns accompanied by unpredictable rainfall patterns are also affecting crop productivity in several countries, including India (Singh G. et al., 2018). The threat perceptions due to the impact of global climate change on agriculture are going to be huge in the coming decades. There is a complex genetic basis for most of the improved traits in wheat related to water-deficient and heat-stress conditions (Sallam et al., 2019). This is because each of these traits is polygenic and each gene has a small effect. Improvement of polygenic traits is itself a difficult task (Ranjan et al., 2021). The genetics behind abiotic stresses is more complex as compared to that of biotic stress. Although many studies have been conducted to elucidate the genetics of these traits, only limited success has been achieved in utilizing the vast wealth of data in the crop improvement programs. Global warming is severely affecting weather patterns, resulting in extremes of temperature, drought, frequent frost, and snowfall in high altitudes (IPCC. Climate Change, 2013). In the last few years, droughts and heatwaves have become frequent in a large part of India, posing a serious threat to future wheat production. In an estimate, the average yield loss of wheat in India due to a 1°C rise in temperature is reported to be 9.1 ± 5.4%, while the global yield loss triggered by the same is projected to be 5.5%, accounting for an aggregate loss of 35 M tons (Wang et al., 2018). This calls for the deployment of varieties that can withstand heat stress during the anthesis and seed setting stages. Furthermore, the accelerated use of wheat germplasm in a sustainable and planned manner is a viable option for addressing biotic, abiotic, and malnutrition threats.

In the last few decades, due to drastic changes in climatic conditions, most of the world faced low water accessibility, especially in South Asia and Africa. Among all the abiotic stresses, drought and terminal heat stress are the major limitations to food production worldwide, including India. Hence, developing genotypes that hold terminal heat tolerance is one of the crucial precedents of wheat improvement programs in India. The continuous shrinking of water resources around the world has further compounded problems, in addition to thermal stress, leading to reduced production and productivity (World Meteorological Organization, 1997), and there is a need for additional sustainable approaches to increasing productivity on restricted land, which will prevent the detrition of biodiversity (Pretty and Bharucha, 2014). Climate change is predicted to have a high impact due to rising temperatures and water scarcity in the densely populated regions of India. In many of the global wheat-growing areas, drought and terminal heat stress cause maximum damage.

2.2.1 Terminal Heat Stress

Around nine million ha of wheat in the subtropical or tropical zone (Lillemo et al., 2005) are heat stressed in countries including India, Bangladesh, Uganda, Nigeria, Sudan, and Egypt that have traditions of cultivating wheat since long ago (Abdelmageed et al., 2019). An estimate suggests that India's 13.5 million ha of wheat cultivated land comes under a heat-stressed zone (Joshi et al., 2007a). Terminal heat stress is one of the measures of sudden remarkable enhancement in temperature during the grain filling

TABLE 1 | Perceptual dissemination of trait-specific Indian wheat germplasm collection.

SI no.	Important trait	Accession example	Reference
1	Drought tolerant	Safed mundri and Lal mundri and Jautri	Pal et al. (2007) and Gupta and Kant (2012)
2	High yield	Jhusia, Kishva, Churi, and Farmi	Panwar et al. (2014)
3	Drought and high biomass	Bhuri mundiya	Mehta et al. (2009)
4	Softness and good biscuit-making quality	Naphal	Ram et al. (2007) and Gupta et al. (2009)
5	Tastier chapati	Lal gehun	Gupta et al. (2009) and Mehta et al. (2009)
6	Dalia and fodder	Rata and Bhati	Gupta et al. (2009)
7	Small grains and long awns	Tank	Gupta et al. (2009) and Mehta et al. (2009)
8	Salt tolerant	Kharchia	Díaz De León et al. (2010) and Goyal et al. (2016)
9	Two forms of ear head color	Kathia	Arora and Koppar (1979)
10	Long culm	Jautri	Arora and Koppar (1979)
11	High elevation adaptation	Bhotia	Tripathi et al. (2018) and Mehta et al. (2019)
12	Valley adaptation	Chanosi	Mehta et al. (2011)
13	Drought tolerant	Dapati	Pande et al. (2016)
14	Excellent chapati quality	Daulatkhani	Mehta et al. (2011) and Mehta et al. (2019)
15	Mid-hill adaptation	Dudh gehun	Panwar et al. (2014) and Mehta et al. (2019)
16	Hailstorm tolerance	Lakha	Tripathi et al. (2018)
17	Mid- to higher-elevation adaptation	Lal mundia	Kumari et al. (2018) and Kumari et al. (2019)
18	Awnletted	Mundia	Tripathi et al. (2018) and Mehta et al. (2019)
19	Grain boldness	Thanga	Mehta et al. (2019)
20	Terminal heat tolerant	Halna	Singh et al. (2005)
21	Grain yield	Bawaji	Gami et al. (2011)
22	Non-shattering	Kankoo and Dharmauri	Panwar et al. (2014)
23	High tillering	Dharnon and Shruin	Panwar et al. (2014)
24	Long spike	Dholia and Katta	Panwar et al. (2014)

stage till maturity. The mean temperature above 31°C during caryopsis ripening in wheat comes under the influence of terminal heat (Kumari et al., 2015; Dubey et al., 2020). Due to climatic fluctuation, the commencement of early summer than normal and late sowing of wheat due to a mixed cropping system are the possible factor for terminal heat stress in wheat (Gupta et al., 2012). Terminal heat stress causes severe damage to wheat, which alters its physiology and grain filling mechanism. Intense high-temperature waves are likely to become more damaging if the current trends continue and future predictions about global warming hold true. Notably, it significantly impacts starch synthesis and accumulation which is a measure of grain filling rate and gross productivity declined by the sudden outbreak of heatwave during caryopsis development (Jenner, 1994; Kumar et al., 2016c). Furthermore, current approaches for crop management utilize the application of irrigation water, which can reduce heat stress on plants (Badaruddin et al., 1999) but is not feasible for large areas. To date, our limited understanding of the complex interaction of cellular/molecular mechanisms with whole-plant adaptation has restricted deterministic approaches to breeding for heat tolerance (Reynolds et al., 2021). Germplasm could be a reliable source of gene or QTL for heat tolerance, especially at the ripening stage and seed maturation. Additionally, it could be managed by the introduction of trait-like late maturity genotype, stay green-harboring germplasm, and their wider use for adaptability against the current scenario.

2.2.2 Drought Stress

Drought is the second most serious abiotic stress limiting wheat production in different parts of the world and occurs with varying frequencies (Boyer, 1982; Chaves et al., 2003). Drought affects wheat crops more frequently in tropical and subtropical regions, where most of the developing countries are situated. Around 17% of the cultivated wheat areas worldwide were affected by drought during the period of 1980–2006 (Dai, 2013). In India, 29% of the total cultivable area faces drought conditions, of which 10% is under severe drought (Anonymous, 2003). This has caused an estimated 20–30% reduction in total wheat yield in stressed areas. Reduced bioavailability of water across the heatwave at the terminal growth phase of the spike is negatively correlated with productivity. Basically, they both occur at the same time, and their additive effect causes aborted grain filling (Sattar et al., 2020). Drought stresses impact on their own or in combination that significantly affect several agronomical features like heading days, the height of the plant, numbers of tiller per plant, and length and occupancy of the spike. However, an indirect correlation was suggested in terms of expressed results of GWAS or mapped QTL possibly due to a paradoxical association between traits and genetic loci (Tahmasebi et al., 2016; Abou-Elwafa and Shehzad, 2021). Therefore, a significant effort will be required, including molecular tools to breed superior drought-tolerant varieties. Whole-genome sequencing for each genotype was not possible earlier, but the commencement of high-throughput sequencing technology makes it accessible for extreme landrace and exotic lines (Khadka et al., 2020). The number of genes contributing directly or indirectly to drought tolerance relies on the associated traits' magnitude and proximity of the genes associated with the markers. Genome-wide association studies (GWAS) or QTL mapping could be used to identify genes, involved in drought tolerance in unexplored germplasm, which could then be used to improve crop drought tolerance (Zeng et al., 2014; Sukumaran et al., 2018). Furthermore, advanced breeding programs for crop improvement are assisted by genomic selection and gene

editing for improving drought tolerance in wheat (Singh S. K. et al., 2015).

2.2.3 Salinity

Among abiotic stresses, increased soil salinity and sodicity pose a challenge to agriculture. The high salt concentrations of the soil can be attributed to the poor land and water management practices as well as the lack of soil reclamation processes in many parts of the world. In India, approximately 8.6 mha of the cultivated land is affected by soil salinity. Furthermore, the areas under salinity are expanding each year due to low precipitation, mixing with the coastline, saline water irrigation, high surface evaporation, and poor cultural practices (Jamil et al., 2011). It has been extrapolated that about 50% of the cultivated land area may be impregnated with salt by the mid of twenty-first century (Mahajan and Tuteja, 2005; Sharma et al., 2012). Salinity tolerance could be accessed by using conventional (El-Hendawy et al., 2005) to modern spectral imaging techniques (Moghimi et al., 2018). Although, most of the affected parameters were known for salt tolerance which limits productivity, inadequate large-scale phenotyping could be a possible factor for a significant outcome (El-Hendawy et al., 2005; Rosenqvist et al., 2019). Finding well-studied genes/transcription factors from wheat germplasm like AVP1, NHX2, DREB, and SHN1 and their associated marker (Díaz De León et al., 2010; Goyal et al., 2016; Singh A. K. et al., 2018; Kumar et al., 2020c) for the utilization for tolerance breeding could be a sustainable approach for generating salt-tolerant wheat genotypes (Choudhary et al., 2021). Hence, there is a compelling need to develop salt-tolerant wheat varieties. Although in the context of salt tolerance germplasm utilization, there is little progress yet, notably germplasm which harbors extreme salt-tolerant genes could be rescued for generation of pre-breeding lines for crop breeding which could stand against the high saline condition.

2.3 Nutritional Quality Traits and Nutrient Use Efficiency

Wheat genetic resources are an ideal solution for addressing the issue of nutritional security in the Indian population. Malnourishment exists both in underprivileged rural populations as well as in wealthier urban populations, where anemia is a major health challenge in children and women (Müller and Krawinkel, 2005; Sethi et al., 2020). Due to its consumption by a major chunk of the Indian population, the development of iron, zinc, and protein fortified wheat is well justified as it can provide these essential micronutrients and proteins through routine edible product intake. (Borrill et al., 2014; Balk et al., 2019). The main objectives for the quality improvement are the enhancement of protein contents, bio-fortifying with essential amino acids which are basically absent in wheat, elevation in flour quality by modifying starch and glutenins, and elimination of anti-nutrient factors like phytic acid and polyphenols (Grewal and Goel, 2015; Adhikari S. et al., 2022). Although basic research on flour quality was documented, a translation aspect for wheat improvements would be fruitful. Screening of massive wheat germplasm for quality traits and their utilization for quality breeding would be an appropriate sustainable solution (Joshi et al., 2007b; Ramadas et al., 2019). However, breeding in wheat is quite difficult due to complex genetic and metabolic networks, differences in wheat plants' micronutrient use efficiency, translocation coherence, source-sink relationship for metabolite allocation and partitioning, and genotype-dependent metabolite translocation (Ramadas et al., 2019). Hence, for efficient breeding, it is necessary to understand the genetic basis of micronutrient accumulation in grains and, accordingly, explore the conserved collection for suitable resources. Plant nutrients, including nitrogen (N), phosphorus (P_2O_5), and potash (K_2O) are the major and most salient nutrients required by the plants (Zörb et al., 2018; Rajičić et al., 2019). The genetic architecture of the plants plays an important role in fertilizer uptake (Sandhu et al., 2021). Therefore, different genotypes respond differently to the amount of supplied nutrients (Mkhabela et al., 2019). Not only agronomic practices but also breeding plays an important role in improving nutrient-use efficiency. As a result, nutrient-use efficient lines/varieties can be developed by modifying root architecture, stem phenology, and leaf phenology (Dharmateja et al., 2021). The variable germplasm with miscellaneous structures, viz., deep root systems, enormous taproots, and the diverse shapes of roots previously adopted to low nutrient soil needs to be assessed under highly precise and uniform conditions. The ratio of a different nutrient may be studied for better uptake and efficiency. Using the precision nutrition platform, a large number of genotypes could be evaluated with high precision and accuracy. Overall, Indian wheat germplasm could be served for such unusual traits (**Table 1**), which could be useful after being incorporated into desired wheat genotypes.

3 THE WEALTH OF INDIAN WHEAT GENETIC RESOURCES AND ITS CONSERVATION STATUS

Archaeological and botanical evidence reveals the domestication center of einkorn (*Triticum monococcum*) and emmer (*Triticum dicoccum*) to be in the Mesopotamian crescent of the Near East at about 7500 BC (uncalibrated) and from there, it spread to the Middle East, Asia, and North Africa, and ultimately Europe, America, and South Africa (Gustafson et al., 2009). In an evolutionary context, the A genome of wheat is predominated and domesticated earliest during wheat evolution. It circumscribed to cultivate as wild einkorn. With the introduction of large-scale genomics analysis like genotype by sequencing, SNP array revealed that the origin of *T. urartu* is the closest genome for subgenome A (Maccaferri et al., 2019; Adhikari L. et al., 2022). Furthermore, with the commencement of remarkable genetic diversity losses in the pericentromeric and donating B genome by *T. speltoides* give a new tetraploid *T. turgidum*. A second hybridization event between the resulted tetraploid and third D genome donor followed by chromosome doubling has occurred to gain hexaploidy or *T. aestivum* (Gustafson et al., 2009; Maccaferri et al., 2019). Widely cultivated with dynamic adaptability,

wheat can be grown at varying altitudes ranging from the sea level to 4500 m above the mean sea level (AMSL) under diverse agro-ecological conditions. Currently, it comprises several high-yielding varieties suitable for a wide range of environments, ranging from the low-humid regions of India, Nigeria, Australia, and Egypt to the highly humid regions of South America (Damania et al., 1997). Currently, the Asian continent is the leading wheat producer. For example, the total area under wheat crops is nearly 31 million ha, divided into three ecozones: the Northern Himalayan Zone, the Central Zone, and the West South Zone (Kulshrestha, 1985; Singh et al., 2007). The crop gene banks came into existence in response to the growing concern over the rapid erosion of agro-biodiversity due to the preference of superior modern cultivars over landraces and indigenous lines (Díez et al., 2018). Recognizing and deploying relevant genetic and genomic variation from wheat germplasm stored at gene banks to breeding programs is an important strategy for sustaining crop genetic improvements and conserving genetic diversity (Mir et al., 2012; Sehgal et al., 2015; Mondal et al., 2016). Recent next-generation studies have charted new approaches for eliminating redundant duplication in large gene bank collections, thus facilitating the availability of manageable collection sizes for effective molecular breeding (Mascher et al., 2019; Singh et al., 2019). Gene banks around the world maintain a huge collection of wheat germplasm (Prada, 2009). The Indian wheat genetic resources are collectively conserved in its National Gene Bank (NGB) located at the NBPGR, New Delhi.

The National Bureau of Plant Genetic Resources (NBPGR) is an official institute at the national level for the governance of plant genetic resources (PGR). Its headquarters is in New Delhi. The Indian NGB housed at ICAR-NBPGR has also been reported to store more than 31,000 wheat accessions, which include landraces, exotic lines, and indigenous collections (Tyagi, 2016). The major mandate of the institute is to intend, assemble, and coordinate exploration and collection of native and exotic plant genetic resources from extreme environments for sustainable agriculture and to introduce, exchange, and supervise intellectual property right-based quarantine of plant genetic resources.

These are of infinite value for agriculture, food, research materials, human resources development for sustainable agricultural growth, boosting the efficient use of genetic and genomic resources of cereals, pulses, and other orphan and ornamental crops, and allied research (Singh S. et al., 2018). In addition, coordinating, capacity building in PGR management, germplasm policy access, and sharing social benefits are also pivotal. Genetic and molecular profiling of agri-horticultural crops, genetically modified plant (GMP) detection technology research, and development of information networks on plant genetic resources (Tyagi, 2016; Singh, 2018) are also mandated activities of NBPGR. Currently, the NGB of India has the largest collection of wheat in the Asian region, with around 34,000 accessions (of 51 species) in its long-term storage (data not ported in Genesys). This collection has over 18,000 indigenous and 14,000 exotic accessions (Tyagi, 2016). The wheat genetic resources are further complemented by other institutes within the National Agricultural Research System (NARS), viz., the Indian Institute of Wheat and Barley Research (IIWBR) in the Karnal district of Haryana, Punjab, Agricultural University in the Ludhiana district of Punjab, and the Indian Agricultural Research Institute (IARI), New Delhi. India's national wheat germplasm collection is genetically rich in its species diversity and indigenous wealth. It has around 2,000 accessions belonging to the category of traditional cultivars'/landraces'/farmers' varieties, drawn from diverse ecological zones within the country. The states of Uttarakhand, Himachal Pradesh, Uttar Pradesh, Rajasthan, Gujarat, and Madhya Pradesh are the major areas from where these indigenous resources have been collected and conserved over the past 5 decades (**Figure 2**). All these genotypes are treasure mines of unique genes related to several economically important traits. For example, as annexed in **Table 1**, drought-tolerant genotypes are Safed and Lal mundri, high yield-producing genotypes are Jhusia, Kishva, and Churi, excellent chapatti-making genotypes including Kankoo, Dharmauri, and Lal gehun, and biscuit-making quality exhibited by Mishri and Naphal. However, Bhati could be utilized as excellent fodder for livestock and Bhuri mundiya for high biomass. Additionally, augmentation efforts have been made through repatriation of Indian origin accessions from the USDA Gene Bank, Australian Grains Gene Bank, and John Innes Institute, United Kingdom. The repatriated germplasm resources comprise landraces, wild species, and relatives, which are the critical components of the wheat improvement program.

These accessions might prove very useful in the development of high-yielding and climate-resilient wheat varieties if the gene/germplasm is deployed in adapted cultivars in a planned way. However, only a small proportion of this collection has been utilized in breeding programs to date, primarily due to a lack of information about the traits and associated genes/markers in this collection. The evaluation of huge gene bank collections for the targeted traits is a costly and labor-intensive task. Although recent efforts have been made to develop core sets based on agro-morphological traits, these may not accurately represent the original collection's diversity because agro-morphological parameters are influenced by environmental conditions (Dutta et al., 2015). Precise characterization and documentation of these valuable germplasm lines are prerequisites for germplasm utilization in breeding and genomics studies (**Table 1**). These germplasms have been characterized by several traits in recent years, indicating that a reasonable number of indigenous germplasm lines are tolerant to both biotic and abiotic stresses because they have co-evolved with their environments for a long time.

4 IDENTIFICATION OF TRAIT-SPECIFIC GENETIC RESOURCES FOR DISEASE RESISTANCE, NUTRITION, AND CLIMATE RESILIENCE

Global warming is severely affecting weather patterns, reflecting extreme heat, drought, frequent frost, and snowfall in high

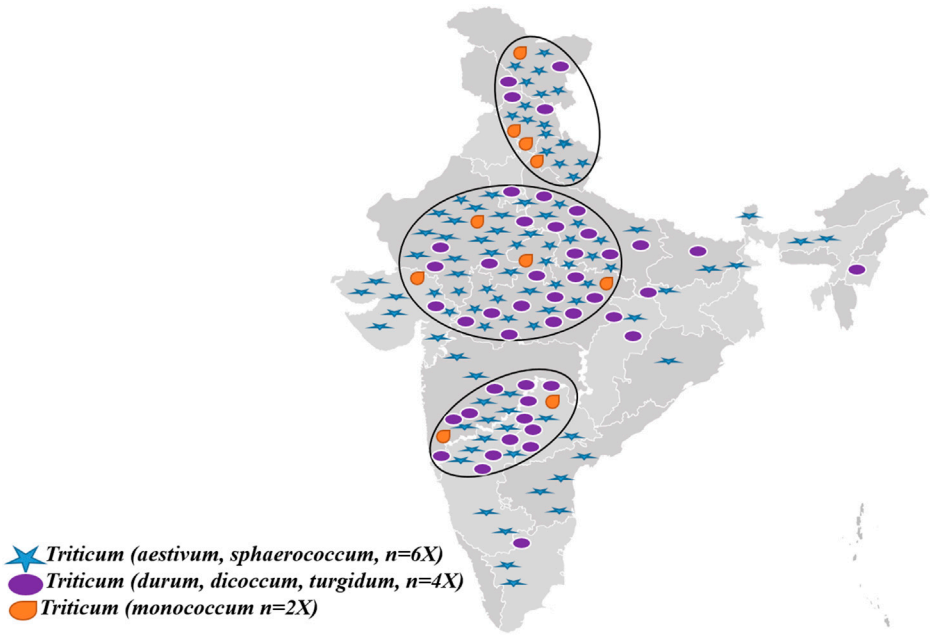


FIGURE 2 | Prevalence and collection sources of diverse ploidy wheat genetic resources in India.

altitudes (IPCC, 2013). In the last few years, drought and heatwaves have become frequent in a large part of India, posing a serious threat to wheat production. During 2014–15 and 2015–16, wheat production was far below the expected target due to severe drought in various regions of the country. Soil salinity and sodicity are also anticipated to increase from the current 6.73 million ha to 20 million ha by 2050 (Sharma et al., 2012). The wheat-breeding program for abiotic stress tolerance, especially for drought and heat stresses seems to be challenging. The slow genetic progress accomplished to date is a consequence of non-adaptive genotypes with concerned environments, yield constituent compensation, the enigmatic origin of drought, and heat tolerance (Maich et al., 2007). Landraces have long served as the source of traits for local adaptation, tolerance to various stresses, yield stability, and optimum nutritional profile. Evaluation of landraces and local germplasm for finding traits pertaining to abiotic stress tolerance, and their deployment in the elite breeding lines could be the best strategy. Although, in the recent decade, scientists from NBPGR and their collaborators worldwide have made huge efforts to harness the genetic potential prevalent in wheat genetic resources (**Table 2**), a large proportion is still untouched in the context of trait identification and omics study.

Hays et al. (2007) report that high temperatures during grain filling can cause a yield potential loss of up to 40% under dreadful stress. Drought is limiting wheat production in different parts of the world (Fahad et al., 2017; Abhinandan et al., 2018). Globally, about 17% of the wheat cultivated area is distressed by drought (Dai, 2013). In India, the fraction of total cultivable land affected by drought is 29%, of which 10% is under severe drought (Anonymous, 2003). Water has emerged as a limiting factor for sustained cultivation of wheat and other crops in various parts of India, even in the water-rich Indo-Gangetic Plains (Joshi et al., 2007a). Therefore, the generation of drought-tolerant varieties through breeding is essential for achieving enhanced crop productivity and food security for the hundreds of millions of people living in rural areas (Ortiz et al., 2008). Excellent drought-specific markers were identified to determine tolerance against droughts such as *Dreb* and *Fehw3* (Rasheed et al., 2016). Consequently, the existence or absence of *Dreb* and *Fehw* markers can be analyzed in any promising germplasm. Several drought-tolerant lines have been identified in India (Kumar et al., 2018), which can be used for favorable allele mining. The identification of novel genetic loci for the improvement of drought tolerance can be achieved by GWAS or QTL mapping using germplasm lines (Zeng et al., 2014; Sukumaran et al., 2018) (**Table 2**).

5 AVAILABILITY/DISCOVERY/ DEVELOPMENT OF GENOMICS RESOURCES

Over the past decade, there has been a substantial advancement in the development of genomic tools and techniques in wheat (Alaux et al., 2018; Purugganan and Jackson, 2021). The wheat gene pool possesses a tremendous amount of genetic variability for a trait of interest. Several high-density genetic and physical maps of wheat have been developed (Chao et al., 2007). The release of the gold standard reference genome assembly of wheat into the public domain will expedite the use of genomic resources in breeding (IWGSC, 2018). Moreover, high-throughput genotyping tools such as SNP arrays and GBS platforms have also been developed. In recent years, there has been an outburst of

TABLE 2 | Recently harnessed wheat germplasm for biotic and abiotic stresses and nutrient-use efficiency to improve crop productivity.

Sl no.	Evaluated germplasm	Trait	Reference
1	19,460 germplasm lines were evaluated for wheat powdery mildew	Powdery mildew resistance	Vikas et al. (2020)
2	Auspicious 47 germplasm accessions, including 15 durum species were studied for heat stress	Heat tolerance	Sareen et al. (2020)
3	169 accessions and wild relatives of wheat, viz., *Elymus* L. (8/139 acc.), *Hordeum* L. (2/4), and *Leymus* Hochst. (2/26) were assessed for cold tolerance	Cold stress tolerance	Pradheep et al. (2019)
4	287 wheat association mapping initiative (WAMI) panels used for spot blotch resistance mapping	Spot blotch resistance	Ahirwar et al. (2018)
5	Germplasm genotypes ET127225, ET127230, EC531185, ET127236, ET127267, and ET127269 exhibit a good level of drought stress tolerance	Drought tolerance	Kumar et al. (2018)
6	Accessions IC564121, IC529684, IC443669, IC443652, IC529962, IC548325, and EC178071-331 were highly resistant to spot blotch	Spot blotch	Kumari et al. (2018)
7	19,460 wheat germplasm accessions evaluated for rusts and spot blotch	Rust and spot blotch resistance	Kumar et al. (2016b)
8	The selected accessions (IC445595, IC543417, IC252650, IC310590, IC539561, IC443636, and IC75246) were evaluated and found superior for grain yield, 1,000-grain weight, and heat stress tolerance	Terminal heat tolerance	Kumari et al. (2015)
9	Out of 267, 239 accessions of *Aegilops tauschii* were resistant to stem rust	Stem rust resistance	Vikas et al. (2014)
10	Phenological and agronomical features of elite germplasm were screened against spikelet fertility	Fertility	Meena et al. (2013)
11	An inter-varietal RIL mapping population of the cross "C306" × "HUW206" was evaluated for drought stress	Drought tolerance	Kumar et al. (2012)
12	Germplasm lines for multiple disease resistance were identified	Rusts, foliar blight, and Karnal bunt	Sharma et al. (2012)
13	Germplasm lines were screened for yellow rust resistance genes (*Yr5, Yr10, Yr15*, and *Yr18*)	Stripe rust resistance	Mukhtar et al. (2015)
14	Cold and drought tolerance was observed in TW 9336, RL 111 P2, and RL 124–2 P2 along with the high grain yield and harvest index	Drought and cold stresses	Gupta et al. (2005)
15	The alleles *Barc 1*, *Barc 26*, *Barc 77*, and *Barc 147* were used to screen 41 genotypes for diversity	Genetic diversity analysis	Kumar et al. (2016a)
16	Six genotypes (IC 542394, IC 542391, IC 542416, IC 542431, IC 542426, and IC 542387) were high in Fe and Zn content	Micronutrient concentration	Krishnappa et al. (2019)
17	Drought stress tolerance was accessed in wheat germplasm	Drought tolerance	Jamla et al. (2017)
18	Six resistant landraces, viz., IC266831, IC266872, IC393109, IC392578, IC444217, and IC589276, were identified against pests	Pest, a weevil (*Sitophilus oryzae*)	Tripathi et al. (2017)
19	Multi-environmental evaluation of wheat germplasm identifies potential donors for disease resistance	Fungal resistance	Kumar et al. (2019b)
20	Exotic line characterization for disease resistance	Rusts and spot blotch	Kumar et al. (2020b)
21	Five genotypes from advanced Indian wheat breeding material were found resistant against rusts	Yellow rust and powdery mildew	Sood et al. (2020)
22	Stripe rust resistance was observed in eight genotypes including DWR 16, VL616, UP212, HD2281, HD2307, K65, Lal Bahadur, and HD2329	Stripe rust	Gupta V et al. (2018)
23	Five landraces VHC(BD)2, VHC6185, VRB-CW-2106, VHC6178, and VAH-CW 3166 revealed seedling and adult plant resistance	Stripe rust	Raghu et al. (2018)
24	IC-368665, IC-78696, IC-75352, IC-104550, IC-75354, IC-36867, IC-572071, IC-104561, 145,953, and IC-59137 exhibit QTL for stay-green trait	Terminal heat	Kumar et al. (2016c)
25	Nutrient use efficiency was observed in BW66, BW103, BW104, BW143, and BW183	Phosphorus-use efficient	Dharmateja et al. (2021)
26	Genotypes Glu3 and PBW343 + Glu acquired allele for grain protein content and test weight	Flour quality	Vishwakarma et al. (2015)
27	About 35 exotic genotypes express slow resistance to stripe rust	Stripe rust	Singh G et al. (2017)
28	Waterlogging tolerance was found in DUCULA 4, CUNDERIN, KRL 105, HD3086, RW3684, BH 1146, DBW39, 52, NW1014, NW 1067, NW 4081, PBW 621, PBW 631, PBW 590, HD 2967, HD 2997, and NW 4083	Waterlogging	Singh R. P et al. (2017) and Singh S et al. (2018)
29	IC611273, IC611071, IC75240, IC416188, IC321906, and J31-170 manifest against abiotic stress	Heat stresses	Kumar et al. (2020a)
30	About 36 wheat genotypes and three triticales were resistant against stem rust pathotype, *Ug99*	*Ug99* (stem rust)	Sharma et al. (2015)
31	Characterized wheat germplasm for puroindoline proteins (antimicrobial)	Antimicrobial properties	Chugh et al. (2015)

innovations in the field of "genomics" which can be employed for the identification of genes or genomic regions for useful traits from a large set of germplasm collections conserved in gene banks (Crossa et al., 2016; He and Li, 2020). The important ones to mention are high-throughput genotyping assays, whole-genome sequencing (WGS), GWAS, and genomic selection (GS) (Muleta et al., 2017). Of these, GS is of special interest and has emerged as a promising approach for genetic improvement of complex traits

TABLE 3 | Identified candidate gene/transcripts, QTL, and MTA, etc., from Indian wheat germplasm.

Candidate gene, QTL, transcript, and MTA	Desired trait	Reference
Pina and Pinb gene	Grain hardness	Chugh et al. (2015) and Kumar R. et al. (2015)
tsn 1 gene	Tan spot resistance	Phuke et al. (2020)
APR gene detection	Leaf rust	Kumar et al. (2019a)
pgd3 gene	Heat stress	Rangan et al. (2020)
(TA 5088 and TA 5638) Alien chromosome	Drought tolerance	Djanaguiraman et al. (2019)
SNPs	Spot blotch resistance	Ahirwar et al. (2018)
Glu-B1	Protein quality	Routray et al. (2007)
Putative *Pm3c*	Powdery mildew	Basandrai et al. (2016)
Alleles 12*, 12.1*, 12.1, 12.2, and 12.3	Glu-D1 locus	Goel et al. (2018)
vrn-B3, Vrn-A1c, GluB3i, GluB3g, and GluA3b	Vernalization and glutenin	Sehgal et al. (2015)
MTAs (2AS, 1BS, 6BS, and 7BL) significant (2NS/2AS translocation)	Head blast resistance	He et al. (2021)
MTAs on 2A, 3A, 1B, 2B, 3B, 4B, 5B, 6B, 2D, and 3D	Stripe rust resistance	Pradhan et al. (2020)
Multi-trait SNPs on Chr2BS, Chr1Ds, and Chr2DS	Agronomic traits	Kumar et al. (2020a)

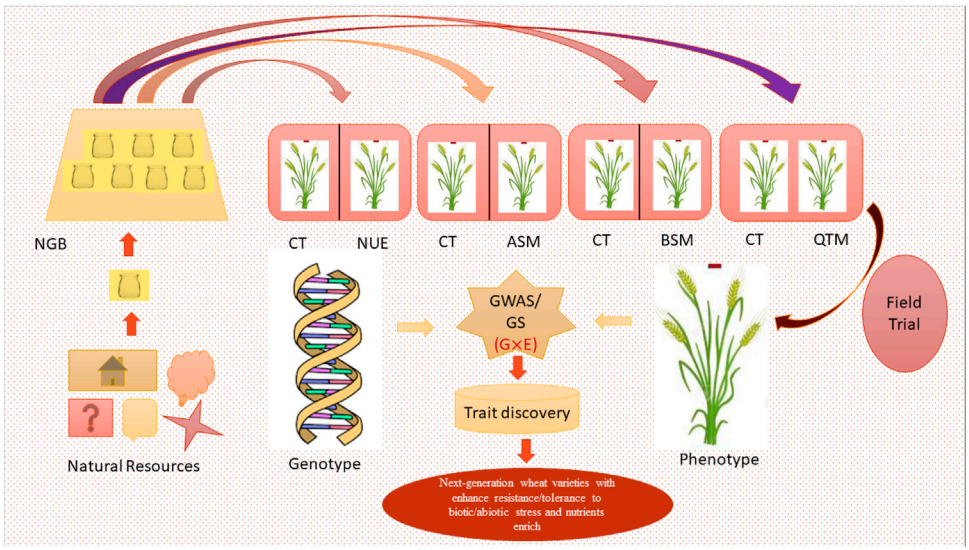


FIGURE 3 | Roadmap for germplasm assessment and trait discovery using integrated analysis of the genotype and phenotype (NGB—National Gene Bank; CT—control; NUE—nutrient-use efficiency; ASM—abiotic stress measurement; BSM—biotic stress measurement; QTM—quality trait measurement; GWAS/GS—genome-wide association study/genomic selection).

(Ali and Borrill, 2020; He and Li, 2020). GS could be used for large plant breeding populations with genome-wide molecular markers to predict the total genetic value for complex or economically important traits such as yield. The key conceptual difference between conventional breeding and genomic selection approaches is that in the former, selections of candidate varieties are based on the observed phenotypic performance, whereas, in the latter, selections are based on the genetic makeup and genotype × environment interaction (Crossa et al., 2016; He and Li, 2020). A robust theoretical and experiential report suggests that GS methods can predict performance with adequate accuracy to allow selection based on molecular markers alone (Ali and Borrill, 2020). Furthermore, GS is a promising approach for accelerating the rate of genetic gain in plant-breeding programs by enabling selection for complex traits (like yield under heat stress) early in the breeding cycle and therefore reducing the cycle time, which increases the annual gain. Genomic selection has a potential breeding strategy to map numerous genetic loci for diverse traits of interest. Various research groups started working on Indian wheat and associated germplasm genotype for crop improvement against biotic resistance (Juliana et al., 2021; Budhlakoti et al., 2022). But still, more accurate prediction from a large genotype reservoir of Indian wheat germplasm is necessary for germplasm-assisted crop improvements for abiotic and quality-related traits.

In GS, genome-wide molecular markers are used to predict total breeding values called genomic estimated breeding value (GEBV) and make selections of individuals or breeding lines before phenotyping (Larkin et al., 2019; Kumar et al., 2021). This approach has several advantages, especially for 1) making

selections before phenotypic evaluation, which reduces the time needed to make selections and 2) increasing the size of breeding populations since genotyping of a large number of lines can be carried out at a lower cost than phenotypic evaluation (Crossa et al., 2016; Muleta et al., 2017; He and Li, 2020). This component aims to develop genomic selection for yield-related traits to accelerate genetic gain. Genomic selection approaches have been proven to be effective for complex or economically important traits such as yield, using an elite set of lines, including germplasm (Roy et al., 2021). From our perspective, there is little information available about Indian wheat germplasm, and a short table was prepared (**Table 3**) in the context of genes, transcripts, and QTL identified so far. In addition to the huge sustaining potential (**Table 1**), the Indian germplasm lines will be used to develop a prediction model using the existing genotyping and multi-location phenotyping data (**Figure 3**). The focus of this review would be to envisage candidate genotypes preserved in the national gene bank, which can be produced in abundance under varying climatic conditions. To the best of our knowledge, this would be the first time that genomics- and physiology-based hypothetical networks would be used to maximize the value of wheat germplasm in India. The collection of representative lines in this study will generate a public resource of elite germplasm lines with well-characterized phenotypic and genotypic information, along with seeds and their genetic constitutions. This resource would lead to determining the optimized configuration of wheat-breeding systems to support coming generations.

6 FUTURE PROSPECTS OF HARNESSING THE GENETIC POTENTIAL OF WHEAT GERMPLASM USING GENOMIC APPROACHES

The outcome of the Indian wheat germplasm genomics initiative would be a comprehensive pipeline connecting germplasm evaluation and genomic information, which could be used to accelerate the utilization of indigenous wheat germplasm in the national breeding programs for the improvement of biotic, abiotic, and quality traits (**Figure 3**). The following envisaged expected output could be: 1) a detailed insight into the extent and pattern of quality traits in the indigenous wheat collection, 2) wheat germplasm and genomic resource database containing phenotypic evaluation data and associated genomic information in the form of SNP markers for large-scale genotyping applications, 3) molecular tags such as markers, genes, and haplotypes associated with important agro-morphological, yield, and grain quality associated traits, 4) novel gene/markers conferring resistance to important wheat diseases (rusts, powdery mildew, and spot blotch, etc.) and tolerance to environmental stresses (heat, drought, and salinity), 5) elite germplasm/accession/genetic stocks based on extensive phenotyping and genomics-based analysis, 6) stable and cross-validated genomic prediction model to calculate the genomic-estimated breeding value for faster genetic gain in elite and pre-breeding lines for various traits (heat, drought, nitrogen use efficiency, rusts, spot blotch, and yield, etc.), 7) integration of physiological traits into the national wheat-breeding program to develop high carbon-capturing pre-breeding lines or candidate varieties, and 8) rematriation of old landraces of wheat to evaluate them at their native place to know their adaptive functionality. Wheat genetic resources include extant cultivars, obsolete cultivars, parental lines, advanced breeding material, mapping populations, and explored germplasm lines. Globally, there is a huge reserve of conserved wheat genetic resources, though much of it remains unexplored for trait-specific information. Genetic data on traits and their association with suitable markers will facilitate the use of wider variability in crop improvement. In this project, such an effort has been put forth to strengthen the Indian wheat-breeding program. The pre-breeding lines generated in the project will have enhanced climate resilience and combining both abiotic and biotic stress tolerance and maximized yield potential. This work would also set a precedent for further enrichment of the national wheat collection with wild and weedy relatives of wheat (wild Triticaceae such as species of *Aegilops*, *Elymus*, and *Eremopyrum*), distributed primarily in the western and north-western Himalayas, for use in future programs on climate resilience. It would also serve as a reference for identifying the required areas for exploration and collection of *Triticum* species based on the gaps identified in the gene bank holdings, especially for trait-specific and unique germplasm accessions.

AUTHOR CONTRIBUTIONS

Conceptualization: SdK, MP, SK, and KS; writing—review and editing: SdK, SJ, RM, VV, PK, TC, SaK, UK, SS, RS, SP, AnS, AmS, JK, MS, and SB; final editing and correction: SdK, TC, and SuK. All authors contributed to the manuscript and approved the submitted version.

ACKNOWLEDGMENTS

The authors are thankful to DBT, Govt. of India, for their funding to the project, "Wheat Germplasm Genomics for Trait Discovery" for evaluating wheat germplasm lines for various biotic, abiotic, quality, and nutritional traits.

REFERENCES

Abdelmageed, K., Chang, X., Wang, D. M., Wang, Y. J., and Yang, S. Y. (2019). Evolution of Varieties and Development of Production Technology in Egypt Wheat: a Review. *Integr. Agric.* 18, 483–495. doi:10.1016/S2095-3119(18)62053-2

Abhinandan, K., Skori, L., Stanic, M., Hickerson, N., Jamshed, M., and Samuel, M. A. (2018). Abiotic Stress Signaling in Wheat–An Inclusive Overview of Hormonal Interactions during Abiotic Stress Responses in Wheat. *Front. plant Sci.* 9, 734. doi:10.3389/fpls.2018.00734

Abou-Elwafa, S. F., and Shehzad, T. (2021). Genetic Diversity, GWAS and Prediction for Drought and Terminal Heat Stress Tolerance in Bread Wheat (*Triticum aestivum* L). *Genet. Resour. Crop Evol.* 68 (2), 711–728. doi:10.1007/s10722-020-01018-y

Adhikari, L., Raupp, J., Wu, S., Wilson, D., Evers, B., Koo, D. H., et al. (2022). Genetic Characterization and Curation of Diploid A-Genome Wheat Species. *Plant physiol.* 188 (4), 2101–2114. doi:10.1093/plphys/kiac006

Adhikari, S., Kumari, J., Jacob, S. R., Prasad, P., Gangwar, O. P., Lata, C., et al. (2022). Landraces-Potential Treasure for Sustainable Wheat Improvement. *Genet. Resour. Crop Evol.* 69, 499–523. doi:10.1007/s10722-021-01310-5

Ahirwar, R. N., Mishra, V. K., Chand, R., Budhlakoti, N., Mishra, D. C., Kumar, S., et al. (2018). Genome-wide Association Mapping of Spot Blotch Resistance in Wheat Association Mapping Initiative (WAMI) Panel of Spring Wheat (*Triticum aestivum* L). *Plos One* 13, e0208196. doi:10.1371/journal.pone.0208196

Alaux, M., Rogers, J., Letellier, T., Flores, R., Alfama, F., Pommier, C., et al. (2018). Linking the International Wheat Genome Sequencing Consortium Bread Wheat Reference Genome Sequence to Wheat Genetic and Phenomic Data. *Genome Biol.* 19 (1), 1–10. doi:10.1186/s13059-018-1491-4

Alexandratos, N., and Bruinsma, J. (2012). *World Agriculture towards 2030/2050. ESA Working Paper No. 12-03*. Rome: Food and Agriculture Organization of the United Nations.

Alford, L., Andrade, T. O., Georges, R., Burel, F., and van Baaren, J. (2014). Could Behavior and Not Physiological Thermal Tolerance Determine Winter Survival of Aphids in Cereal Fields? *Plos One* 9, e114982. doi:10.1371/journal.pone.0114982

Ali, M. W., and Borrill, P. (2020). Applying Genomic Resources to Accelerate Wheat Biofortification. *Heredity* 125 (6), 386–395. doi:10.1038/s41437-020-0326-8

Arora, R. K., and Koppar, M. N. (1979). *Collection of Local Wheat Germplasm from Madhya Pradesh*. New Delhi: National Bureau of Plant Genetic Resources NBPGR.

Ayana, G. T., Ali, S., Sidhu, J. S., Hernandez, J. L. G., Turnipseed, B., and Sehgal, S. K. (2018). Genome-Wide Association Study for Spot Blotch Resistance in Hard Winter Wheat. *Front. Plant Sci* 9, 926. doi:10.3389/fpls.2018.00926

Azeez, M. A., Adubi, A. O., and Durodola, F. A. (2018). "Landraces and Crop Genetic Improvement," in *Rediscovery of Landraces as a Resource for the Future* (IntechOpen). doi:10.5772/intechopen.75944

Badaruddin, M., Reynolds, M. P., and Ageeb, O. A. A. (1999). Wheat Management in Warm Environments: Effect of Organic and Inorganic Fertilizer, Irrigation Frequency, and Mulching. *Agron. J.* 91, 975–983. doi:10.2134/agronj1999.916975x

Balk, J., Connorton, J. M., Wan, Y., Lovegrove, A., Moore, K. L., Uauy, C., et al. (2019). Improving Wheat as a Source of Iron and Zinc for Global Nutrition. *Nutr. Bull.* 44 (1), 53–59. doi:10.1111/nbu.12361

Basandrai, D., Basandrai, A. K., Rana, S. K., Sharma, B. K., Singh, A., Singh, D., et al. (2016). Resistance to Powdery Mildew (Blumeria Graminis F. Sp. Tritici E. Marchal.) in Bread Wheat, Durum, Dicoccum, and Triticale Genotypes. *Indian J. Genet. Plant Breed.* 76 (2), 205–208. doi:10.5958/0975-6906.2016.00024.9

Bennett, F. G. A. (1984). Resistance to Powdery Mildew in Wheat; a Review of its Use in Agriculture and Breeding Programmes. *Plant Pathol.* 33, 279–300. doi:10.1111/j.1365-3059.1984.tb01324.x

Beres, B. L., Hatfield, J. L., Kirkegaard, J. A., Eigenbrode, S. D., Pan, W. L., Lollato, R. P., et al. (2020). Toward a Better Understanding of Genotype× Environment× Management Interactions—A Global Wheat Initiative Agronomic Research Strategy. *Front. Plant Sci.* 11, 828. doi:10.3389/fpls.2020.00828

Bhardwaj, S. C., Singh, G. P., Gangwar, O. P., Prasad, P., and Kumar, S. (2019). Status of Wheat Rust Research and Progress in Rust Management in Indian Context. *Agronomy* 9, 892. doi:10.3390/agronomy9120892

Borrill, P., Connorton, J., Balk, J., Miller, T., Sanders, D., and Uauy, C. (2014). Bio Fortification of Wheat Grain with Iron and Zinc: Integrating Novel Genomic Resources and Knowledge from Model Crops. *Front. plant Sci.* 5, 53. doi:10.3389/fpls.2014.00053

Bourke, P. M., Evers, J. B., Bijma, P., Van Apeldoorn, D., Smulders, M. J., Kuyper, T., et al. (2021). Breeding beyond Monoculture: Putting The'intercrop'into Crops. *Front. plant Sci.* 18;12, 734167. doi:10.3389/fpls.2021.734167

Boyer, J. S. (1982). Plant Productivity and Environment. *Science* 218, 443–448. doi:10.1126/science.218.4571.44310.1126/science.218.4571.443

Brar, G. S., Dhariwal, R., and Randhawa, H. S. (2018). Correction to: Resistance Evaluation of Differentials and Commercial Wheat Cultivars to Stripe Rust (*Puccinia Striiformis*) Infection in Hot Spot Regions of Canada. *Eur. J. Plant Pathol.* 152, 269. doi:10.1007/s10658-018-1452-5

Brar, G. S., AnitaBrûlé-Babel, L., Ruan, Y., Henriquez, M. A., Pozniak, C. J., Kutcher, H. R., et al. (2019). Genetic Factors Affecting Fusarium Head Blight Resistance Improvement from Introgression of Exotic Sumai 3 Alleles (Including Fhb1, Fhb2, and Fhb5) in Hard Red Spring Wheat. *BMC Plant Biol.* 19, 179. doi:10.1186/s12870-019-1782-2\

Budhlakoti, N., Kushwaha, A. K., Rai, A., Chaturvedi, K. K., Kumar, A., Pradhan, A. K., et al. (2022). Genomic Selection: A Tool for Accelerating the Efficiency of Molecular Breeding for Development of Climate Resilient Crops. *Front. Genet.* 13, 832153. doi:10.3389/fgene.2022.832153

Chao, S., Zhang, W., Dubcovsky, J., and Sorrells, M. (2007). Evaluation of Genetic Diversity and Genome-wide Linkage Disequilibrium Among US Wheat (*Triticum aestivum* L.) Germplasm Representing Different Market Classes. *Crop Sci.* 47 (3), 1018–1030. doi:10.2135/cropsci2006.06.0434

Chaves, M. M., Maroco, J. P., and Pereira, J. S. (2003). Understanding Plant Responses to Drought – from Genes to the Whole Plant. *Funct. Plant Biol.* 30, 239–264. doi:10.1071/fp02076

Chen, X. M. (2013). Review Article: High-Temperature Adult-Plant Resistance Key for Sustainable Control of Stripe Rust. *Am. J. Plant Sci.* 4, 608–627. doi:10.4236/ajps.2013.43080

Choudhary, A., Kaur, N., Sharma, A., and Kumar, A. (2021). Evaluation and Screening of Elite Wheat Germplasm for Salinity Stress at the Seedling Phase. *Physiol. Plant.* 173 (4), 2207–2215. doi:10.1111/ppl.13571

Chugh, V., Kaur, K., Singh, D., Kumar, V., Kaur, H., and Dhaliwal, H. S. (2015). Molecular Characterization of Diverse Wheat Germplasm for Puroindoline Proteins and Their Antimicrobial Activity. *Turkish J. Biol.* 39 (3), 359–369. doi:10.3906/biy-1405-30

Crossa, J., Jarquín, D., Franco, J., Pérez-Rodríguez, P., Burgueño, J., Saint-Pierre, C., et al. (2016). Genomic Prediction of Gene Bank Wheat Landraces. *G3 Genes, Genomes, Genet.* 6 (7), 1819–1834. doi:10.1534/g3.116.029637

Dai, A. (2013). Increasing Drought under Global Warming in Observations and Models. *Nat. Clim. Change* 3, 52–58. doi:10.1038/NCLIMATE1633

Damania, A. B. (1997). *Te Origin of Agriculture and Crop Domestication (ICARDA)*. Aleppo, Syrian Arab Republic 325.

Das, M. K., Rajaram, S., Mundt, C. C., Kronstad, W. E., and Singh, R. P. (1992). Inheritance of Slow Rusting Resistance in Wheat. *Crop Sci.* 32, 1452–2145. doi:10.2135/cropsci1992.0011183X003200060028x

Dev, S. M., and Sharma, A. N. (2010). *Food Security in India: Performance, Challenges, and Policies*. India: Oxfam.

Dhaliwal, G. S., Jindal, V., and Dhawan, A. K. (2010). Insect Pest Problems and Crop Losses: Changing Trends. *Indian J. Ecol.* 37, 1–7. doi:10.12968/denu.2010.37.6.364

Dharmateja, P., Kumar, M., Pandey, R., Mandal, P. K., Babu, P., Bainsla, N. K., et al. (2021). Deciphering the Change in Root System Architectural Traits under Limiting and Non-limiting Phosphorus in Indian Bread Wheat Germplasm. *Plos One* 16 (10), e0255840. doi:10.1371/journal.pone.0255840

Díaz De León, J., Escoppinichi, R., Zavala-Fonseca, R., Castellanos, T., Röder, M., and Mujeeb-Kazi, A. (2010). Phenotypic and Genotypic Characterization of Salt-Tolerant Wheat Genotypes. *Cereal Res. Commun.* 38 (1), 15–22. doi:10.1556/CRC.38.2010.1.2

Díez, M. J., De la Rosa, L., Martín, I., Guasch, L., Cartea, M. E., Mallor, C., et al. (2018). Plant Genebanks: Present Situation and Proposals for Their Improvement. The Case of the Spanish Network. *Front. plant Sci.* 9, 1794. doi:10.3389/fpls.2018.01794.eCollection2018

Djanaguiraman, M., Prasad, P. V. V., Kumari, J., Sehgal, S. K., Friebe, B., and Djalovic, I. (2019). Alien Chromosome Segment from Aegilops Speltoides and Dasypyrum Villosum Increases Drought Tolerance in Wheat via Profuse and Deep Root System. *BMC plant Biol.* 19 (1), 1–15. doi:10.1186/s12870-019-1833-8

Dubey, R., Pathak, H., Chakrabarti, B., Singh, S., Gupta, D. K., and Harit, R. C. (2020). Impact of Terminal Heat Stress on Wheat Yield in India and Options for Adaptation. *Agric. Syst.* 181, 102826. doi:10.1016/j.agsy.2020.102826

Dutta, M., Phogat, B. S., Kumar, S., Kumar, N., Kumari, J., Pandey, A. C., et al. (2015). "Development of Core Set of Wheat (Triticum spp.) Germplasm Conserved in the National Genebank in India," in *Advances in Wheat*

Genetics: From Genome to Field (Tokyo: Springer), 33–45. doi:10.1007/978-4-431-55675-6_4

Dvorjak, D. S. (2014). *Fusarium Head Blight Resistance and Agronomic Perform Performance in Soft Red Winter Wheat Populations, Theses, and Dissertations-Pplant and Soil Sciences*.

El-Hendawy, S. E., Hu, Y., Yakout, G. M., Awad, A. M., Hafiz, S. E., and Schmidhalter, U. (2005). Evaluating Salt Tolerance of Wheat Genotypes Using Multiple Parameters. *Eur. J. Agron.* 22 (3), 243–253. doi:10.1016/j.eja.2004.03.002

Fahad, S., Bajwa, A. A., Nazir, U., Anjum, S. A., Farooq, A., Zohaib, A., et al. (2017). Crop Production under Drought and Heat Stress: Plant Responses and Management Options. *Front. plant Sci.* 8, 1147. doi:10.3389/fpls.2017.01147. eCollection.2017

FAO (1983). *Commission on Plant Genetic Resources. Resolution 8/83 of the 22nd Session of the FAO Conference*. Rome: FAO Publication.

Friebe, B., and Heun, M. (1989). C-banding Pattern and Powdery Mildew Resistance of Triticum Ovatum and Four *T. aestivum-T.Ovatum* Addition Lines. *Theor. Appl. Genet.* 78, 417–424. doi:10.1007/BF00265306

Fuentes-Dávila, G., and Rajaram, S. (1994). Sources of Resistance to *Tilletia Indica* in Wheat (*Triticum aestivum*). *Crop Prot.* 13, 20–24. doi:10.1016/0261-2194(94)90131-7

Gami, R. A., Tank, C. J., Patel, S. S., Chauhan, R. M., and Patel, H. N. (2011). Combining Ability Analysis for Grain Yield and Quality Component Traits in Durum Wheat (Triticum Durum Desf). *Res. Crops.* 12, 502–504.

Gauchan, D., and Joshi, B. K. (2019). *Seeds without Borders Initiative for Enhanced Food and Nutrition Security in South Asia*.

Goel, S., Yadav, M., Singh, K., Jaat, R. S., and Singh, N. K. (2018). Exploring Diverse Wheat Germplasm for Novel Alleles in HMW-GS for Bread Quality Improvement. *J. food Sci. Technol.* 55 (8), 3257–3262. doi:10.1007/s13197-018-3259-y

Goyal, E., Amit, S. K., Singh, R. S., Mahato, A. K., Chand, S., Kanika, K., et al. (2016). Transcriptome Profiling of the Salt-Stress Response in *Triticum aestivum* Cv. Kharchia Local. *Sci. Rep.* 6, 1–14. doi:10.1038/srep27752

Grewal, S., and Goel, S. (2015). *Current Research Status and Future Challenges to Wheat Production in India*.

Grote, U., Fasse, A., Nguyen, T. T., and Erenstein, O. (2021). Food Security and the Dynamics of Wheat and Maize Value Chains in Africa and Asia. *Front. Sustain. Food Syst.* 4, 317. doi:10.3389/fsufs.2020.617009

Guo, J., Zhang, X. L., Hou, Y. L., Cai, J. J., Shen, X. R., Zhou, T., et al. (2015). High-density Mapping of the Major FHB Resistance Gene Fhb7 Derived from Thinopyrum Ponticum and its Pyramiding with Fhb1 by Marker-Assisted Selection. *Theor. Appl. Genet.* 128, 2301–2316. doi:10.1007/s00122-015-2586-x

Gupta, H. S., and Kant, L. (2012). Wheat Improvement in Northern Hills of India. *Agric. Res.* 1, 100–116. doi:10.1007/s40003-012-0020-z

Gupta, D., Mittal, R. K., Singh, M., and Kant, A. (2005). Evaluation of Triticale X Bread Wheat Derivatives for Agro-Morphological, Physiological and Biochemical Traits in Relation to Drought and Cold Stress. *Ann. Biol.* 21 (2), 143.

Gupta, A., Kant, L., Mahajan, V., Saha, S., and Gupta, H. S. (2009). Tank': Local Landrace of Wheat for Protection of Wheat Field against Monkey Raids–An Indigenous Technical Knowledge. *Curr. Sci.* 96, 467–468. Available at: https://www.jstor.org/stable/i24104487.

Gupta, P. K., Balyan, H. S., Gahlaut, V., and Kulwal, P. L. (2012). Phenotyping, Genetic Dissection, and Breeding for Drought and Heat Tolerance in Common Wheat: Status and Prospects. *Plant Breed. Rev.* 36, 85–168. doi:10.1002/9781118358566.ch2

Gupta, P. K., Chand, R., and Vasistha, N. K. (2018). Spot Blotch Disease of Wheat: the Current Status of Research on Genetics and Breeding. *Plant Pathol.* 67, 508–531. doi:10.1111/ppa.12781

Gupta, V., Mahajan, S., Razdan, V. K., Fatima, K., and Sharma, S. (2018). Characterization of Wheat Germplasm for Slow Rust Resistance against Puccinia Striiformis. *Archives Phytopathology Plant Prot.* 51 (3-4), 207–216. doi:10.1080/03235408.2018.1446293

Gustafson, P., Raskina, O., Ma, X., and Nevo, E. (2009). "Wheat Evolution, Domestication, and Improvement," in *Wheat: Science and Trade* (Danvers: Wiley), 5–30.

Hasanuzzaman, M., Nahar, K., and Hossain, M. A. (2019). *Wheat Production in Changing Environments*. New York: Springer.

Hatfield, J. L., and Beres, B. L. (2019). Yield Gaps in Wheat: Path to Enhancing Productivity. *Front. plant Sci.* 10, 1603. doi:10.3389/fpls.2019.01603

Hays, D. B., Do, J. H., Mason, R. E., Morgan, G., and Finlayson, S. A. (2007). Heat Stress Induced Ethylene Production in Developing Wheat Grains Induces Kernel Abortion and Increased Maturation in a Susceptible Cultivar. *Plant Sci.* 172, 1113–1123. doi:10.1016/j.plantsci.2007.03.004

He, T., and Li, C. (2020). Harness the Power of Genomic Selection and the Potential of Germplasm in Crop Breeding for Global Food Security in the Era with Rapid Climate Change. *Crop J.* 8 (5), 688–700. doi:10.1016/j.cj.2020.04.005

He, X., Juliana, P., Kabir, M. R., Roy, K. K., and Islam, R. (2021). Screening and Mapping for Head Blast Resistance in a Panel of CIMMYT and South Asian Bread Wheat Germplasm. *Front. Genet.* 12, 782. doi:10.3389/fgene.2021.679162

Hedden, P. (2003). Te Genes of the Green Revolution. *Trends Genet.* 19, 5–9. doi:10.1016/s0168-9525(02)00009-4

Hsam, S., Lapochkina, I., and Zeller, F. (2003). Chromosomal Location of Genes for Resistance to Powdery Mildew in Common Wheat (*Triticum aestivum* L. Em Thell). 8. Gene Pm32 in a Wheat-Aegilops Speltoides Translocation Line. *Euphytica* 133, 367–370. doi:10.1023/A:1025738513638

Hussain, B. (2015). Modernization in Plant Breeding Approaches for Improving Biotic Stress Resistance in Crop Plants. *Turkish J. Agric. For.* 39 (4), 515–530. doi:10.3906/tar-1406-176

ICAR-IIWBR (2021). *Annual Report of ICAR- Indian Institute of Wheat and Barley Research*. doi:10.13140/RG.2.2.23821.87528

IPCC. Climate Change (2013). *The Physical Science Basis*. Geneva, Switzerland: IPCC.

IWGSC (The International Wheat Genome Sequencing Consortium) (2018). Shifting the Limits in Wheat Research and Breeding Using a Fully Annotated Reference Genome. *Science* 361, 7191. doi:10.1126/science.aar7191

Jacob, S. R., Tyagi, V., Agrawal, A., Chakrabarty, S. K., and Tyagi, R. K. (2015). Indian Plant Germplasm on the Global Platter: an Analysis. *Plos One* 10 (5), e0126634. doi:10.1371/journal.pone.0126634

Jamil, A., Riaz, A. M., and Foolad, M. R. (2011). Gene Expression Profiling of Plants under Salt Stress. *Crit. Rev. Plant Sci.* 30, 435–458. doi:10.1080/07352689.2011.605739

Jamla, M., Gaikwad, A. B., Jacob, S. R., and Archak, S. (2017). Text-mining for Identifying Abiotic Stress Candidate Genes to Screen Bread Wheat (*Triticum aestivum*) Germplasm. *Indian J. Plant Genet. Resour.* 30 (2), 162–164. doi:10.5958/0976-1926.2017.00023.7

Jenner, C. F. (1994). Starch Synthesis in the Kernel of Wheat under High-Temperature Conditions. *Funct. Plant Biol.* 21 (6), 791–806. doi:10.1071/pp9940791

Johnson, R., and Law, C. N. (1973). Cytogenetic studies on the resistance of the wheat variety Berse´e to Puccinia striiformis. *Cereal Rusts Bull.* 1, 38–43.

Jorgensen, I. H. (1992). Discovery, Characterization, and Exploitation of Mlo Powdery Mildew Resistance in Barley. *Euphytica* 63, 141–152. doi:10.1007/BF00023919

Joshi, A. K., Mishra, B., Chatrath, R., Ferrara, G. O., and Singh, R. P. (2007a). Wheat Improvement in India: Present Status, Emerging Challenges, and Future Prospects. *Euphytica* 157, 431–446. doi:10.1007/s10681-007-9385-7

Joshi, A. K., Ortiz-Ferrara, G., and Crossa, J. (2007b). Combining Superior Agronomic Performance and Terminal Heat Tolerance with Resistance to Spot Blotch (Bipolaris Sorokiniana) of Wheat in the Warm Humid Gangetic Plains of South Asia. *F. Crop Res.* 103, 53–61. doi:10.1016/j.fcr.2007.04.010

Juliana, P., He, X., Marza, F., Islam, R., and Anwar, B. (2021). Genomic Selection for Wheat Blast in a Diversity Panel, Breeding Panel and Full-Sibs Panel. *Front. plant Sci.* 12, 745379. doi:10.3389/fpls.2021.745379

Khadka, K., Raizada, M. N., and Navabi, A. (2020). Recent Progress in Germplasm Evaluation and Gene Mapping to Enable Breeding of Drought-Tolerant Wheat. *Front. Plant Sci.* 11, 1149. doi:10.3389/fpls.2020.01149

Krishnappa, G., Srinivasan, K., Singh, A. M., Mishra, A., and Ahlawat, A. K. (2019). Exploratory Studies on Components of Variability for Seed Longevity and Quality Traits in Bread (*Triticum aestivum*) and Durum (Triticum Durum) Wheat. *Indian J. Agric. Sci.* 89 (3), 515–521.

Kulshrestha, V. P. (1985). History and Ethnobotany of Wheat in India. *J. Traditional Agric. Appl. Bot.* 32, 61–71. doi:10.3406/jatba.1985.3929

Kumar, S., Sehgal, S. K., Kumar, U., Prasad, P. V., and Joshi, A. K. (2012). Genomic Characterization of Drought Tolerance-Related Traits in Spring Wheat. *Euphytica* 186, 265–276. doi:10.1007/s10681-012-0675-3

Kumar, R., Arora, S., Singh, K., and Garg, M. (2015). Puroindoline Allelic Diversity in Indian Wheat Germplasm and Identification of New Allelic Variants. *Breed. Sci.* 65 (4), 319–326. doi:10.1270/jsbbs.65.319

Kumar, S., Röder, M. S., and Tripathi, S. B. (2015). Mendelization and Fine Mapping of a Bread Wheat Spot Blotch Disease Resistance QTL. *Mol. Breed.* 35, 218. doi:10.1007/s11032-015-0411-5

Kumar, S., Kumar, V., Kumari, P., Singh, A. K., and Singh, R. (2016a). DNA Finger-Printing and Genetic Diversity Studies in Wheat Genotypes Using SSR Markers. *J. Environ. Biol.* 37 (2), 319–326.

Kumar, S., Archak, S., Tyagi, R. K., Kumar, J., Vk, V., Jacob, S. R., et al. (2016b). Evaluation of 19,460 Wheat Accessions Conserved in the Indian National Genebank to Identify New Sources of Resistance to Rust and Spot Blotch Diseases. *PLoS One* 11 (12), e0167702. doi:10.1371/journal.pone.0167702

Kumar, S., Mishra, D. C., Bansal, R., Kumari, J., and Kumar, S. (2016c). Molecular Characterization of Indian Wheat Germplasm Lines for Stay Green & Other Heat Tolerance Genes Using Linked SSR Markers. *Biotech Today Int. J. Biol. Sci.* 6 (1), 90–94. doi:10.5958/2322-0996.2016.00014.4

Kumar, S., Kumari, J., Bansal, R., Kuri, B. R., Upadhyay, D., Srivastava, et al. (2018). Multi-Environmental Evaluation of Wheat Genotypes for Drought Tolerance. *Indian J. Genet.* 78, 26–35.

Kumar, S., Phogat, B. S., Vikas, V. K., Sharma, A. K., Saharan, M. S., Singh, A. K., et al. (2019a). Mining of Indian Wheat Germplasm Collection for Adult Plant Resistance to Leaf Rust. *Plos One* 14 (3), e0213468. doi:10.1371/journal.pone.0213468

Kumar, S., Singroha, G., Bhardwaj, S. C., Bala, R., Saharan, M. S., Gupta, V., et al. (2019b). Multienvironmental Evaluation of Wheat (*Triticum aestivum* L.) Germplasm Identifies Donors with Multiple Fungal Disease Resistance. *Genet. Resour. Crop Evol.* 66 (4), 797–808. doi:10.1007/s10722-019-00751-3

Kumar, S., Kumari, J., Bhusal, N., Pradhan, A. K., and Budhlakoti, N. (2020a). Genome-wide Association Study Reveals Genomic Regions Associated with Ten Agronomical Traits in Wheat under Late-Sown Conditions. *Front. plant Sci.* 11, 1420. doi:10.3389/fpls.2020.549743

Kumar, S., Singroha, G., Bhardwaj, S. C., Saharan, M. S., Gangwar, O. P., Mishra, C. N., et al. (2020b). Characterization of Exotic Germplasm Lines for Resistance to Wheat Rusts and Spot Blotch. *Indian Phytopathol.* 73 (2), 237–243. doi:10.1007/s42360-020-00232-z

Kumar, R., Masthigowda, M. H., Kaur, A., Bhusal, N., Pandey, A., Kumar, S., et al. (2020c). Identification and Characterization of Multiple Abiotic Stress Tolerance Genes in Wheat. *Mol. Biol. Rep.* 47 (11), 8629–8643. doi:10.1007/s11033-020-05906-5

Kumar, N., Rana, M., Kumar, B., Chand, S., and Shiv, A. (2021). Genomic Selection for Wheat Improvement. *Physiological, Mol. Genet. Perspect. Wheat Improv.*, 175–207. doi:10.1007/978-3-030-59577-7_9

Kumari, J., Phogat, B. S., Singh, T. P., Kumar, S., Kaur, V., Kumar, S., et al. (2015). Correlated Response of Component Traits towards Grain Yield in Cultivated Wheat Germplasm under Late Sown Condition. *J. Agro Ecol. Nat. Resour. Manag.* 2, 110–112.

Kumari, J., Kumar, S., Singh, N., Vaish, S. S., Das, S., Gupta, A., et al. (2018). Identification of New Donors for Spot Blotch Resistance in Cultivated Wheat Germplasm. *Cereal Res. Commun.* 46, 467–479. doi:10.1556/0806.46.2018.028

Kumari, J., Kumar, S., Singh, N., Vaish, S. S., Das, S., Gupta, A., et al. (2019). IC564121 (IC0564121; INGR18014), a Wheat (*Triticum aestivum*) Germplasm Highly Resistant to Spot Blotch. *Indian J. Plant Genet. Resour.* 32, 422–423.

Larkin, D. L., Lozada, D. N., and Mason, R. E. (2019). Genomic Selection—Considerations for Successful Implementation in Wheat Breeding Programs. *Agronomy* 9 (9), 479. doi:10.3390/agronomy9090479

León, D. D. J., Escoppinichi, R., Zavala-Fonseca, R., Castellanos, T., Röder, M., and Mujeeb-Kazi, A. (2010). Phenotypic and Genotypic Characterization of Salt-Tolerant Wheat Genotypes. *Cereal Res. Commun.* 38 (1), 15–22. Available at: https://www.jstor.org/stable/23790376.

Leonard, K. J., and Szabo, L. J. (2005). Stem Rust of Small Grains and Grasses Caused by Puccinia Graminis. *Mol. Plant Pathol.* 6, 99–111. doi:10.1111/j.1364-3703.2005.00273.x

Li, T., Zhang, D. D., Zhou, X. L., Bai, G. H., Li, L., and Gu, S. L. (2016). Fusarium Head Blight Resistance Loci in a Stratified Population of Wheat Landraces and Varieties. *Euphytica* 207, 551–561. doi:10.1007/s10681-015-1539-4

Lillemo, M., Van Ginkel, M., Trethowan, R. M., Herna´ndez, E., and Crossa, J. (2005). Differential Adaptation of CIMMYT Bread Wheat to Global High Temperatures. *Crop Sci.* 45, 2443–2453. doi:10.2135/cropsci2004.0663

Lillemo, M., Joshi, A. K., and Prasad, R. (2013). QTL for Spot Blotch Resistance in Bread Wheat Line Saar Co-locate to the Biotrophic Disease Resistance Loci Lr34 and Lr46. *Theor. Appl. Genet.* 126, 711–719. doi:10.1007/s00122-012-2012-6

Lobell, D. B., Schlenker, W., and Costa-Roberts, J. (2011). Climate Trends and Global Crop Production since 1980. *Science* 333, 616–620. doi:10.1126/science.1204531

Lu, C. B., Cheng, S. H., Wu, R. L., Hu, Y. H., Fan, J. P., Wang, C. S., et al. (2010). Breeding the Lines with Resistance to Fusarium Head Blight of Wheat Cultivar Yangmai 13 by Molecular Marker-Assisted Selection. *J. Triticeae Crops.* 30, 1058–1064.

Lu, P., Liang, Y., and Li, D. (2016). Fine Genetic Mapping of Spot Blotch Resistance Gene Sb3 in Wheat (*Triticum aestivum*). *Theor. Appl. Genet.* 129, 577. doi:10.1007/s00122-015-2649-z

Lv, C., Song, Y. X., Gao, L. F., Yao, Q., Zhou, R. H., Xu, R. G., et al. (2014). Integration of QTL Detection and Marker-Assisted Selection for Improving Resistance to Fusarium Head Blight and Important Agronomic Traits in Wheat. *Crop J.* 2, 70–78. doi:10.1016/j.cj.2013.10.004

Maccaferri, M., Harris, N. S., Twardziok, S. O., Pasam, R. K., and Gundlach, H. (2019). Durum Wheat Genome Highlights Past Domestication Signatures and Future Improvement Targets. *Nat. Genet.* 51 (5), 885–895. doi:10.1038/s41588-019-0381-3

Mahajan, S., and Tuteja, N. (2005). Cold, Salinity and Drought Stresses: An Overview. *Biochem. Biophys.* 444, 139–158. doi:10.1016/j.abb.2005.10.018

Maich, R., Ortega, D., Masgrau, A., and Manera, G. (2007). "Genetic Achievements under Rainfed Conditions," in *Wheat Production in Stressed Environments* (Dordrecht: Springer), 321–329. doi:10.1007/1-4020-5497-1_40

Malhi, G. S., Kau, r. M., and Kaushik, P. (2021). Impact of Climate Change on Agriculture and its Mitigation Strategies: A Review. *Sustainability* 13 (3), 1318. doi:10.3390/su13031318

Marone, D., Russo, M. A., Mores, A., Ficco, D., Laidò, G., Mastrangelo, A. M., et al. (2021). Importance of Landraces in Cereal Breeding for Stress Tolerance. *Plants* 10 (7), 1267. doi:10.3390/plants10071267

Mascher, M., Schreiber, M., Scholz, U., Graner, A., Reif, J. C., and Stein, N. (2019). Genebank Genomics Bridges the Gap between the Conservation of Crop Diversity and Plant Breeding. *Nat. Genet.* 51, 1076–1081. doi:10.1038/s41588-019-0443-6

McIntosh, R. A. (1998). "Breeding Wheat for Resistance to Biotic Stresses," in *Wheat Prospects for Global Improvement*. Editor H. J. Braun (Dordrecht, Netherlands: Kluwer Academic Press), 71–86.

McIntosh, R. A. (2020). *Catalogue of Gene Symbols for Wheat- 2020 Supplement Catalogue of Gene Symbols for Wheat*. Annual Wheat Newsletter 66.

McMullen, M., Bergstrom, G., de-Wolf, E., Dill-Macky, R., Hershamn, D., Shaner, G., et al. (2012). A Unified Effort to Fight an Enemy of Wheat and Barley, Fusarium Head Blight. *Plant Dis.* 96, 1712–1728. doi:10.1094/PDIS-03-12-0291-FE

Meena, B. L., Singh, A. K., Phogat, B. S., and Sharma, H. B. (2013). Effects of Nutrient Management and Planting Systems on Root Phenology and Grain Yield of Wheat (*Triticum aestivum*). *Indian J. Agric. Sci.* 83, 627–632.

Mehta, P. S., Muneem, K. C., and Negi, K. S. (2009). Traditional Wheat (*Triticum aestivum* L.) Genetic Resources for Subsistence in District Pauri of Uttarakhand. *Indian J. Plant Genet. Resour.* 22, 70–73. doi:10.5958/0976-1926.2019.00023.8

Mehta, P. S., Kumar, D., Ojha, S. N., and Negi, K. S. (2011). On-farm Dynamics of Wheat Genetic Resources in Nainital District of Uttarakhand. *Indian J. Plant Genet. Resour.* 24, 305–308.

Mehta, P. S., Arya, M., Singh, R., Rawat, R., and Bisht, I. S. (2019). Wheat (*Triticum aestivum* L.) Landrace Diversity in Traditional Production Landscapes of Uttarakhand Himalaya in North-Western India. *Indian J. Plant Genet. Resour.* 32, 181–191. doi:10.5958/0976-1926.2019.00023.8

Mir, R. R., Kumar, J., Balyan, H. S., and Gupta, P. K. (2012). A Study of Genetic Diversity Among Indian Bread Wheat (*Triticum aestivum* L.) Cultivars Released during Last 100 Years. *Genet. Resour. crop Evol.* 59 (5), 717–726. doi:10.1007/s10722-011-9713-6

Mitra, M. (1931). A New Bunt on Wheat in India. *Ann. Appl. Biol.* 18, 178–179. doi:10.1111/j.1744-7348.1931.tb02294.x

Mkhabela, S. S., Shimelis, H., Odindo, A. O., and Mashilo, J. (2019). Response of Selected Drought-Tolerant Wheat (*Triticum aestivum* L.) Genotypes for Agronomic Traits and Biochemical Markers under Drought-Stressed and Non-stressed Conditions. *Acta Agric. Scand. Sect. B—Soil Plant Sci.* 69 (8), 674–689. doi:10.1080/09064710.2019.1641213

Moghimi, A., Yang, C., Miller, M. E., Kianian, S. F., and Marchetto, P. M. (2018). A Novel Approach to Assess Salt Stress Tolerance in Wheat Using Hyperspectral Imaging. *Front. plant Sci.* 9, 1182. doi:10.3389/fpls.2018.01182

Mondal, S., Rutkoski, J. E., Velu, G., Singh, P. K., Crespo-Herrera, L. A., Guzman, C., et al. (2016). Harnessing Diversity in Wheat to Enhance Grain Yield, Climate Resilience, Disease, and Insect Pest Resistance and Nutrition through Conventional and Modern Breeding Approaches. *Front. plant Sci.* 7, 991. doi:10.3389/fpls.2016.00991.eCollection2016

Morgounov, A., Tufan, H. A., Sharma, R., Akin, B., Bagci, A., Braun, H.-J. Y., et al. (2012). Global Incidence of Wheat Rusts and Powdery Mildew during 1969–2010 and Durability of Resistance of Win-Ter Wheat Variety Bezostaya 1. *Eur. J. Plant Pathol.* 132, 323–340. doi:10.1007/s10658-011-9879-y

Müller, O., and Krawinkel, M. (2005). Malnutrition and Health in Developing Countries. *Cmaj* 173 (3), 279–286. doi:10.1503/cmaj.050342

Mukhtar, S., Khan, M. A., Paddar, B. A., Anjum, A., and Zaffar, G. (2015). Molecular Characterization of Wheat Germplasm for Stripe Rust Resistance Genes (*Yr5, Yr10, Yr15* & *Yr18*) and Identification of Candidate Lines for Stripe Rust Breeding in Kashmir. *Indian J. Biotech.* 14, 241–284.

Muleta, K. T., Bulli, P., Zhang, Z., Chen, X., and Pumphrey, M. (2017). Unlocking Diversity in Germplasm Collections via Genomic Selection: a Case Study Based on Quantitative Adult Plant Resistance to Stripe Rust in Spring Wheat. *Plant Genome* 10 (3), 15. doi:10.3835/plantgenome2016.12.0124

Mwale, M. M., Tang, X. L., and Chilembwe, E. (2017). Molecular Detection of Disease Resistance Genes to Powdery Mildew (Blumeria Graminis F. Sp. Tritici) in Wheat (*Triticum aestivum*) Cultivars. *Afr. J. Biotechnol.* 16, 22–31. doi:10.5897/AJB2016.15720

Nagarajan, S., and Joshi, L. M. (1985). "Epidemiology in the Indian Subcontinent," in *Alan P. Roelfs, William R. Bushnell, Diseases, Distribution, Epidemiology, and Control* (Academic Press), 371–402. doi:10.1016/B978-0-12-148402-6.50020-1

Nagarajan, S. (2005). Can India Produce Enough Wheat Even by 2020? *Curr. Sci.* 89, 1467–1471. Available at: https://www.jstor.org/stable/24110907.

Ortiz, R., Braun, H. J., Crossa, J., Crouch, J. H., Davenport, G., Dixon, J., et al. (2008). Wheat Genetic Resources Enhancement by the International Maize and Wheat Improvement Center (CIMMYT). *Genet. Resour. Crop Evol.* 55, 1095–1140. doi:10.1007/s10722-008-9372-4

Pal, D., Kumar, S., and Rana, J. C. (2007). Collection and Characterization of Wheat Germplasm from North-West Himalaya. *Indian J. Plant Genet. Resour.* 20 (2), 170–173. doi:10.5958/0976-1926.2014.00005.9

Pande, P. C., Awasthi, P., Bargali, K., and Bargali, S. S. (2016). Agro-biodiversity of Kumaun Himalaya, India: a Review. *Curr. Agric. Res. J.* 4, 16. doi:10.12944/carj.4.1.02

Pandey, G. K., Chinnusamy, V., and Lenka, S. K. (2021). Genes, Genomes, and Germplasm for Climate-Smart Agriculture-Part-I. *Curr. Genomics* 22, 2–3. doi:10.2174/1389202922012104121613 26

Panwar, N. S., Bhatt, K. C., Dhariwal, O. P., Pandey, A., and Jacob, S. (2014). Spatial Distribution of Trait-specific Diversity in Indian Wheat Collections. *Indian J. Plant Genet. Resour.* 27, 280–286. doi:10.5958/0976-1926.2014.00027.8

Phogat, B. S., Kumar, S., Kumari, J., Kumar, N., Pandey, A. C., Singh, T. P., et al. (2021). Characterization of Wheat Germplasm Conserved in the Indian National Genebank and Establishment of a Composite Core Collection. *Crop Sci.* 1, 604–620. doi:10.1002/csc2.20285

Phuke, R. M., He, X., Juliana, P., Bishnoi, S. K., and Singh, G. P. (2020). Association Mapping of Seedling Resistance to Tan Spot (Pyrenophora Tritici-Repentis Race 1) in CIMMYT and South Asian Wheat Germplasm. *Front. plant Sci.* 11, 1309. doi:10.3389/fpls.2020.01309

Pillay, D. P. K., and Kumar, T. M. (2018). Food Security in India: Evolution, Efforts, and Problems. *Strateg. Anal.* 42 (6), 595–611. doi:10.1080/09700161.2018.1560916

Pistorius, R. (1997). Scientists, Plants and Politics: A History of the Plant Genetic Resources Movement. *Bioversity Int.* 13.

Prada, D. (2009). Molecular Population Genetics and Agronomic Alleles in Seed Banks: Searching for a Needle in a Haystack? *J. Exp. Bot.* 60, 2541–2552. doi:10.1093/jxb/erp130

Pradhan, A. K., Kumar, S., Singh, A. K., Budhlakoti, N., Mishra, D. C., Chauhan, D., et al. (2020). Identification of QTLs/defense Genes Effective at Seedling Stage against Prevailing Races of Wheat Stripe Rust in India. *Front. Genet.* 11, 572975. doi:10.3389/fgene.2020.572975

Pradheep, K., Singh, M., Sultan, S. M., Singh, K., Parimalan, R., and Ahlawat, S. P. (2019). Diversity in Wild Relatives of Wheat: An Expedition Collection from Cold-Arid Indian Himalayas. *Genet. Resour. Crop Evol.* 66, 275–285. doi:10.1007/s10722-018-0706-6

Pretorius, Z. A., Singh, R. P., Wagoire, W. W., and Payne, T. S. (2000). Detection of Virulence to Wheat Stem Rust Resistance Gene Sr31 in Puccinia Graminis F. Sp. Tritici in Uganda. *Plant Dis.* 84, 203. doi:10.1094/pdis.2000.84.2.203b

Pretty, J., and Bharucha, Z. (2014). Sustainable Intensification in Agricultural Systems. *Ann. Bot.* 114, 1571–1596. doi:10.1093/aob/mcu205

Purugganan, M. D., and Jackson, S. A. (2021). Advancing Crop Genomics from Lab to Field. *Nat. Genet.* 53 (5), 595–601. doi:10.1038/s41588-021-00866-3

Raghu, B. R., Gangwar, O. P., Bhardwaj, S. C., and Jain, S. K. (2018). Seedling and Adult Plant Resistance to Stripe Rust in Uttarakhand Landraces of Wheat. *Appl. Biol. Res.* 20 (3), 291–301. doi:10.5958/0974-4517.2018.00039.3

Rajičić, V., Milivojević, J., Popović, V., Branković, S., Đurić, N., Perišić, V., et al. (2019). Winter Wheat Yield and Quality Depending on the Level of Nitrogen, Phosphorus and Potassium Fertilization. *Agric. For.* 65 (2), 79–88. doi:10.17707/agricultforest.65.2.06

Ram, S., Shoran, J., and Mishra, B. (2007). Nap Hal, an Indian Landrace of Wheat, Contains Unique Genes for Better Biscuit Making Quality. *J. plant Biochem. Biotechnol.* 16 (2), 83–86. doi:10.1007/bf03321979

Ramadas, S., Kumar, T. K., and Singh, G. P. (2019). "Wheat Production in India: Trends and Prospects," in *Recent Advances in Grain Crops Research* (IntechOpen).

Rangan, P., Furtado, A., and Henry, R. (2020). Transcriptome Profiling of Wheat Genotypes under Heat Stress during Grain-Filling. *J. Cereal Sci.* 91, 102895. doi:10.1016/j.jcs.2019.102895

Ranjan, R., Yadav, R., Jain, N., Sinha, N., Bainsla, N. K., Gaikwad, K. B., et al. (2021). Epistatic QTLs Play a Major Role in Nitrogen Use Efficiency and its Component Traits in Indian Spring Wheat. *Agriculture* 11 (11), 1149. doi:10.3390/agriculture11111149

Rasheed, A., Wen, W., Gao, F., Zhai, S., Jin, H., Liu, J., et al. (2016). Development and Validation of KASP Assays for Genes Underpinning Key Economic Traits in Bread Wheat. *Theor. Appl. Genet.* 129, 1843–1860. doi:10.1007/s00122-016-2743-x

Reynolds, M. P., Lewis, J. M., Ammar, K., Basnet, B. R., Crespo-Herrera, L., Crossa, J., et al. (2021). Harnessing Translational Research in Wheat for Climate Resilience. *J. Exp. Bot.* 72 (14), 5134–5157. doi:10.1093/jxb/erab256

Rosenqvist, E., Großkinsky, D. K., Ottosen, C. O., and Van de Zedde, R. (2019). The Phenotyping Dilemma—The Challenges of a Diversified Phenotyping Community. *Front. Plant Sci.* 10, 163. doi:10.3389/fpls.2019.00163

Routray, P., Basha, O., Garg, M., Singh, N. K., and Dhaliwal, H. S. (2007). Genetic Diversity of Landraces of Wheat (*Triticum aestivum* L.) from Hilly Areas of Uttaranchal, India. *Genet. Resour. Crop Evol.* 54 (6), 1315–1326. doi:10.1007/s10722-006-9112-6

Roy, C., Juliana, P., Kabir, M. R., Roy, K. K., Gahtyari, N. C., Marza, F., et al. (2021). New Genotypes and Genomic Regions for Resistance to Wheat Blast in South Asian Germplasm. *Plants* 10 (12), 2693. doi:10.3390/plants10122693

RustTracker.org (2019). *A Global Wheat Rust Monitoring System*.

Rush, C. M., Stein, J. M., Bowden, R. L., Riemenschneider, R., Boratynski, T., and Royer, M. H. (2005). Status of Karnal bunt of wheat in the United States 1996 to 2004. *Plant Disease* 89 (3), 212–223. doi:10.1094/PD-89-0212

Saari, E. E. (1998). "Leaf Blight Disease and Associated Soil Borne Fungal Pathogens of Wheat in South and South East Asia," in *Helminthosporium Blights of Wheat: Spot Blotch and Tan Spot*. Editors E. Duveiller, H. J. Dubin, J. Reeves, and A. McNab (Mexico, D.F.: CIMMYT), 37–51.

Sallam, A., Alqudah, A. M., Dawood, M. F., Baenziger, P. S., and Börner, A. (2019). Drought Stress Tolerance in Wheat and Barley: Advances in Physiology,

Breeding and Genetics Research. *Int. J. Mol. Sci.* 20 (13), 3137. doi:10.3390/ijms20133137

Sandhu, N., Sethi, M., Kumar, A., Dang, D., Singh, J., and Chhuneja, P. (2021). Biochemical and Genetic Approaches Improving Nitrogen Use Efficiency in Cereal Crops: A Review. *Front. Plant Sci.* 12, 757. doi:10.3389/fpls.2021.657629

Sareen, S., Bhusal, N., Kumar, M., Bhati, P. K., Munjal, R., Kumari, J., et al. (2020). Molecular Genetic Diversity Analysis for Heat Tolerance of Indigenous and Exotic Wheat Genotypes. *J. Plant Biochem. Biotechnol.* 29, 15–23. doi:10.1007/s13562-019-00501-7

Sattar, A., Sher, A., Ijaz, M., Ul-Allah, S., Rizwan, M. S., Hussain, M., et al. (2020). Terminal Drought and Heat Stress Alter Physiological and Biochemical Attributes in Flag Leaf of Bread Wheat. *Plos One* 15 (5), e0232974. doi:10.1371/journal.pone.0232974

Sawhney, R. N. (1995). "Genetics of Wheat-Rust Interaction," in *Plant Breeding Reviews 13*. Editor J. Janick (USA: John Wiley & Sons).

Sehgal, D., Vikram, P., Sansaloni, C. P., Ortiz, C., Pierre, C. S., Payne, T., et al. (2015). Exploring and Mobilizing the Gene Bank Biodiversity for Wheat Improvement. *Plos One* 10, e0132112. doi:10.1371/journal.pone.0132112

Sethi, V., de Wagt, A., Bhanot, A., Singh, K. D., Agarwal, P., Murira, Z., et al. (2020). Levels and Determinants of Malnutrition Among India's Urban Poor Women: An Analysis of Demographic Health Surveys 2006 and 2016. *Maternal child Nutr.* 16 (3), e12978. doi:10.1111/mcn.12978

Sharma, R. C., and Duveiller, E. (2006). Farmer Participatory Evaluation Confirms Higher Grain Yields in Spring Wheat Using a Selection Index for Spot Blotch Resistance, Maturity, and Kernel Weight. *Euphytica* 150, 307–317. doi:10.1007/s10681-006-0446-0

Sharma, A. K., Saharan, M. S., Sharma, I., Singh, M., Kundu, S., Phogat, B. S., et al. (2012). Identification of Diverse Sources of Multiple Disease Resistance in Wheat. *Indian J. Plant Genet. Resour.* 25, 238–245. doi:10.1094/PDIS-10-10-0760

Sharma, A. K., Saharan, M. S., Bhardwaj, S. C., Prashar, M., Chatrath, R., Tiwari, V., et al. (2015). Evaluation of Wheat (*Triticum aestivum*) Germplasm and Varieties against Stem Rust (Puccinia Graminis F. Sp. Tritici) Pathotype Ug99 and its Variants. *Indian Phytopathol.* 68 (2), 134–138.

Sharma, S., Schulthess, A. W., Bassi, F. M., Badaeva, E. D., Neumann, K., Graner, A., et al. (2021). Introducing Beneficial Alleles from Plant Genetic Resources into the Wheat Germplasm. *Biology* 10 (10), 982. doi:10.3390/biology10100982

Singh, J., and Kaur, D. J. (2015). Characterization of Cereal Cyst Nematode (Heterodera Avenae) Infecting Wheat in Punjab. *J. Wheat Res.* 7 (1), 48–54.

Singh, R. P., and Rajaram, S. (2002). *Breeding for Disease Resistance in Wheat*, 30. FAO plant production and protection series.

Singh, H., Grewal, T. S., Pannu, P. P. S., and Dhaliwal, H. S. (1999). Genetics of Resistance to Karnal Bunt Disease of Wheat. *Euphytica* 105, 125–131. doi:10.1023/a:1003425729370

Singh, N. B., Singh, Y. P., and Singh, V. P. N. (2005). Variation in Physiological Traits in Promising Wheat Varieties under Late Sown Condition. *Indian J. Plant Physiology* 10, 171.

Singh, S. K., Arun, B., and Joshi, A. K. (2007). Comparative Evaluation of Exotic and Adapted Germplasm of Spring Wheat for Floral Characteristics in the Indo-Gangetic Plains of Northern India. *Plant Breed.* 126, 559–564. doi:10.1111/j.1439-0523.2007.01396.x

Singh, R. P., Hodson, D. P., Huerta-Espino, J., Jin, Y., Njau, P., Wanyera, R., et al. (2008). Will Stem Rust Destroy the World's Wheat Crop? *Adv. Agron.* 98, 271–309. doi:10.1016/s0065-2113(08)00205-8

Singh, R. P., Hodson, D. P., Jin, Y., Lagudah, E. S., Ayliffe, M. A., Bhavani, S., et al. (2015). Emergence and Spread of New Races of Wheat Stem Rust Fungus: Continued Threat to Food Security and Prospects of Genetic Control. *Phytopathology* 105, 872–884. doi:10.1094/phyto-01-15-0030-fi

Singh, S. K., Venkatesh, K., Mishra, C. N., Tiwari, V., and Sharma, I. (2015). *Utilization of Local Wheat Biodiversity for Sustainable Development*. New Delhi: New India Publishing agency.

Singh, G., Singh, M. K., Tyagi, B. S., Singh, J. B., and Kumar, P. (2017). Germplasm Characterization and Selection Indices in Bread Wheat (*Triticum aestivum*) for Waterlogged Soils in India. *Indian J. Agric. Sci.* 87 (9), 1139–1148.

Singh, R. P., Srivastava, P., Sharma, A., and Bains, N. S. (2017). Bread Wheat Cultivar PBW 343 Carries Residual Additive Resistance against Virulent Stripe Rust Pathotype. *J. Crop Improv.* 31 (2), 183–191. doi:10.1080/15427528.2016.1263262

Singh, G., Kumar, P., Gupta, V., Tyagi, B. S., Singh, C., Sharma, A. K., et al. (2018). Multivariate Approach to Identify and Characterize Bread Wheat (*Triticum aestivum*) Germplasm for Waterlogging Tolerance in India. *Field Crops Res.* 221, 81–89. doi:10.1016/j.fcr.2018.02.019

Singh, S., Vikram, P., Sehgal, D., Burgueño, J., Sharma, A., Singh, S. K., et al. (2018). Harnessing Genetic Potential of Wheat Germplasm Banks through Impact-Oriented-Prebreeding for Future Food and Nutritional Security. *Sci. Rep.* 8, 1–11. doi:10.1038/s41598-018-30667-4

Singh, A. K., Chaurasia, S., Kumar, S., Singh, R., Kumari, J., Yadav, et al. (2018). Identification, Analysis and Development of Salt Responsive Candidate Gene Based SSR Markers in Wheat. *BMC plant Biol.* 18 (1), 1–15. doi:10.1186/s12870-018-1476-1

Singh, N., Wu, S., Raupp, W. J., Sehgal, S., Arora, S., Tiwari, V., et al. (2019). Efficient Curation of Genebanks Using Next-Generation Sequencing Reveals Substantial Duplication of Germplasm Accessions. *Sci. Rep.* 9 (1), 1–10. doi:10.1038/s41598-018-37269-0

Singh, K., Gupta, K., Tyagi, V., and Raj Kumar, S. (2020). Plant Genetic Resources in India: Management and Utilization. *Vavilov J. Genet. Breed.* 24, 306. doi:10.18699/vj20.622

Singh, K. (2018). *Role of ICAR-NBPGR in PGR Management*. New Delhi: ICAR-National Bureau of Plant Genetic Resources.

Smiley, R. W., Dababat, A. A., Iqbal, S., Jones, M. G., Maafi, Z. T., Peng, D., et al. (2017). Cereal Cyst Nematodes: A Complex and Destructive Group of Heterodera Species. *Plant Dis.* 101 (10), 1692–1720. doi:10.1094/pdis-03-17-0355-fe

Sood, T., Basandrai, D., Basandrai, A. K., Sohu, V. S., Rana, V., Mehta, A., et al. (2020). Stable Sources of Resistance to Yellow Rust and Powdery Mildew in Indian and Exotic Wheat Germplasm. *J. Cereal Res.* 12 (1), 23–28. doi:10.25174/2582-2675/2020/100835

Steiner, B., Buerstmayr, M., Michel, S., Schweiger, W., Lemmens, M., and Buerstmayr, H. (2017). Breeding Strategies and Advances in Line Selection for Fusarium Head Blight Resistance in Wheat. *Trop. Plant Pathol.* 42, 165–174. doi:10.1007/s40858-017-0127-7

Sukumaran, S., Reynolds, M. P., and Sansaloni, C. (2018). Genome-Wide Association Analyses Identify QTL Hotspots for Yield and Component Traits in Durum Wheat Grown under Yield Potential, Drought, and Heat Stress Environments. *Front. Plant Sci.* 9, 81. doi:10.3389/fpls.2018.00081

Tahmasebi, S., Heidari, B., Pakniyat, H., and McIntyre, C. L. (2016). Mapping QTLs Associated with Agronomic and Physiological Traits under Terminal Drought and Heat Stress Conditions in Wheat (*Triticum aestivum* L). *Genome* 60 (1), 26–45. doi:10.1139/gen-2016-0017

Tripathi, K., Chauhan, S. K., Gore, P. G., Mehta, P. S., Bisht, I. S., and Bhalla, S. (2017). Evaluation of Wheat Landraces of North-Western Himalaya against Rice Weevil, Sitophilus Oryzae L. Vis-À-Vis Physical Seed Parameters. *Plant Genet. Resour.* 15 (4), 321–326. doi:10.1017/s1479262115000672

Tripathi, K., Gore, P. G., Kumar, S., Bhalla, S., and Bisht, I. S. (2018). Population Structure and Genetic Diversity of Wheat Landraces from Northwestern Indian Himalaya. *Indian J. Plant Genet. Resour.* 31, 169–177. doi:10.5958/0976-1926.2018.00020.7

Tyagi, V., Kak, K. A., Pedapati, A., and Brahmi, P. (2015). Utilization of Exotic Plant Genetic Resources in Wheat Registered Germplasm. *Vegetos* 29 (2), 11–18. doi:10.5958/2229-4473.2016.00014.8

Tyagi, R. K. (2016). Conservation of PGR for Effective Utilization: New Initiatives at NBPGR. *Indian J. Plant Genet. Resour.* 29, 272–277. doi:10.5958/0976-1926.2016.00044.9

Vaibhav, K., Singh, G. P., Singh, P. K., Harikrishna, R., and Gogoi, R. (2017). Assessment of Slow Rusting Resistance Components to Stripe Rust Pathogen in Some Exotic Wheat Germplasm. *Indian Phytopathol.* 70 (1), 52–57. doi:10.1016/j.tplants.2018.08.012

Vasistha, N. K., Balasubramaniam, A., and Mishra, V. K. (2017). Molecular Introgression of Leaf Rust Resistance Gene Lr34 Validates Enhanced Effect on Resistance to Spot Blotch in Spring Wheat. *Euphytica* 213, 262. doi:10.1007/s10681-017-2051-9

Vernooy, R. (2019). *Strengthening National Capacities to Implement the International Treaty on Plant Genetic Resources for Food and Agriculture: Project Highlights*.

Vikas, V. K., Kumar, S., Archak, S., Tyagi, R. K., Kumar, J., Jacob, S., et al. (2020). Screening of 19,460 Genotypes of Wheat Species for Resistance to Powdery Mildew and Identification of Potential Candidates Using Focused Identification of Germplasm Strategy (FIGS). *Crop Science* 60 (6), 2857–2866. doi:10.1002/csc2.20196

Vikas, V. K., Sivasamy, M., Kumar, J., Jayaprakash, P., Kumar, S., Parimalan, R., et al. (2014). Stem and Leaf Rust Resistance in Wild Relatives of Wheat with D Genome (Aegilops spp). *Genet. Resour. crop Evol.* 61, 861–874. doi:10.1007/s10722-014-0085-6

Villareal, R. L., Toro, E. del., Mujeeb-Kazi, A., and Rajaram, S. (1995). The 1BL/1RS Chromosome Translocation Effect on Yield Characteristics in a *Triticum aestivum* L. Cross. *Plant Breed.* 114, 497–500. doi:10.1111/j.1439-0523.1995.tb00843.x

Vishwakarma, M. K., Mishra, V. K., Yadav, P. S., Saxesena, R. R., and Arun, B. (2015). Molecular Validation of Quality Trait Linked Markers for Exploring the Wheat Germplasm of Eastern Gangetic Plains of India. *Curr. Trends Biotechnol. Pharm.* 9 (1).

Wang, J., Vanga, S. K., Saxena, R., Orsat, V., and Raghavan, V. (2018). Effect of Climate Change on the Yield of Cereal Crops: a Review. *Climate* 6 (2), 41. doi:10.3390/cli6020041

Wang, H., Sun, S., Ge, W., Zhao, L., Hou, B., Wang, K., et al. (2020). Horizontal Gene Transfer of Fhb7 from Fungus Underlies Fusarium Head Blight Resistance in Wheat. *Science* 368 (6493), eaba5435. doi:10.1126/science.aba5435

Xiong, W., Reynolds, M. P., Crossa, J., Schulthess, U., Sonder, K., Montes, C., et al. (2021). Increased Ranking Change in Wheat Breeding under Climate Change. *Nat. Plants* 7 (9), 1207–1212. doi:10.1038/s41477-021-00988-w

Zeibig, F., Kilian, B., and Frei, M. (2021). *The Grain Quality of Wheat Wild Relatives in the Evolutionary Context.* Theoretical and Applied Genetics, 1–20.

Zeng, Z. B., Teulat, B., Merah, O., Sirault, X., Borries, C., Waugh, R., et al. (2014). Detection and Validation of Novel QTL for Shoot and Root Traits in Barley (Hordeum Vulgare L). *J. Exp. Bot.* 9, 171–180. doi:10.1007/s11032-014-0122-3

Zörb, C., Ludewig, U., and Hawkesford, M. J. (2018). Perspective on Wheat Yield and Quality with Reduced Nitrogen Supply. *Trends plant Sci.* 23 (11), 1029–1037.

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.

Copyright © 2022 Kumar, Jacob, Mir, Vikas, Kulwal, Chandra, Kaur, Kumar, Kumar, Sharma, Singh, Prasad, Singh, Singh, Kumari, Saharan, Bhardwaj, Prasad, Kalia and Singh. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 04 July 2022
doi: 10.3389/fgene.2022.912251

Genome-Wide Identification and Characterization of *Receptor-Like Protein Kinase 1 (RPK1)* Gene Family in *Triticum aestivum* Under Drought Stress

Amna Abdul Rahim[1,2], *Muhammad Uzair*[2], *Nazia Rehman*[2], *Obaid Ur Rehman*[2], *Nageen Zahra*[2] *and Muhammad Ramzan Khan*[1,2]*

[1]National Centre for Bioinformatics (NCB), Quaid-e-Azam University, Islamabad, Pakistan, [2]National Institute for Genomics and Advanced Biotechnology (NIGAB), National Agricultural Research Centre (NARC), Islamabad, Pakistan

OPEN ACCESS

Edited by:
Vijay Gahlaut,
Institute of Himalayan Bioresource Technology (CSIR), India

Reviewed by:
Shumayla,
Panjab University, India
Gautam Saripalli,
University of Maryland, United States

***Correspondence:**
Muhammad Ramzan Khan
mrkhan@parc.gov.pk

Specialty section:
This article was submitted to Plant Genomics, a section of the journal Frontiers in Genetics

Received: *04 April 2022*
Accepted: *24 May 2022*
Published: *04 July 2022*

Citation:
Rahim AA, Uzair M, Rehman N, Rehman OU, Zahra N and Khan MR (2022) Genome-Wide Identification and Characterization of Receptor-Like Protein Kinase 1 (RPK1) Gene Family in Triticum aestivum Under Drought Stress. Front. Genet. 13:912251. doi: 10.3389/fgene.2022.912251

Receptor-like protein kinase 1 (RPK1) genes play crucial roles in plant growth and development processes, root architecture, and abiotic stress regulation. A comprehensive study of the *RPK1* gene family has not been reported in bread wheat (*Triticum aestivum*). Here, we reported the genome-wide identification, characterization, and expression patterns of the *RPK1* gene family in wheat. Results confirmed 15 *TaRPK1* genes, classified mainly into three sub-clades based on a phylogenetic tree. The *TaRPK1* genes were mapped on chromosomes 1–3 in the respective A, B, and D genomes. Gene structure, motif conservation, collinearity prediction, and synteny analysis were carried out systematically. A Gene ontology study revealed that *TaRPK1* genes play a vital role during molecular and biological processes. We also identified 18 putative miRNAs targeting *TaRPK1* genes, suggesting their roles in growth, development, and stress responses. *Cis*-Regulatory elements interpreted the presence of light-related elements, hormone responsiveness, and abiotic stress-related motifs in the promoter regions. The SWISS_MODEL predicted the successful models of TaRPK1 proteins with at least 30% identity to the template, a widely accepted threshold for successful modeling. *In silico* expression analysis in different tissues and stages suggested that *TaRPK1* genes exhibited the highest expression in root tissues. Moreover, qRT-PCR further validated the higher expression of *TaRPK1* genes in roots of drought-tolerant varieties compared to the drought-susceptible variety. Collectively, the present study renders valuable information on the functioning of *TaRPK1* genes in wheat that will be useful in further functional validation of these genes in future studies.

Keywords: receptor-like protein kinase 1 (RPK1), abiotic stress, genome-wide studies, phylogenetic relationship, expression patterns, *Triticum aestivum*

INTRODUCTION

Globally, wheat is a staple food and a source of nutrition. In the last 2 decades, the production of wheat increased by up to 1% annually (Manès et al., 2012), but this increase is not enough to meet the demand of the population, which will increased from 7.8 billion to 9.7 billion in 2050 (Roser and Ortiz-Ospina, 2013). The climatic changes including abiotic and biotic stresses are the main causes which extremely effect the quality and yield of the crops. To face these challenges, it is vital to explore the crop genotypes that can stand up to all of these hurdles. Plants are immobile in nature; they don't move here and there in search of food, but their roots do. The root is the major organ that has a crucial role in the adaptation of the plant to its unfavorable environment. Root systems captivate the water and nutrients essential for the growth and maintenance of plant (Alahmad et al., 2019; Grzesiak et al., 2019). Hence, improved root system overcomes the challenges of the harsh environment and might enhance crop production (Djanaguiraman et al., 2019; Danakumara et al., 2021; Rasool et al., 2021).

Drought is one of the major abiotic stresses caused due to scarce rainfall that affects productivity. An increase in drought in the coming 30 years will have adverse effects on crop yield with 6–12 bushels/acre (Zargar et al., 2017). Creation of drought tolerance is a very complicated because many genes such as *TaER1, 2,* and *3, TaZFP34, TaWRKY1, 10, 33, 44*, and *93*, *TaDR O 1*, and *TaRAP2.1* directly or indirectly involved. In animals, *receptor protein kinases (RPKs)* are the genes which play a significant role in the stimulation of hormones and other growth factors (Fantl et al., 1993). In plants, similar to animals' *RPKs,* there is a *receptor-like protein kinase (RLK)* gene family. The *RLK* family is a huge family of genes found in many plants. The typical *RLK* structure comprises an extracellular domain at the N-terminal, a membrane helix, and an intracellular conserved kinase domain (KD) at the C-terminal. The extracellular domains of the *RLK* family are highly diverged, which results in the differentiation of *RLKs* into 17 distinct subfamilies, including the *receptor-like kinases* (Mishra et al., 2021). The *LRRKs* (*leucine-rich repeat kinase*) represents biggest subfamily of *RLK* with 531 *TaLRRK* genes in wheat (Sharma et al., 2016), comprising of ECD (extracellular domain) to receive signals, TM (transmembrane) region to bound it to cell membrane and cytoplasmic kinase domain for phosphorylation of substrate (Gou et al., 2010; Dievart et al., 2020). The *LRRKs* has numerous roles in plants as it is involved in initiating innate defense at front-line against microbial pathogens (Nejat and Mantri, 2017), morphogenesis, organogenesis, hormone signaling, abiotic, and biotic stress regulation in plants (Diévart and Clark, 2003; Li and Tax, 2013; Dufayard et al., 2017). Later on the role of *LRR-RLKs* in pathogen sensing and activation of downstream defense response has been reviewed deeply (Nejat and Mantri, 2017). Due to the indispensable roles of *LRR-RLKs* in plants, they have been classified into two main classes (Diévart and Clark, 2003). First, the *LRR-RLK* is crucial for morphogenesis, organogenesis, hormone signaling, signifying development, and growth regulation. Secondly, numerous *LRR-RLK* members respond to biotic and abiotic stresses like Fusarium wilt, drought, salt, and cold, and hence are associated with defense (Afzal et al., 2008; Cao et al., 2020). Some of the *LRR-RLKs* have dual roles that might be because of the cross-talk among development and defense cascades or due to the binding of several ligands to a receptor (Afzal et al., 2008).

The *RPK1* gene is a calcium independent Serine-Threonine (Ser-Thr) kinase that belongs to the subfamily of *leucine-rich receptor kinases* (*LRR kinases*) and family of *Receptor-Like Kinases* (*RLK*) (Zou et al., 2014). *RPK1* is one of the short subfamilies with few genes that regulates abiotic stresses and root system architecture. The *RPK1* comprises of extracellular six LRR motifs, a transmembrane domain, extracellular ligand-binding domain, and single cytoplasmic kinase conserved domain in rice (Hong et al., 1997; Cheng et al., 2009; Motte et al., 2014). Studies in rice have shown that *RPK1* is involved in root system architecture (RSA) via regulating negatively polar auxin transport (PAT) and accumulation of auxin in roots (Zou et al., 2014). In other studies of rice, it was also reported that auxin defective mutants showed stunted growth and shorter roots (Uzair et al., 2021). Down-regulation of *RPK1* endorsed the growth and enhanced the height of the plant and number of tillers, whereas up-regulation resulted in immature lateral roots, adventitious roots, and a decreased apical meristem of roots (Zou et al., 2014). In *Arabidopsis*, the inhibition of *AtRPK1* displayed greater salt tolerance than normal plants, while overexpressed plants exhibited lesser salt tolerance degrees (Shi et al., 2014). The levels of *AtRPK1* were enhanced ominously under less water, abscisic acid (ABA), high salt and lower temperature (Hong et al., 1997). In *Arabidopsis thaliana*, inhibition of *RPK1* delayed ABA-induced senescence significantly (Lee et al., 2011). *AtRPK1* is also prerequisite for cotyledon primordial initiation of cotyledons during embryogenesis in *Arabidopsis thaliana* (Nodine et al., 2007; Nodine and Tax, 2008). *AtRPK1* positively regulates *CaM1* gene expression, which in turn regulates ROS (reactive oxygen species) production, leaf senescence, and ABA response (Dai et al., 2018).

Triticum aestivum L., commonly known as "bread wheat", is a cereal and staple food grain all over the globe. Being a most consumed cereal crop, it was grown on a large scale of 240 million hectares in 2016 (Milner et al., 2018). However, due to water scarcity, nutrient deficiency, and abiotic stresses, wheat yield is curtailed (Mondal et al., 2015; Abbas et al., 2022). Wheat is a drought sensitive crop. Therefore, in order to meet the global demand, that is 50% of the grain in 20 years approximately, the varieties of wheat with effective utilization of minerals and water are requisite (Odegard and Van der Voet, 2014). Since roots are the main structures for the minerals and water uptake and decipher stress stimuli from soil (Fang et al., 2017). Hence, identification of stress-tolerant genes within the root system could be propitious.

Since the genome of *T. aestivum* has been sequenced, it is feasible to carry out a genome-wide analysis of different genes. In this study, 15 *TaRPK1* genes were analyzed for their structure, chromosomal location within the genome, phylogenetic relationships, conserved motifs, synteny, and *cis*-regulatory elements. Additionally, the patterns of expression of all

15 *TaRPK1* members were also studied *in silico*. RT-PCR expression analysis of *TaRPK1* members was also performed in Pakistan-13, Galaxy (drought tolerant), and Shafaq (drought susceptible) wheat varieties under normal and drought conditions. The current study enlightens the role of *TaRPK1* genes in plant developmental processes under drought conditions and provides a solid foundation for the functional characterization of the wheat *RPK1* gene family.

MATERIALS AND METHODS

Identification of *RPK* Gene Family Members in *T. aestivum*

The sequence IDs of *Arabidopsis* and rice *RPK1* genes were acquired from the available literature (Shi et al., 2014; Zou et al., 2014). These sequences were retrieved from Ensembl plants and NCBI, which were then used as queries for the Basic-Local Alignment Search tool (BlastP and BlastN) against IWGSC (INSDC Assembly GCA_900519105.1 July 2018 database version 106.4), NCBI (https://www.ncbi.nlm.nih.gov/) and Ensembl plants (plantshttps://plants.ensembl.org/index.html) for *T. aestivum*. For all of the candidate *RPK1* genes, the kinase domain presence was substantiated with Pfam (http://pfam.sanger.ac.uk), and by SMART (http://smart.embl-heidelberg.de/) (Letunic and Bork, 2018) databases. The sequences in which the kinase domain was absent were removed (**Supplementary Table S1**). *In silico* based putative protein information of *RPK1* genes (physio-chemical) was analyzed through the Protparam (https://web.expasy.org/protparam/) tool. The subcellular localization of RPK1 proteins was predicted via Plant-mSubP and pLoc-mPlant (http://bioinfo.usu.edu/Plant-mSubP/; http://www.jci-bioinfo.cn/pLoc-mPlant/) (Cheng et al., 2017; Sahu et al., 2020).

Chromosomal Location of *TaRPK1* Genes

The chromosomal locations of all candidate *RPK1* genes in *T. aestivum* were acquired from Ensembl (http://plants.ensembl.org/Triticum_aestivum/Info/Index). The gene map of *TaRPK1* genes was drawn with the help of MapChart and confirmed through TBtools.

Phylogenetic Analysis of RPK1 Proteins

To retrieve the RPK1 protein sequences, the amino acid sequences of 15 TaRPK1 members were used as queries to blast (BLASTP) against the *Triticum turgidium*, *Triticum dicoccoides*, *Titicum urartu*, *Triticum speltoides*, *Aegilops tauschii*, *Hordeum vulgare*, *Arabidopsis thaliana,* and different species of *Oryza* (*rufipigon*, *japonica*, *indica,* and *glaberrima*). The sequences with > 60% identities were retrieved from Ensembl (http://plants.ensembl.org). The phylogenetic trees were made by means of MEGA-X software with NJ (neighbor-joining method) (Kumar et al., 1994). The parameter Poisson model and pairwise deletion were used with replicates of 1,000 bootstraps for assessment of node significance.

Prediction of Gene Structure and Conserved Motifs in TaRPK1 Proteins

The number of exons and introns was predicted by the gene structure display server (GSDS, http://gsds.cbi.pku.edu.cn/) and the genomic sequences and coding sequences were aligned using ClustalW. Conserved motifs in RPK1 proteins of *T. aestivum* were analyzed using MEME, a multiple-EM for motif elicitation program (http://meme-suite.org/tools/meme) (Bailey et al., 2009). The execution of MEME search was done with default parameters apart from motif maximum number, which was set to 10, and optimum motif width of ≥6 and ≤200 was selected.

Gene Ontology Enrichment Analysis

The analysis of *TaRPK1* gene ontology was performed by TaRPK1 protein sequences via the online gProfiler tool (https://biit.cs.ut.ee/gprofiler/gost) with default parameters (Raudvere et al., 2019).

miRNA Prediction in Wheat *RPK1* Family Genes

The miRNA prediction was performed as mentioned formerly (Yan et al., 2019). The *TaRPK1* sequences were submitted for potential miRNA prediction through a search against the available wheat miRNA reference by means of the psRNATarget Server (https://www.zhaolab.org/psRNATarget/), using default settings (Dai and Zhao, 2011). The visualization of the interaction network of the predicted miRNA with their corresponding *TaRPK1* target genes was done by Cytoscape software (https://cytoscape.org/) with default settings (Shannon et al., 2003).

Interpretation of Putative Regulatory *Cis*-Acting Elements

The sequence size of 2 kb in the upstream region were dug out from all *TaRPK1* genes of *T. aestivum* that acted as promoters for the regulatory *cis* acting elements prediction through the PlantCare (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) database (Lescot et al., 2002).

Collinearity Prediction and Synteny Analysis

The GFF3 files and proteomes of *Triticum aestivum* and its ancestors, including *Aegilopes tauschii, Triticum spelta*, *Triticum turgidum,* and *Triticum dicoccoides,* were used from the Ensembl Plants database for collinearity prediction via the MCScanX algorithm (Wang et al., 2012). Synteny scrutiny of *RPK1* family members was performed via Tbtools (Chen et al., 2018).

Three-Dimensional Protein Structure Prediction

The TaRPK1 protein structures were modeled via amino acid sequence using the SWISS-MODEL database (https://www.swissmodel.expasy.org) (Biasini et al., 2014), and for

TABLE 1 | *In silico* prediction of identified *RPK1* genes in wheat and sequence characteristics.

Sr. no.	New name	Gene ID	Chr No.	Chr location	Orientation	CDS (bp)	No. of exons	Coding exons	No. of introns
1	TaRPK1.1 (TaRPK1)	TraesCS1A02G304200	1A	497,503,763–497,509,507	R	2,895	19	19	18
2	TaRPK1.2 (TaRPK2)	TraesCS1B02G314700	1B	539,546,762–539,552,423	R	2,895	19	19	18
3	TaRPK1.3 (TaRPK3)	TraesCS1D02G303700	1D	401,666,525–401,672,077	R	2,895	19	19	18
4	TaRPK1.4 (TaRPK4)	TraesCS3A02G340100	3A	587,403,291–587,408,585	F	2,676	17	17	16
5	TaRPK1.5 (TaRPK5)	TraesCS3B02G371700	3B	584,546,469–584,551,744	F	2,676	17	17	16
6	TaRPK1.6 (TaRPK6)	TraesCS3D02G333600	3D	445,633,883–445,639,177	F	2,676	17	17	16
7	TaRPK1.10 (TaRPK10)	TraesCS2A02G176500	2A	136,053,228–136,056,887	R	3,372	2	1	0
8	TaRPK1.11 (TaRPK11)	TraesCS2B02G202900	2B	182,708,242–182,711,907	R	3,372	2	1	0
9	TaRPK1.12 (TaRPK12)	TraesCS2D02G183900	2D	129,186,794–129,190,494	R	2,949	2	2	1
10	TaRPK1.13 (TaRPK13)	TraesCS2A02G260600	2A	410,851,518–410,855,096	F	2,187	2	1	0
11	TaRPK1.14 (TaRPK14)	TraesCS2B02G281400	2B	388,595,342–388,597,540	R	1830	2	2	1
12	TaRPK1.15 (TaRPK15)	TraesCS2D02G263100	2D	320,280,150–320,283,723	R	2,199	2	1	0
13	TaRPK1.7 (TaRPK7)	TraesCS3A02G340000	3A	587,396,690–587,401,627	F	2,772	18	18	17
14	TaRPK1.8 (TaRPK8)	TraesCS3B02G371600	3B	584,539,043–584,544,883	F	2,874	19	19	18
15	TaRPK1.9 (TaRPK9)	TraesCS3D02G333500	3D	445,627,122–445,632,371	F	2,775	18	18	17

Chr, chromosome; F, forward strand; R, reverse strand; CDS, coding sequence; bp, base pairs.

TABLE 2 | *In silico*-based putative protein information of *RPK1* genes identified in *T. aestivum.*

Sr. no.	New name	Sequence ID	PL (Aa)	Domain loc	Mol. wt. (Kda)	pI	II	AI	GRAVY	SCL
1	TaRPK1.1 (TaRPK1)	TraesCS1A02G304200	964	632–901	104	8	28.37	91.76	-0.032	Cell membrane
2	TaRPK1.2 (TaRPK2)	TraesCS1B02G314700	964	632–901	104	8	29.41	91.05	-0.046	Cell membrane
3	TaRPK1.3 (TaRPK3)	TraesCS1D02G303700	964	632–901	104	8	30.18	90.54	-0.04	Cell membrane
4	TaRPK1.4 (TaRPK4)	TraesCS3A02G340100	891	563–832	97	6	37.72	89.55	-0.108	Cell membrane
5	TaRPK1.5 (TaRPK5)	TraesCS3B02G371700	891	563–832	97	6	38.74	88.99	-0.109	Cell membrane
6	TaRPK1.6 (TaRPK6)	TraesCS3D02G333600	891	563–832	97	6	36.85	89.33	-0.099	Cell membrane
7	TaRPK1.10 (TaRPK10)	TraesCS2A02G176500	1,123	841–1,048	120	8	42.51	103.46	0.121	Cell membrane
8	TaRPK1.11 (TaRPK11)	TraesCS2B02G202900	1,123	841–1,112	120	7	42.79	104.32	0.138	Cell membrane
9	TaRPK1.12 (TaRPK12)	TraesCS2D02G183900	982	700–971	104	8	42.82	106.38	0.135	Cell membrane
10	TaRPK1.13 (TaRPK13)	TraesCS2A02G260600	728	499–720	80	9	52.17	101.35	-0.063	Cell membrane
11	TaRPK1.14 (TaRPK14)	TraesCS2B02G281400	609	330–601	67	9	47	98.74	-0.054	Cell membrane
12	TaRPK1.15 (TaRPK15)	TraesCS2D02G263100	732	453–724	80	9	51.16	100.26	-0.035	Cell membrane
13	TaRPK1.7 (TaRPK7)	TraesCS3A02G340000	923	593–863	101	7	31.67	89.51	-0.147	Cell membrane
14	TaRPK1.8 (TaRPK8)	TraesCS3B02G371600	957	627–897	104	7	31.92	90.2	-0.146	Cell membrane
15	TaRPK1.9 (TaRPK9)	TraesCS3D02G333500	924	594–864	101	7	31.27	87.93	-0.179	Cell membrane

Loc, location; Mol. wt., molecular weight; pI, isoelectric point; II, instability index; AI, aliphatic index; GRAVY, grand average of hydropathicity; SCL, sub-cellular localization.

visualization of 3D structure Pymol software (https://pymol.org/2/) was applied. The verification and validation of the predicted 3D structures of TaRPK1 proteins were assessed using the Ramachandran Plot—Zlab, (https://zlab.umassmed.edu/bu/rama/) (Anderson et al., 2005).

In Silico Differential Expression Patterns of *RPK* Genes

In silico expression analysis was performed using the wheat-expression browser (www.wheat-expression.com) at different wheat stages (Kaur et al., 2017). The data were unruffled in the course of developing seedling, vegetative, and reproductive stages from different organs of wheat such as roots, leaf sheath, leaf blade, shoot, spike, and grain. The heatmap was then created from the composed data, based on the expression values of genes (in TPM) by means of Tbtool.

Interaction Network and Co-Expression Analysis

For interaction network studies, String (https://string-db.org/) was used by selecting *Triticum aestivum* as a platform species. For visualization of the molecular library, Cytoscape was used.

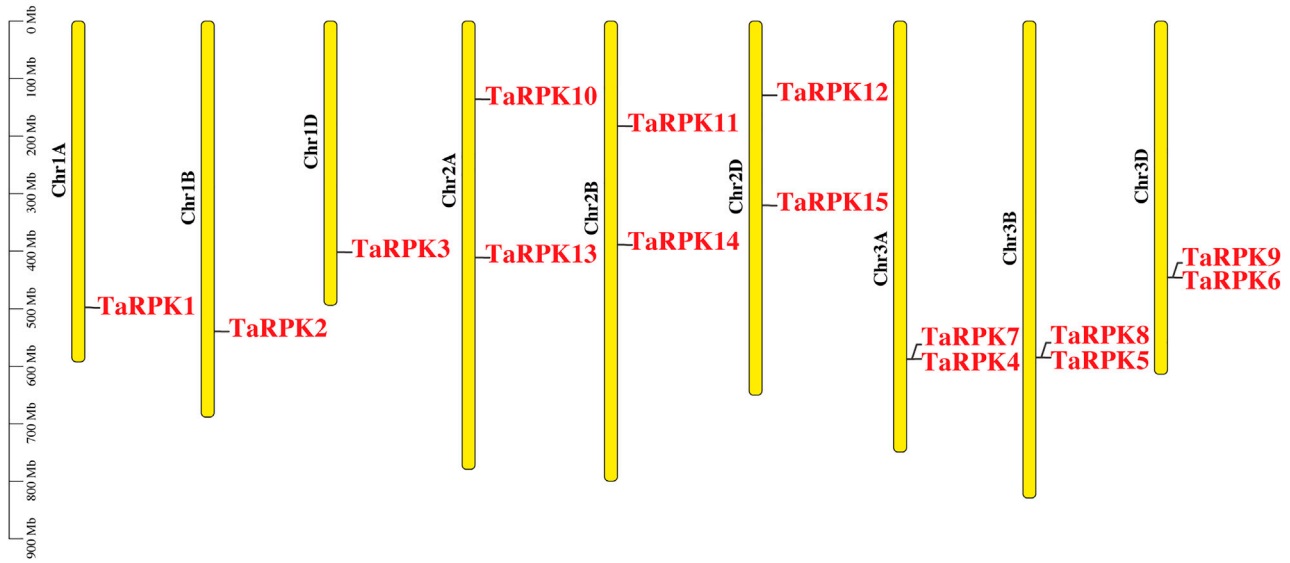


FIGURE 1 | Chromosomal location of *T. aestivum RPK* genes on chromosomes in A, B, and D genomes. Respective chromosome numbers are written as Chr IA to Chr 3D on the top of each chromosome. Gene position can be estimated using the scale (in megabase; Mb) on the left of the figure.

Correlation coefficients on the basis of verities, treatments, and tissues were calculated in R 3.4.0. These coefficients indicate the degree of association among the terms and provide linkages among the *TaRPK1* members.

Plant Material and Stress Treatment

Previously, Pakistan-13, Galaxy, and Shafaq were studied under drought stress and categorized as drought tolerant and susceptible varieties, respectively (Shabbir et al., 2015; Ulfat et al., 2017; Ahmad et al., 2021; Wasaya et al., 2021; Iqra et al., 2022). So, seeds of these varieties were obtained and sown under controlled glass-house conditions at the National Institute for Genomics and Advanced Biotechnology (NIGAB), National Agriculture Research Center (NARC), Islamabad, Pakistan. After 2 weeks of sowing (seedling stage), the roots and leaves tissues were collected. At growth stage 8 (tillering stage), roots, stems, and leaf tissues were collected. At the grain filling stage (14 days after flowering), sampling for roots, stems, leaf, and grains was done (Hyles et al., 2020). For expression profile analysis under drought stress, seeds of selected varieties were first surface sterilized with sodium hypochlorite followed by three washings, then soaked in distilled water in a growth chamber (16 h light/8 h dark cycle at 22°C). After 2 weeks, young seedlings were treated with 20% polyethylene glycol (PEG) 6,000 (v/v). The root and leaf tissues of seedlings were harvested after 12 h of exposure to stress conditions. All the samples were collected in three replicates, and samples were frozen immediately in liquid nitrogen, and placed in −80°C storage for RNA extraction.

RNA Isolation and qRT-PCR Analysis

Approximately 100 mg of tissues were taken for total RNA extraction using an RNA mini kit (Cat # 12183018A, Invitrogen, Thermo Fischer Scientific) followed by the manufacturer's instructions. Through agarose gel electrophoresis, the quality and concentration of RNA were determined, followed by optical density measurement through a spectrophotometer. With the help of the RT Prime-Script Reagent Kit, the cDNA was made from 1 ug of RNA. Specific primers were designed for *TaRPK1* genes manually, followed by confirmation via NCBI Primer Blast software (http://www.ncbi.nlm.gov/tools/primer-blast), provided in **Supplemental Table S2**. The qRT-PCR was accomplished with SYBR Green I (Roche) Master Mix. Wheat *β-Actin* was used as a control reference gene. Three independent biological replicates were analyzed for each sample. The values were means and standard deviations (SD) were calculated from biological replicates. The relative expression levels of each gene were studied by means of $2^{-\Delta\Delta Ct}$ (Schmittgen and Livak, 2008).

RESULTS

Analysis and Sequence Identification of *RPK1* Genes in *T. aestivum*

A set of 15 candidate *RPK1* genes were retrieved from *Triticum aestivum* based on BlastP and BlastN. A domain search by the SMART tool with the corresponding RPK1 candidate amino acid sequences confirmed the S_TKc Domain (SM00220). Thus, a total of 15 *TaRPK1* with complete structures were analyzed in *T. aestivum* (**Table 1**). Subsequent sequence identification of 15 TaRPK1 showed the protein length of 609–1,123 amino acids and a molecular mass ranged from 67–120 kDa. The iso-electric points (PI) of these proteins were 6–9. The Instability Index (II) ranged from 28.37–52.17, the Aliphatic Index (AI) was 87.93–106.38, and the grand-average of hydropathicity

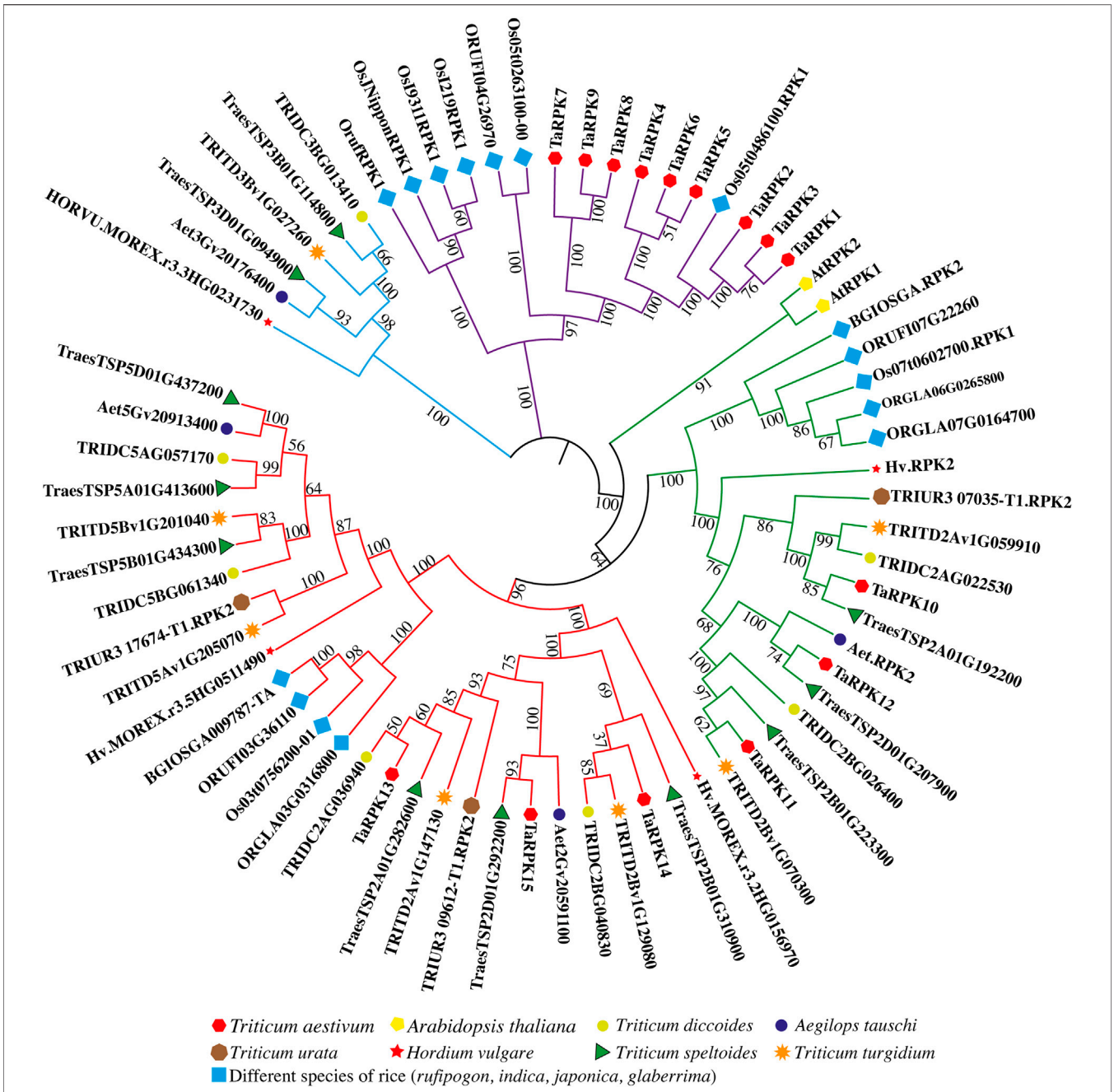


FIGURE 2 | Comparative phylogenetic tree of *RPK* genes between *Triticum aestivum*, *Triticum dicoccoides*, *Triticum turgidum*, *Triticum speltoides*, *Aegilops tauschii*, *Triticum urata*, *Hordeum vulgare*, *Arabidopsis thaliana,* and different species of *Oryza* (*rufipigon*, *indica*, *japonica*, and *glaberrima*). 1,000 replicates were used for the bootstrap test, and the replication percentage is presented next to the branches.

(GRAVY) −0.032–0.138. The instability index of group I was less than 40, representing stable proteins, whereas proteins of groups II and III showed instability index values of more than 40, indicating unstable proteins. The AI signified that all of the TaRPK1 proteins are thermally stable. The GRAVY indicated TaRPK1 proteins to be hydrophilic proteins except for TaRPK10, TaRPK11, and TaRPK12, which showed a value less than zero, representing them as hydrophobic proteins. The sub-cellular localizations of the TaRPK1 were anticipated, which showed that all the TaRPK1 were localized to the cell membrane (**Table 2**).

A detailed protein alignment of structural predictions showed that all TaRPK1 proteins are composed of the leucine-rich repeat N terminal (LRRNT_2) domain, leucine-rich repeat (LRR) domains, transmembrane domain I, and a serine–threonine kinase (S_TKc) domain. However, the LRR domains were

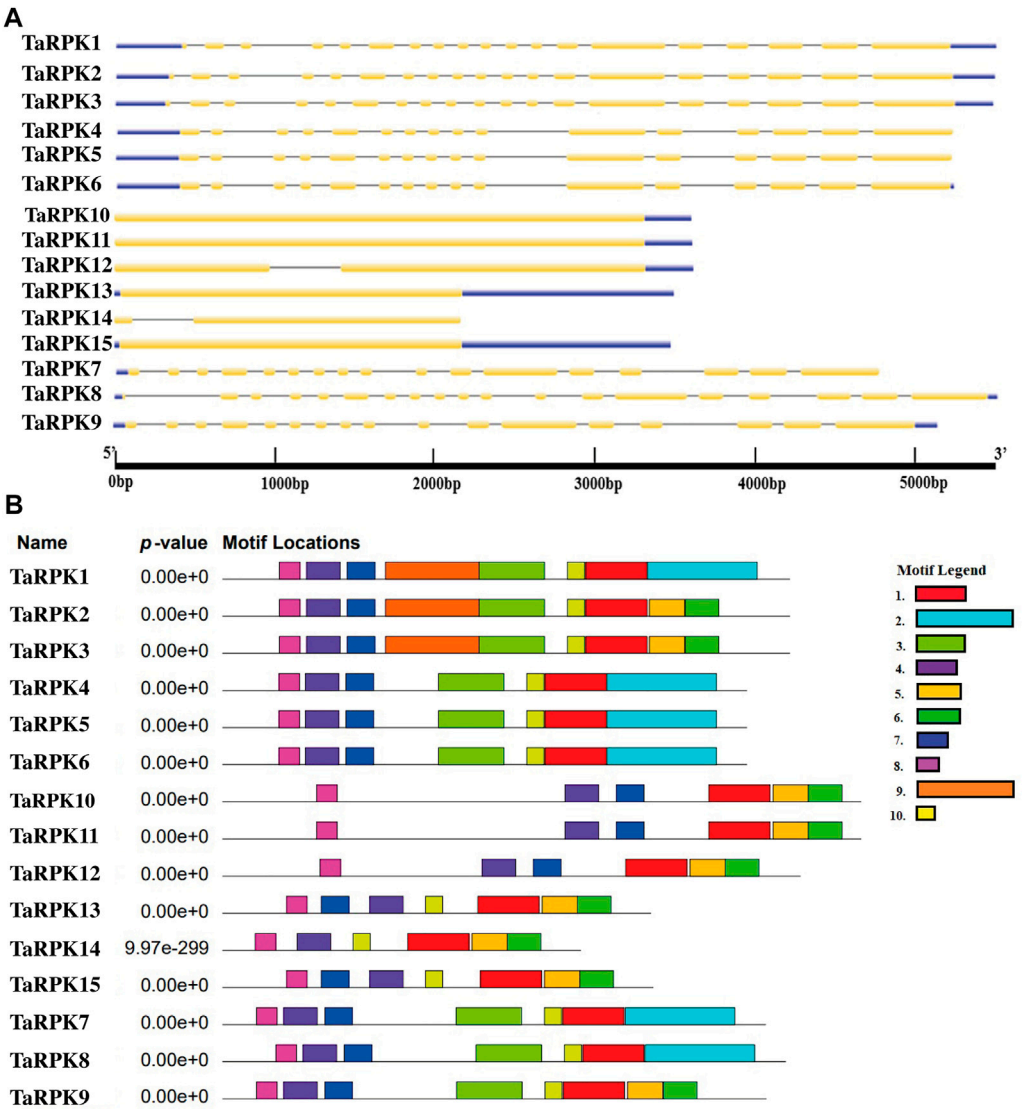


FIGURE 3 | (A) Exon–intron structures of *T. aestivum RPK1* genes. Yellow boxes denote exons, straight black lines represent introns, and blue boxes denote upstream/downstream. **(B)** Schematic representation of identified motifs in *T. aestivum* RPK1 proteins using the MEME motif search tool. Different colors indicate different motifs.

missing in TaRPK1, TaRPK2, and TaRPK3 sequences (**Supplemental Figure S1**).

Chromosomal Distribution of *RPK* Genes

The physical location of *RPK* genes in *T. aestivum*, to the corresponding chromosomes, is shown in **Figure 1**. A total of 15 *RPK* genes were mapped on 9 out of 21 chromosomes in wheat. The genes were mainly mapped on chromosomes 1, 2, and 3 on the respective A, B, and D genomes. No *RPK* genes were found on the rest of the chromosomes.

Phylogenetic Analysis of TaRPK1 Proteins

Of the 15 identified *TaRPK1* genes in this study in *Triticum aestivum*, two *RPK* genes from *Arabidopsis thaliana*, 16 *RPK* genes from rice, seven *RPK* genes from *Triticum dicoccoides*, three *RPK* genes from *Triticum urata*, seven *RPK* genes from *Triticum turgidium*, four *RPK* genes from *Aegilops tauschii*, 11 *RPK* genes from *Triticum speltoides*, and four *RPK* genes from *Hordeum vulgare* were used to construct a neighbor-joining based tree with MEGA X software in order to study the evolutionary relationships (**Figure 2**). The phylogenetic tree generated on the basis of similarities with protein sequences distributed *RPK* members into four main groups, with *TaRPK1* members in three groups. Overall group I possessed nine *TaRPK1* members (*TaRPK1-9*), that were closely associated with *RPK* members of rice. Group II (*TaRPK10-12*) and Group III (*TaRPK13-15*) exhibited three *TaRPK1* members each, that exhibited close association with *Triticum turgidium, Triticum speltoides, Aegilops tauschii* and *Triticum dicoccoides*.

Analysis of *TaRPK1* Gene Structure and Conserved Motif

The intron–exon number and arrangements of the *RPK1* members were envisaged through comparing the coding sequence with the genomic DNA sequence. All of the *TaRPK1* genes in group I consisted of 16–18 introns, except for the groups II and III that contained 0 and 1 intron (**Figure 3A**). Furthermore, the conserved motifs within TaRPK1 proteins were predicted by online MEME software. Ten conserved motifs (1-10) were analyzed (**Figure 3B** and **Supplementary Figure S2**). The motifs 1, 3, 4, 7, 8, and 10 were present in all of the *RPK1* sequences. However, group II did not display motifs 3 and 10, and the motif three was also missing in group III sequences.

Gene Ontology of *RPK1* Genes

GO annotation analysis was conducted for the functional analysis of *RPK1* genes. *In-silico* functional prediction was performed, and the results displayed two types of processes involved, that is, molecular processes (MPs) and biological processes (BPs) (**Figure 4** and **Supplementary Table S3**). Biological processes indicate that *RPK* members are involved actively in various metabolic processes. The molecular processes suggested the *RPK1* member's catalytic activity. Such outcomes clearly denote *RPK1* genes' significant role in growth and development via modulation of molecular and biological processes.

MicroRNA Targeting *TaRPK1* Genes

We also identified putative 18 miRNAs targeting *TaRPK1* genes for the generation of interaction networks by Cytoscape software in order to better understand the underlying miRNA mechanism involved in the modulation of *TaRPK1* genes (**Figure 5** and **Supplementary Table S4**). In the connection distribution and regulation network, *TaRPK1*, *TaRPK2,* and *TaRPK3* were found targeted by single miRNAs, which are tae-miR9782, tae-miR9776, and tae-miR1122c-3p, respectively. *TaRPK10* and *TaRPK11* are the most targeted *RPK1* wheat genes by tae-miR1134, tae-miR9774, tae-miR9661-5p, tae-miR9664-3p and tae-miR9777 targeting *TaRPK10,* and tae-miR9774, tae-miR9777, tae-miR9664-3p, tae-miR395a and tae-miR9661-5p targeting *TaRPK11* genes. However, no miRNA was found targeting *TaRPK13*, *TaRPK14*, and *TaRPK15* genes.

Regulatory *Cis*-Element Interpretation in *T. aestivum*

The promoter regions contain *cis*-modulatory elements which are critical for the binding of transcription factors for transcription initiation, which has an essential function in the expression of genes. The promoter regions of *RPK1* members were used for the *cis*-regulatory element prediction (**Figure 6A**). The results indicated that the *cis*-regulatory elements can be distributed into several categories, such as hormone related elements, light-related elements, developmental responsive elements, abiotic stress responsive elements, promoter-related motifs, and other motifs. Amid them, the elements chiefly present were associated with photoreaction, hormone responsiveness, and abiotic stress-related motifs. The photoreaction responsive cis-regulatory elements included ACE, AE-Box, ATCT, G-Box, GATA, GT1, SP1, AT1, Box 4, Box II, I-Box, TCT, GA, L-Box, TCCC, and ATC motif. The most abundant light-responsive elements were found in *TaRPK11* and *TaRPK13,* which had 17 and 12 members, respectively. Hormone responsive elements were also copiously present in the *RPK1* promoter, mostly comprising abscisic acid response elements. The three extensively distributed cis elements were related to abiotic stress response, among which drought responsive elements were profuse. Other elements correlated to abiotic stress were also identified.

Syntenic Relationship Analysis

In order to understand the evolutionary relationship and origin of *Triticum aestivum* (tr) with *Triticum turgidum* (tg), *Aegilops tauschii* (at), *Triticum speltoides* (ts) and *Tritium dicoccoides* (td), a comparative synteny scrutiny of RPK protein sequences was performed. The proteins were closely related among five species and exhibited significant similarity in analysis of evolutionary correlation. It was observed that the *TaRPK1* genes of *T. aestivum* have similar origins of evolution to other *Triticum* species (**Figure 6B** and **Supplementary Table S5**).

In silico 3D-Structure Prediction of TaRPK1 Proteins

Three-dimensional (3D) structures of TaRPK1 proteins were predicted by using SWISS_MODEL online computational software. 3D structures of target proteins were anticipated based on homology modeling. The SWISS MODEL predicted 15 successful models of TaRPK1 proteins with at least 30% identity to the template (4mn8.1. A, 5hyx.1. A, 5xkj.1. C, 6mOu.1. A, 4mna.1. A, 4oh4.1. A, 6cth.1. A, 7brc.1. A, and 5tos.1. A) that was a widely recognized threshold for effective modeling (Xiang, 2006). However, TaRPK2 and TaRPK3 showed sequence identity of 27.84% and 29.47%, respectively, with the template, which was less than 30%. The highest sequence identity of 45% with the template was observed by TaRPK4, TaRPK5 and TaRPK6 (**Figure 7**). The verification and validation of the predicted 3D structure of TaRPK1 were assessed via Ramachandran Plots (Anderson et al., 2005) that validated the backbone diahedral angles of the targeted protein. The Ramachandran plot assessment showed that 92–98% of the regions of TaRPK1 protein showed highly favorable regions, which indicates the stability and good quality of the predicted protein structure (**Supplemental Table S6**).

Genome Wide Expression Patterns of *RPK* Genes

The data of RNA-seq for all of the 15 *RPK* sequences were obtained from online database. A heatmap was generated showing expression levels of *RPK* members at different stages, namely seedling stage, vegetative stage, and reproductive stage

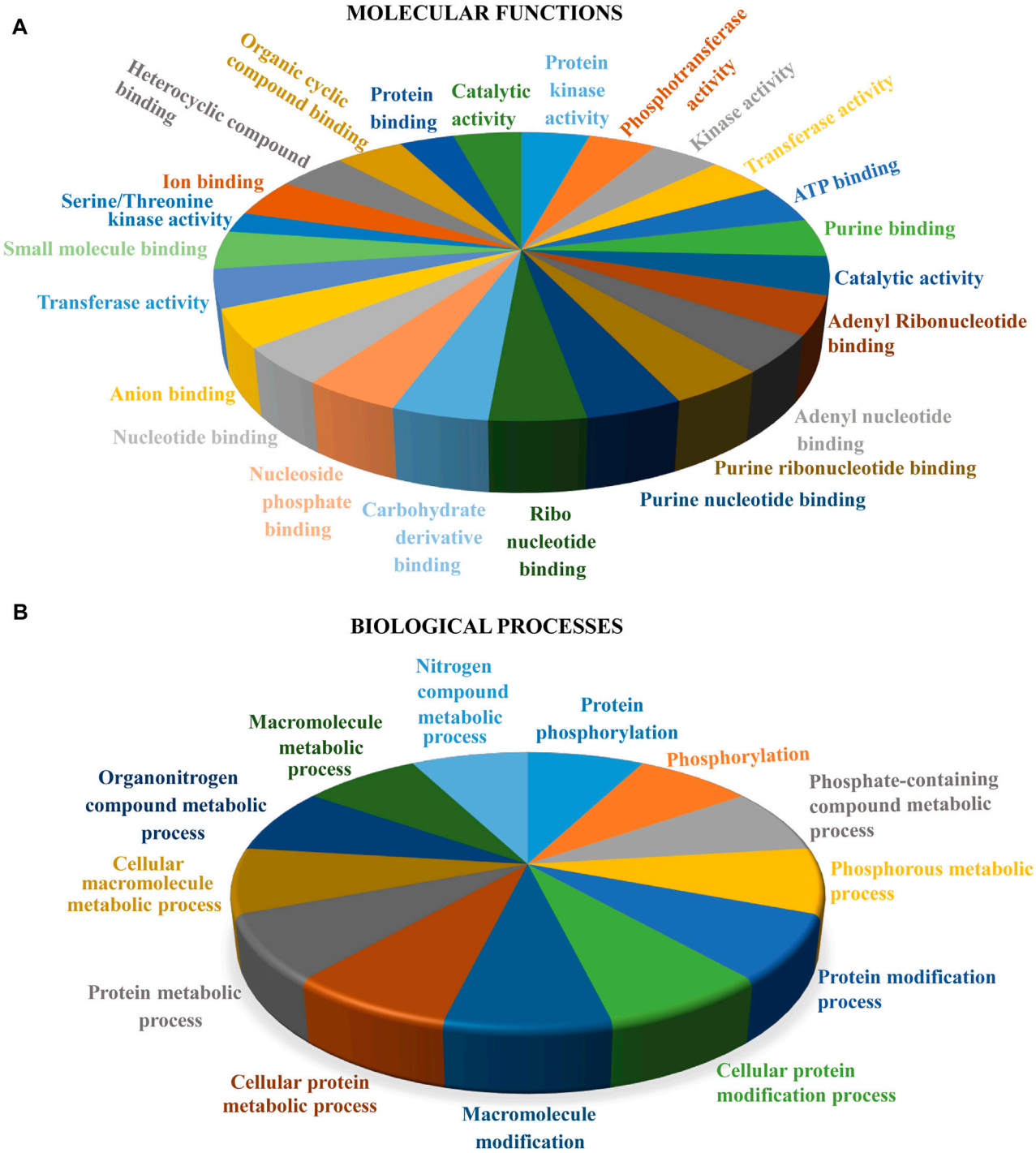


FIGURE 4 | Gene Ontology prediction of *RPK1* genes. The data represent **(A)** molecular functions and **(B)** biological processes.

(**Figure 8** and **Supplementary Table S7**) and in various organs (root, leaf, shoot, spike, and grain) of wheat. The highest expression of *TaRPK1* members was observed in root tissues compared to other tissues. *TaRPK1, TaRPK2,* and *TaRPK3* exhibited the highest expression patterns in roots at seedling, vegetative, and reproductive stages. Higher to moderate expression was observed in grain at the developing reproductive stage by *TaRPK13* and *TaRPK14*, respectively. Spikes, leaves, and shoots showed moderate to low expression in all of the *TaRPK1* members in wheat.

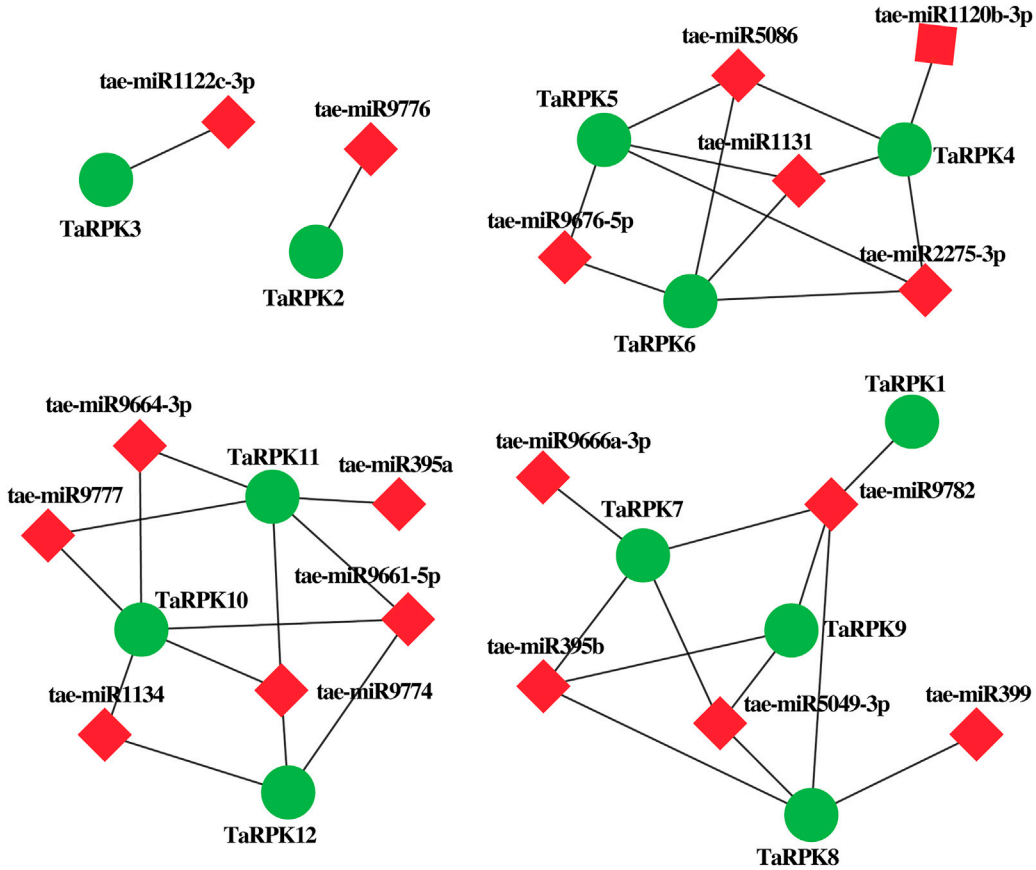


FIGURE 5 | MicroRNA and their targeted *TaRPK1* genes. Regulatory network relationship between miRNA and their targeted *TaRPK1* genes.

Expression Analysis of *RPK1* Genes in *T. aestivum*

The *TaRPK1* gene expression was determined in drought-tolerant (Pakistan 13 and Galaxy) and drought-susceptible (Shafaq) varieties under normal growth conditions in order to get a baseline expression profile. The expression pattern in all of the three varieties was examined in various developmental stages, including seedling stage, tillering stage, and heading stage and in different tissues such as root, stem, leaf, and grain (**Figure 9**). The *TaRPK1*, *TaRPK2,* and *TaRPK3* showed significant expression in the roots at the heading and seedling stages of the Pakistan-13 and Galaxy varieties. The *TaRPK13* exhibited higher expression in grain tissues of all varieties compared to other *TaRPK1*. The *TaRPK1* genes displayed higher expression in roots whereas they showed less expression in leaves and stems compared to the grain and root expression in developmental stages. Our results indicated that *TaRPK1* genes had similar expression patterns in both Pakistan 13 and Galaxy varieties, unlike the Shafaq variety. The higher expression of *TaRPK1* genes was observed in the heading > seedling > tillering stages in Pakistan 13, Galaxy, and Shafaq varieties. Overall, *TaRPK1* exhibited significant expression in root tissues compared to leaf, shoot, and grain tissues.

Roots are a good source to study the drought mechanism. To further confirm this, qRT-PCR showed the expression of *TaRPK1* members in the leaves and roots of two-week-old seedlings with drought stress through PEG simulation. PEG-6000 treatment induced an upregulated expression in roots and leaf tissues in comparison to the susceptible genotype. Higher expression was observed in root seedlings in comparison to the leaf seedlings, except for *TaRPK4* and *TaRPK7*, where higher expression was detected in the leaf tissues compared to the root tissues under drought stress (**Figure 10**). The *TaRPK1* genes displayed higher expression in Pakistan 13 > Galaxy > Shafaq varieties. Furthermore, we also performed co-expression (**Supplementary Figure S3**) and interaction network (**Supplementary Figure S4** and **Supplementary Table S8**) analyses and the results revealed that all the *RPK1* members showed highly significant associations. These results indicate *TaRPK1* gene involvement in drought stress regulation.

DISCUSSION

RPK1 is a serine/threonine protein kinase and belongs to the subfamily LRRKs, which is the largest subfamily of RLK. The LRRKs play a crucial role in a large number of biological activities, from development and growth to stress management in plants (Dufayard et al., 2017). *RPK* genes play significant roles in root

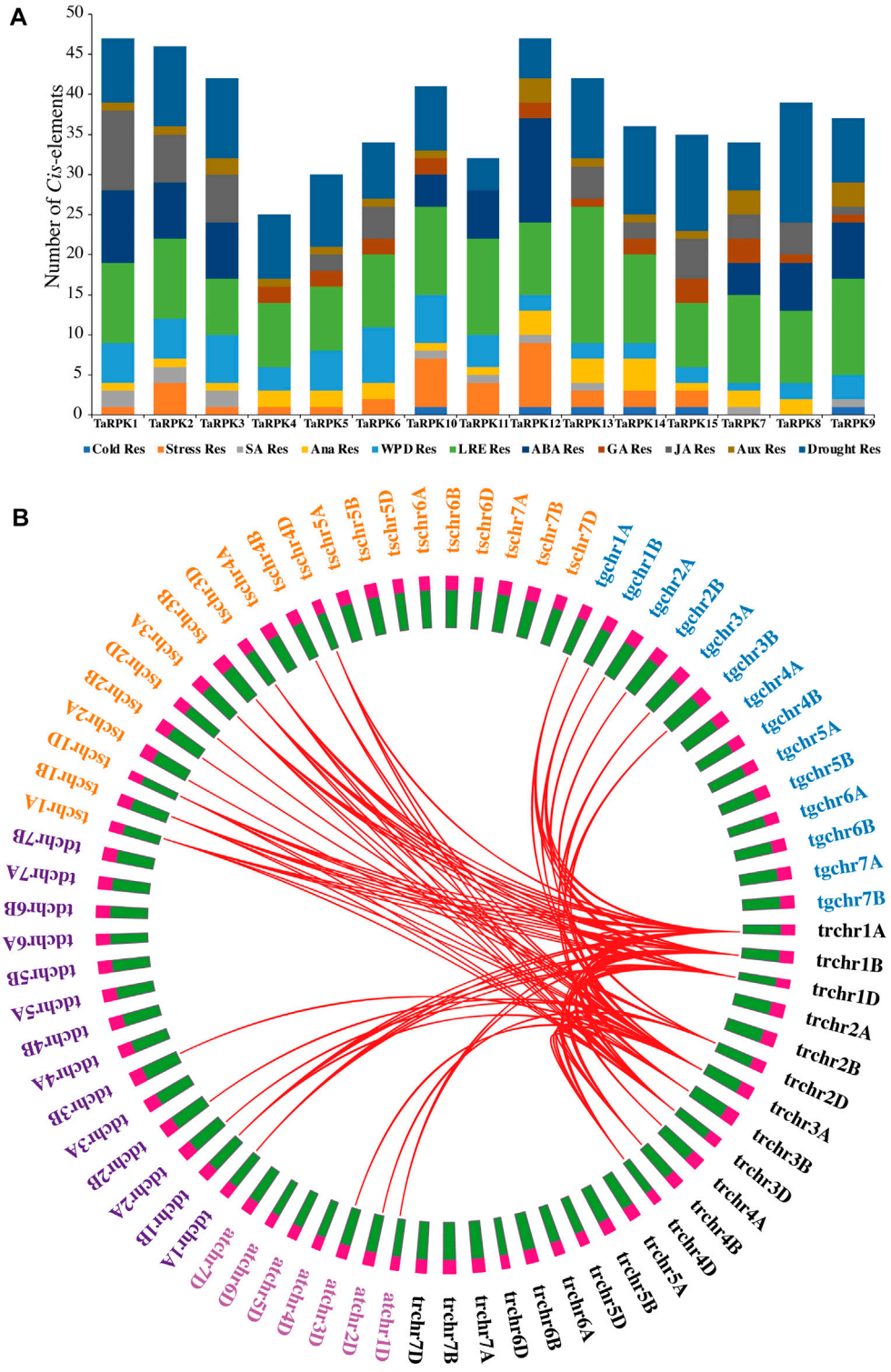


FIGURE 6 | *Cis*-Elements and evolutionary conservation. **(A)** Regulatory *Cis*-element prediction of 2-Kb sequence upstream of *RPK1* genes in *T. aestivum*. **(B)** Syntenic relationship between *Triticum aestivum* (*tr,* black), *Triticum dicoccoides* (*td*, purple), *Triticum turgidum* (*tg*, blue), *Aegilops tauschii* (*at,* pink), and *Triticum speltoides* (*ts*, orange).

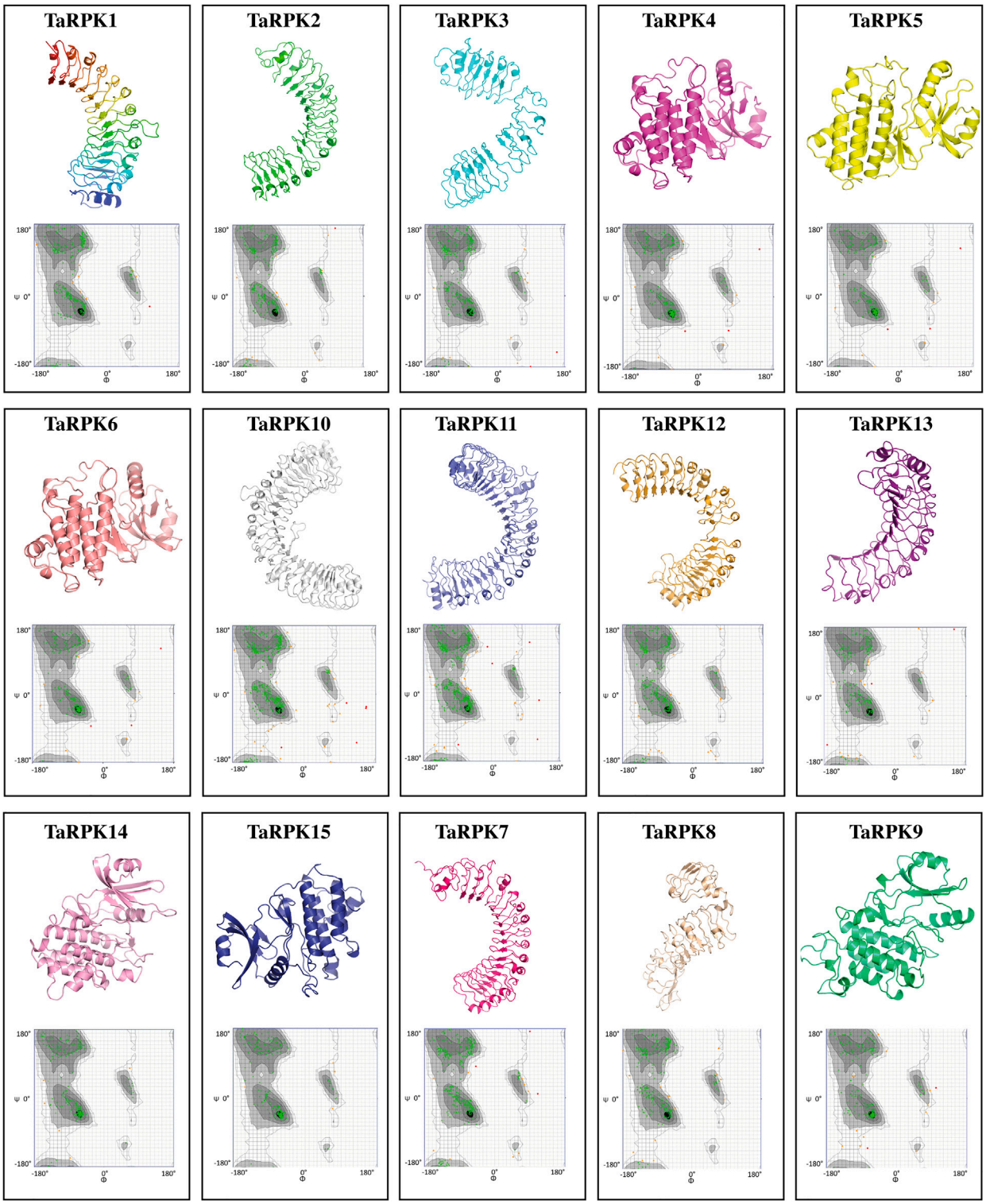


FIGURE 7 | 3D structure of *TaRPK1* proteins along with Ramachandran plots in *T. aestivum*. In all 3D protein structures, the spirals are helices, broad strips with arrow heads are beta-pleated sheets, and thin loops are coils. In Ramachandran plots, dark black, gray, and light gray represent highly preferred conformations with Delta ≥ −2. White with a black grid denotes preferred conformations with −2 > Delta ≥ −4. White with gray grid symbolizes questionable conformations with Delta < −4. The green crosses signify highly preferred observations, brown triangles specify preferred observations, and red circles represent unfavorable observations.

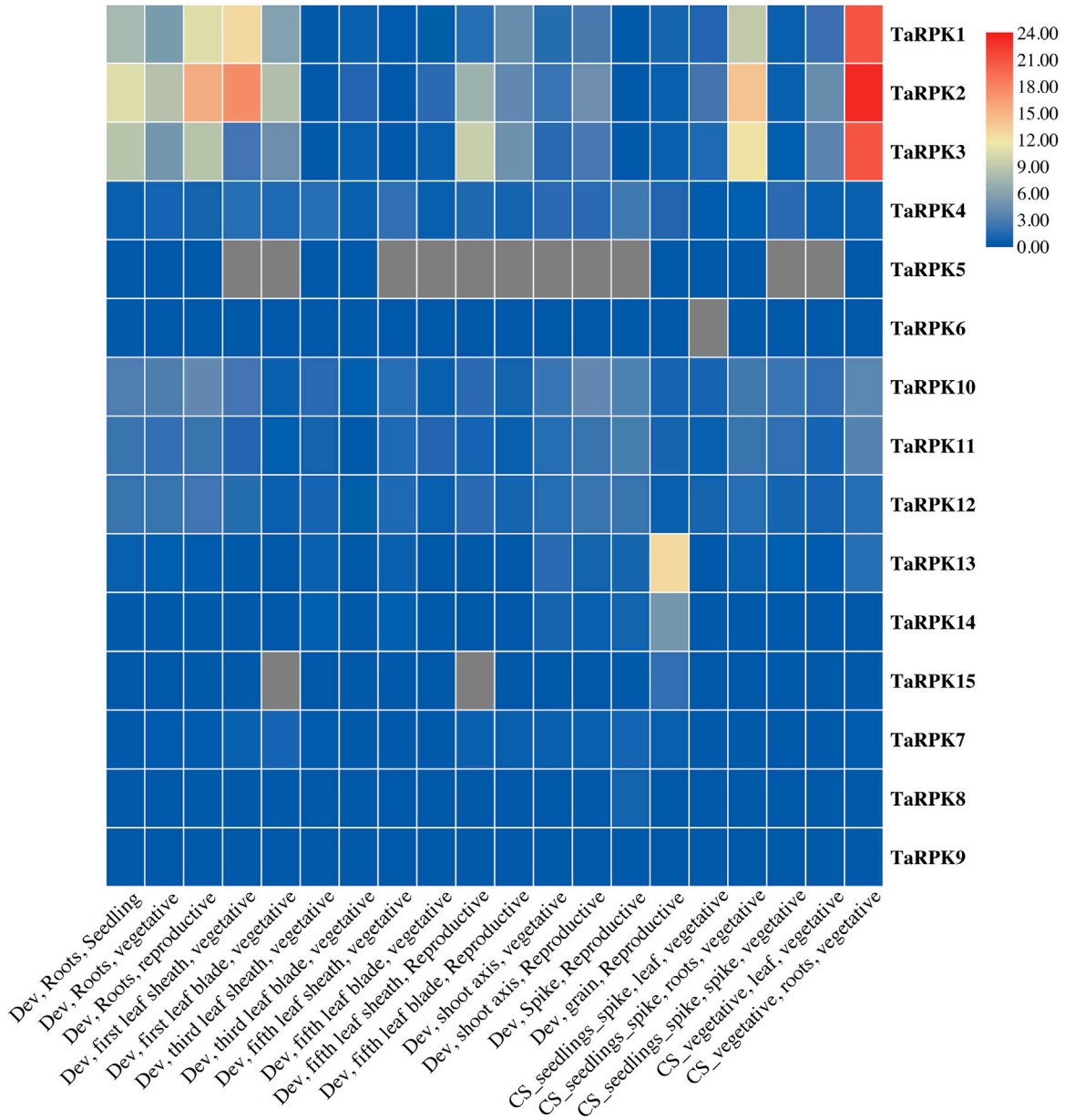


FIGURE 8 | *In silico* analysis of *RPK* members in different tissues (root, shoot, leaf, grain, and spike) was generated using TB tool. Dev = developmental stage, CS = seedlings spike; Chinese Spring seedlings (leaves, roots) and spikes at anthesis and CS vegetative; Chinese Spring leaves and roots from seven leaf stages. The heatmap was constructed from transcripts per 10 million values with the scale bar displaying expression of the genes. The blue and red colors denote lower and higher expressions of the transcripts, respectively.

system architecture (RSA), plant height, number of tillers, salt tolerance, cotyledon primordial initiation of cotyledons during embryogenesis, ABA-induced senescence, and ROS production (Shi et al., 2014; Zou et al., 2014; Dai et al., 2018). The functional characterization of *RPK1* members in wheat has not been reported in previous studies. The reason for it could be a complex allohexaploid (2n = 6x = 42) genome and other factors. Since *RPK1* genes are accountable for essential roles in plants, therefore a comprehensive study was performed to determine the chromosomal location, phylogenetic analysis, gene structure and expression of these genes in *T. aestivum*.

The standard process for the sequence identification of a new gene family is by a BLAST search of sequences of known proteins in model plants. A similar BLAST method was used to characterize two *RPK* genes from *Arabidopsis thaliana*, 16 *RPK* genes from different species of rice, seven *RPK* genes from *Triticum dicoccoides*, three *RPK* genes from *Triticum urata*, seven *RPK* genes from *Triticum turgidum*, four *RPK* genes from

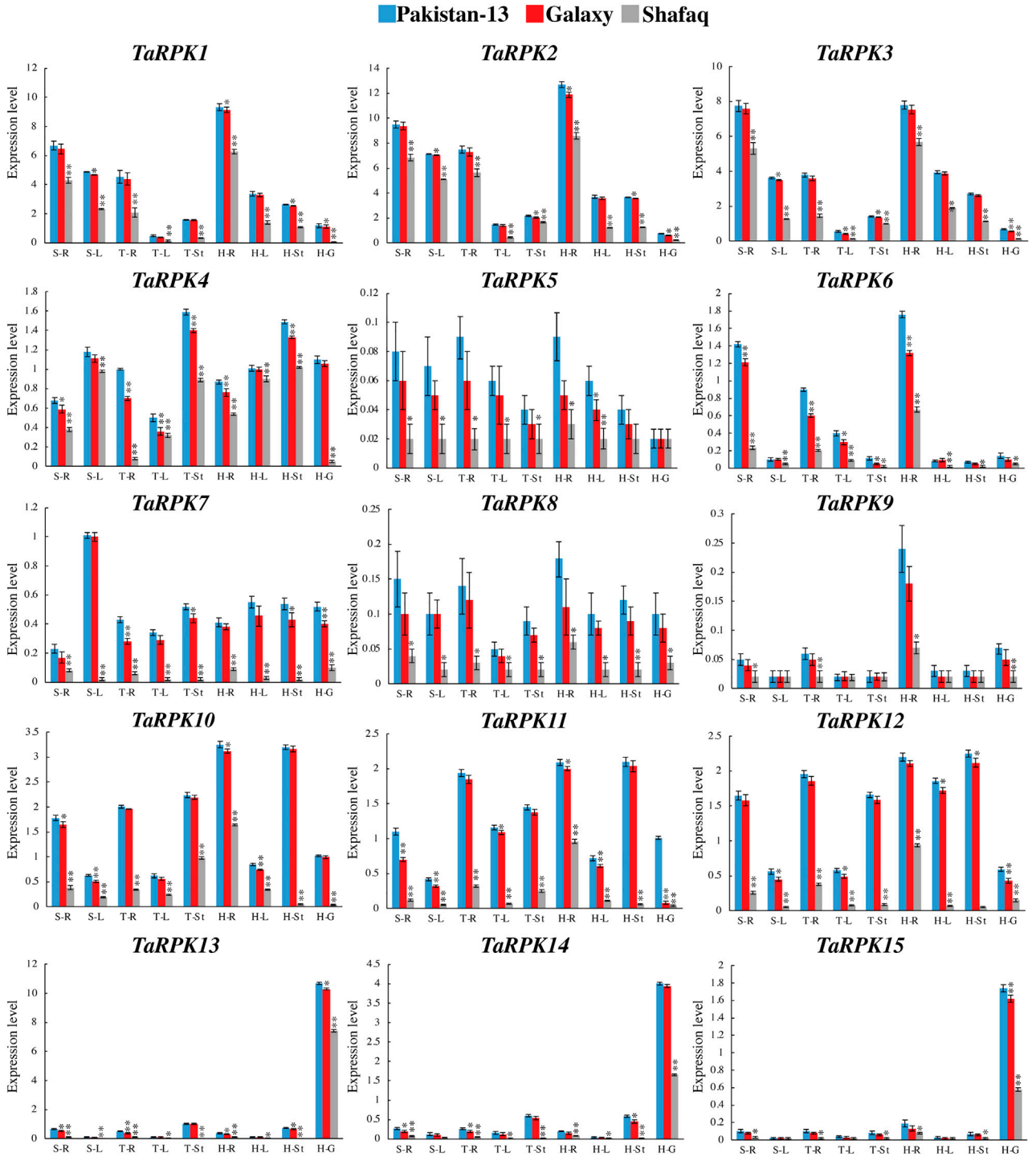


FIGURE 9 | Real time PCR-expression analysis of *TaRPK1* genes in wheat varieties; Pakistan 13 (blue), Galaxy (red), and Shafaq (gray). The wheat plant was germinated and grown in soil under normal conditions. Expressions of *TaRPK1* genes were determined in root (R), stem (St), leave (L), and grain (G) at seedling (S), tillering (T), and heading (H) stages. Standard deviation (SD) of three biological replicates is represented by the error bars. Significance was assessed by using a *t*-test (*$p < 0.05$, **$p < 0.01$, and ns = non-significant).

Aegilops tauschii, 11 *RPK* genes from *Triticum speltoides*, four *RPK* genes from *Hordeum vulgare*, and 15 *TaRPK1* genes in *Triticum aestivum*. The number of *TaRPKs* identified in *T. aestivum* is similar to that of *RPKs* in *Triticum speltoides* (11). The identified *RPK* genes were confirmed for the conserved domains by the SMART database. A higher number of

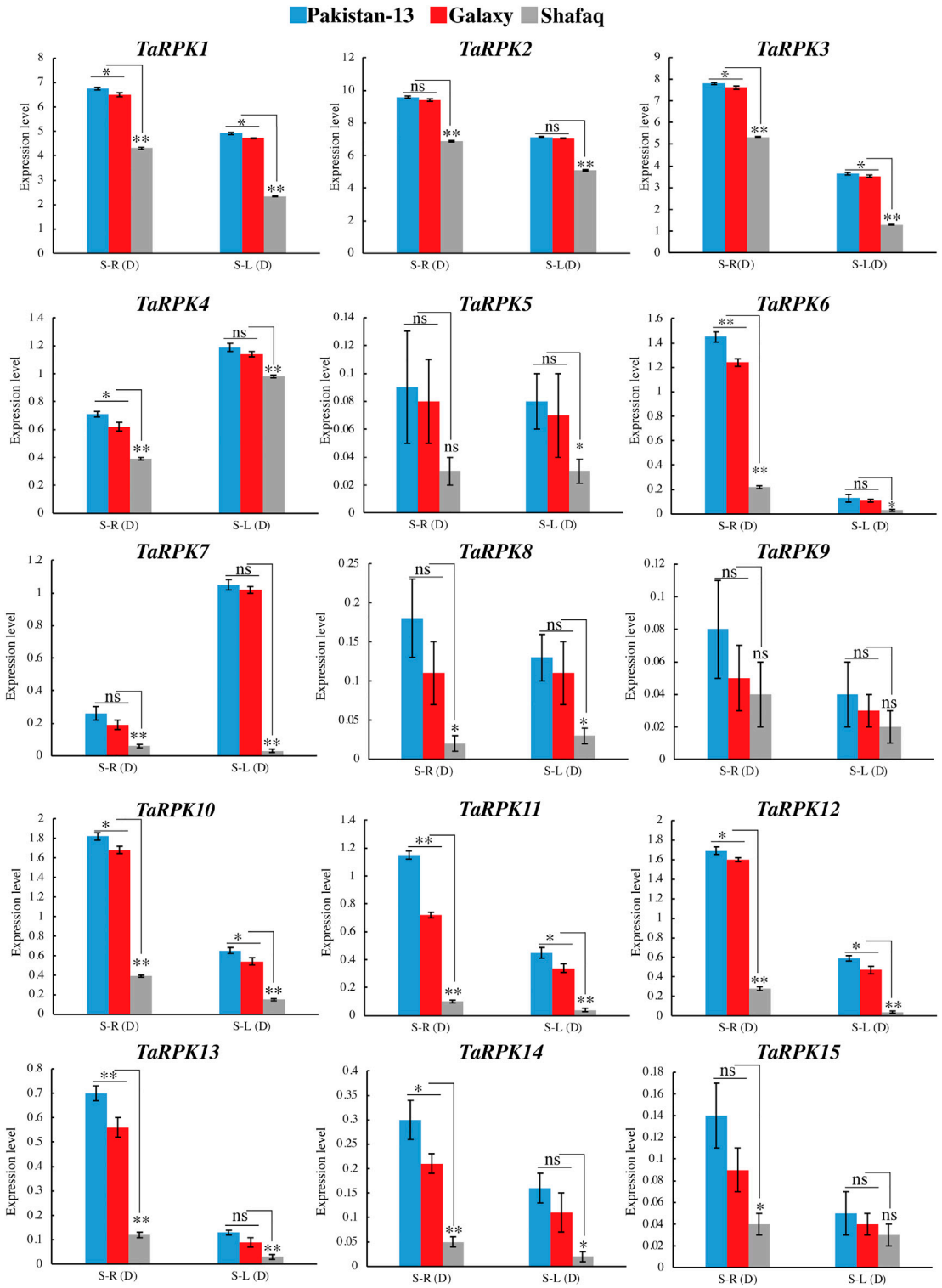


FIGURE 10 | Expression profiling of *TaRPK1* genes under 20% PEG stress in Pakistan 13, Galaxy, and Shafaq varieties. Expressions of *TaRPK1* genes were determined in root (S–R) and leave (S–L) at the seedling stage (S). Error bars denote standard errors of three biological replicates. Significance was assessed by using a *t*-test (*$p < 0.05$, **$p < 0.01$, and ns = non-significant).

TaRPK1 genes might be because of the large allohexaploid nature of the bread wheat genome.

The allohexaploid *T. aestivum* genome was originated due to the 3A, B, and D diploid sub genomes hybridization (Marcussen et al., 2014). Three homoeologous genes at a minimum should be for each *T. aestivum* gene, that is, one from each sub genome, also named as homoeologous genes for their homologous chromosomal localization (Sharma et al., 2016). The genome wide analysis displayed that *TaRPK1* genes along with the homoeologous genes were located mainly on chromosomes 1, 2, and 3 on A, B, and D sub-genomes, which showed that there might be no deletion of *TaRPK1* genes in the course of the acclimatization and evolution process of *T. aestivum*. The *TaRPK1* genes were found to be with maximum number on chromosome 2 and 3 (**Figure 1**) which was very similar to other studied crops. Crops such as *Triticum dicoccoides, Aegilops tauschii*, *Hordeum vulgare*, *Triticum speltoides,* and *Triticum turgidum* also showed the distribution of *RPKs* on chromosome 2 and 3, in addition to chromosome 5. However, the *RPKs* were distributed on chromosomes 3, 4, and 7 in rice, and in *Arabidopsis thaliana* they were on chromosomes 1 and 3.

The phylogenetic relationship was studied using complete TaRPK1 protein sequences, as it indicated evolutionary inference. The known homoeologous sequences were clustered closely (**Figure 2**), which indicated further evolutionary relationships and homology of sequences among them. The putative paralogous sequences were grouped together by those that specified similar origins. Similarity in organization and architecture of domains and motifs in clades designates functional association between these proteins. The gene structure analysis revealed intron numbers in *TaRPK1* genes that ranged from 0 to 1 and 16-18 (**Figure 3A**). The difference in the number of exons in *TaRPK1* was analogous to the one observed in other crops. *Triticum dicoccoides, Aegilops tauschii*, *Hordeum vulgare,* and *Triticum speltoides* exhibited one to two coding exons, and *Triticum turgidum* had one to three coding exons. *Oryza* species also exhibited one to three exons except for *Oryza rufipogon*; ORUFI04G26970 had 102 exons and Os05t0486100-01 *RPK1* exhibited 18 exons. This points toward evolutionary conservation and hence expression of genes between these species.

Prediction of protein domain configuration revealed the similarity to the previously studied RPK proteins (Cheng et al., 2009), with conserved C-terminal Ser/Thr kinase, a transmembrane domain suggesting membrane-bound features of TaRPK1 proteins and LRR domains. The LRR domains were absent in TaRPK1, TaRPK2, and TaRPK3 proteins. However, all *TaRPK1* members showed an additional LRRNT_2 (leucine-rich repeat N-terminal) domain in the N-terminal region of the amino acid (**Supplemental Figure S1**). In addition to sequence alignment, motif analysis also displayed the conservation of the motif at the initial N-terminal region and kinase domain with the motif that remained conserved in all of the 15 TaRPK1 protein sequences (**Figure 3B**). For the functional analysis of *TaRPK1* genes, the gene ontology enrichment analysis was performed (**Figure 4**). *In silico* prediction showed that *TaRPK1* members were involved in several processes of development through regulation of molecular functions (MFs) and biological processes (BPs), and exhibited response to environmental stresses. Several prior studies also described that through monitoring expression of genes, microRNAs respond to stress stimuli (Yan et al., 2019; Rasool et al., 2021; Rehman et al., 2022). The microRNAs are 21–24 nucleotides long endogenous non-coding RNAs that regulates development, growth, and adaptive response against abiotic stresses via monitoring target genes at posttranscriptional level or translation level of protein synthesis (Bai et al., 2017). In this study, we recognized microRNAs and their target genes in order to explore specific transcripts involved in development and growth processes and in response to different stress environments. We identified that miRNAs are majorly involved in cleavage mechanisms rather than translation inhibition (**Figure 5**).

The *cis*-regulatory elements identified in *TaRPK1* were mostly related to light responsiveness (**Figure 6A**). Other distributed *cis*-regulatory elements were related to stress factors, such as drought, cold stress, anaerobic response, wounding pathogens, and defensive elements. Functional relation of other *cis* elements was linked to plant hormones comprising auxins, abscisic, gibberellin and salicylic acid. Thus, the occurrence of various groups of *cis*-regulatory elements functioning in diverse physiological processes is suggestive of the dynamic *RPK1* gene regulation in *T. aestivum.* Synteny analysis with other ancestral *Triticum* species revealed that the *RPK1* gene family converges to a single ancestor (**Figure 6B**). This relationship validates that *RPKs* with analogous evolutionary status might have similar functions in plant growth and development. Homology models for 15 TaRPK1 proteins were made and evaluated with homologous templates. The TaRPK1 proteins exhibited 28%–45% identity to the template, which is a widely accepted threshold for successful modeling. The Ramachandran plots verification and validation displayed that a very higher percentage of all 15 TaRPK1 protein regions showed highly favorable regions that denote good quality protein structure prediction (**Figure 7**). Previous studies have shown similar 3D structure of TATrx proteins in wheat through homology modeling along with Ramachandran plot. The proteins were compared to 2iwt.1. A, 2vlt.1. A, 1fb0.1. A, 3d22.1. A, 2vm1.3. A, and 1faa.1. A templates, and the Ramachandran plot showed more than 95% of the thioredoxin amino acids lying in the most favored area (Bhurta et al., 2022). Another study in wheat has shown similar three-dimensional structure prediction of twenty-one TaEIL proteins via SWISS-MODEL along with Ramachandran plot analysis. The prediction model on the basis of templates heuristically enhanced percentage identification, alignment range, and confidence score of test sequences. The Ramachandran plot analysis confirmed 80% of residuals in the allowed area, signifying the quality of the model (Yi-Qin et al., 2020).

The gene expression in a specific tissue can be used as an information source for function identification in that tissue. Studies have revealed that *OsRPK1* overexpression altered the total architecture of roots in transgenic seedlings along with height, tillering numbers, and apical meristem of roots (Zou et al., 2014). The larger root system might result in a substantial upsurge in water and nutrient uptake. The relative expression level in different

tissues of *OsRPK1* was studied, which indicated higher to lower expression in the pattern of root tips > leaf blades > roots > leaf sheath > stem (Zou et al., 2014). Alike expression pattern was also detected in *TaRPK1* genes. The heatmap generated showed significant expression of *RPK1* genes in the root tissIes in comparison to the other tissues studied (**Figure 9**). As gene expression profiles are always related to their function, we further investigated their expression profiles in various tissues and varieties under normal and drought stress responses. Results of real-time quantitative PCR indicated that *TaRPK1* showed higher expression levels in root tissues at seedling and heading stages under normal conditions. *TaRPK1* exhibited tissue specific expression and showed higher expression in drought stress treatment in root tissues (**Figure 10**). This high expression in particular organs like roots indicates their particular roles in the root development and function of that tissue. The higher expression of *TaRPK1* genes was observed in Pakistan 13 > Galaxy > Shafaq varieties, which indicates the vital role of *TaRPK1* in plant growth and development. The sequence similarity and conserved domains of these protein kinases from *Arabidopsis*, rice, and wheat combined with the evidence from *in-silico* expression analysis and RT-PCR suggest that *TaRPK1* might share analogous functions in root development and hence yield. Future functional validation of these genes will be required.

CONCLUSION

We completely investigated the properties, developmental, location on chromosomes, *cis*-components, synteny, and expression profiles of *TaRPK1* members. An aggregate of 15 *TaRPK1*s were distinguished in the *T. aestivum* genome. This work can fill in as an initial phase in the complete useful portrayal of *RPK1* genes by reversible genetic methodologies. This study provides helpful assets to future investigations on the design and function of *RPK1* genes and for distinguishing and describing these genes in different species. Consequently, the outcomes might offer important data to examine the role of *TaRPK1* genes being developed and stress reactions through present-day practical genomics tools (next-generation sequencing) and genome editing, henceforth clearing the way toward genetic improvement of wheat.

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

MK and AR perceived the idea and planned the experiments. AR performed the research and wrote the manuscript. NZ and OR helped in data analysis. NR and MU helped in manuscript revision. MK supervised the research and acquired resources.

FUNDING

This research was conducted with funds from the Agricultural Linkages Program of Pakistan (ALP project number CS-399).

ACKNOWLEDGMENTS

We acknowledge the research facilities provided by the National Institute for Genomics and Advanced Biotechnology (NIGAB), Pakistan. We also thank the Ensembl plants, NCBI, and various other tools for the availability of sequences and data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.912251/full#supplementary-material

REFERENCES

Abbas, H., Naeem, M. K., Rubab, M., Widemann, E., Uzair, M., Zahra, N., et al. (2022). Role of Wheat Phosphorus Starvation Tolerance 1 Genes in Phosphorus Acquisition and Root Architecture. *Genes.* 13, 487. doi:10.3390/genes13030487

Afzal, A. J., Wood, A. J., and Lightfoot, D. A. (2008). Plant Receptor-like Serine Threonine Kinases: Roles in Signaling and Plant Defense. *Mpmi* 21, 507–517. doi:10.1094/mpmi-21-5-0507

Ahmad, Z., Ishaq, M., Tahir, I., Khan, A., and Jehan, K. (2021). Heat Tolerance Stability of Bread Wheat Genotypes under Early and Late Planting Environments through Stress Selection Indices. *Curr. Res. Agric. Sci.* 8, 37–46. doi:10.18488/journal.68.2021.81.37.46

Alahmad, S., El Hassouni, K., Bassi, F. M., Dinglasan, E., Youssef, C., Quarry, G., et al. (2019). A Major Root Architecture QTL Responding to Water Limitation in Durum Wheat. *Front. Plant Sci.* 10, 436. doi:10.3389/fpls.2019.00436

Anderson, R. J., Weng, Z., Campbell, R. K., and Jiang, X. (2005). Main-chain Conformational Tendencies of Amino Acids. *Proteins* 60, 679–689. doi:10.1002/prot.20530

Bai, J.-F., Wang, Y.-K., Wang, P., Duan, W.-J., Yuan, S.-H., Sun, H., et al. (2017). Uncovering Male Fertility Transition Responsive miRNA in a Wheat Photo-Thermosensitive Genic Male Sterile Line by Deep Sequencing and Degradome Analysis. *Front. Plant Sci.* 8, 1370. doi:10.3389/fpls.2017.01370

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

Bhurta, R., Hurali, D. T., Tyagi, S., Sathee, L., Adavi, B., Singh, D., et al. (2022). Genome-Wide Identification and Expression Analysis of the Thioredoxin (Trx) Gene Family Reveals its Role in Leaf Rust Resistance in Wheat (*Triticum aestivum* L.). *Front. Genet.*, 563. doi:10.3389/fgene.2022.836030

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., et al. (2014). SWISS-MODEL: Modelling Protein Tertiary and Quaternary Structure Using Evolutionary Information. *Nucleic acids Res.* 42, W252–W258. doi:10.1093/nar/gku340

Cao, Y., Liu, M., Long, H., Zhao, Q., Jiang, L., and Zhang, L. (2020). Hidden in Plain Sight: Systematic Investigation of Leucine-Rich Repeat Containing Genes Unveil the Their Regulatory Network in Response to Fusarium Wilt in Tung Tree. *Int. J. Biol. Macromol.* 163, 1759–1767. doi:10.1016/j.ijbiomac.2020.09.106

Chen, C., Chen, H., He, Y., and Xia, R. (2018). TBtools, a Toolkit for Biologists Integrating Various Biological Data Handling Tools with a User-Friendly Interface. BioRxiv 289660.

Cheng, X., Xiao, X., and Chou, K.-C. (2017). pLoc-mPlant: Predict Subcellular Localization of Multi-Location Plant Proteins by Incorporating the Optimal GO Information into General PseAAC. *Mol. Biosyst.* 13, 1722–1727. doi:10.1039/c7mb00267j

Cheng, Y., Qi, Y., Zhu, Q., Chen, X., Wang, N., Zhao, X., et al. (2009). New Changes in the Plasma-Membrane-Associated Proteome of Rice Roots under Salt Stress. *Proteomics* 9, 3100–3114. doi:10.1002/pmic.200800340

Dai, C., Lee, Y., Lee, I. C., Nam, H. G., and Kwak, J. M. (2018). Calmodulin 1 Regulates Senescence and ABA Response in Arabidopsis. *Front. Plant Sci.* 9, 803. doi:10.3389/fpls.2018.00803

Dai, X., and Zhao, P. X. (2011). psRNATarget: a Plant Small RNA Target Analysis Server. *Nucleic acids Res.* 39, W155–W159. doi:10.1093/nar/gkr319

Danakumara, T., Kumari, J., Singh, A. K., Sinha, S. K., Pradhan, A. K., Sharma, S., et al. (2021). Genetic Dissection of Seedling Root System Architectural Traits in a Diverse Panel of Hexaploid Wheat through Multi-Locus Genome-wide Association Mapping for Improving Drought Tolerance. *Ijms* 22, 7188. doi:10.3390/ijms22137188

Diévart, A., and Clark, S. E. (2003). Using Mutant Alleles to Determine the Structure and Function of Leucine-Rich Repeat Receptor-like Kinases. *Curr. Opin. Plant Biol.* 6 (5), 507–516. doi:10.1016/s1369-5266(03)00089-x

Dievart, A., Gottin, C., Périn, C., Ranwez, V., and Chantret, N. (2020). Origin and Diversity of Plant Receptor-like Kinases. *Annu. Rev. Plant Biol.* 71, 131–156. doi:10.1146/annurev-arplant-073019-025927

Djanaguiraman, M., Prasad, P. V. V., Kumari, J., Sehgal, S. K., Friebe, B., Djalovic, I., et al. (2019). Alien Chromosome Segment from Aegilops Speltoides and Dasypyrum Villosum Increases Drought Tolerance in Wheat via Profuse and Deep Root System. *BMC Plant Biol.* 19, 242–315. doi:10.1186/s12870-019-1833-8

Dufayard, J. F., Bettembourg, M., Fischer, I., Droc, G., Guiderdoni, E., Périn, C., et al. (2017). New Insights on Leucine-Rich Repeats Receptor-like Kinase Orthologous Relationships in Angiosperms. *Front. Plant Sci.* 8, 381. doi:10.3389/fpls.2017.00381

Fang, Y., Du, Y., Wang, J., Wu, A., Qiao, S., Xu, B., et al. (2017). Moderate Drought Stress Affected Root Growth and Grain Yield in Old, Modern and Newly Released Cultivars of Winter Wheat. *Front. Plant Sci.* 8, 672. doi:10.3389/fpls.2017.00672

Fantl, W. J., Johnson, D. E., and Williams, L. T. (1993). Signalling by Receptor Tyrosine Kinases. *Annu. Rev. Biochem.* 62, 453–481. doi:10.1146/annurev.bi.62.070193.002321

Gou, X., He, K., Yang, H., Yuan, T., Lin, H., Clouse, S. D., et al. (2010). Genome-wide Cloning and Sequence Analysis of Leucine-Rich Repeat Receptor-like Protein Kinase Genes in *Arabidopsis thaliana*. *BMC genomics* 11, 19–15. doi:10.1186/1471-2164-11-19

Grzesiak, M. T., Hordyńska, N., Maksymowicz, A., Grzesiak, S., and Szechyńska-Hebda, M. (2019). Variation Among Spring Wheat (*Triticum aestivum* L.) Genotypes in Response to the Drought Stress. II-Root System Structure. *Plants* 8, 584. doi:10.3390/plants8120584

Hong, S. W., Jon, J. H., Kwak, J. M., and Nam, H. G. (1997). Identification of a Receptor-like Protein Kinase Gene Rapidly Induced by Abscisic Acid, Dehydration, High Salt, and Cold Treatments in *Arabidopsis thaliana*. *Plant physiol.* 113, 1203–1212. doi:10.1104/pp.113.4.1203

Hyles, J., Bloomfield, M. T., Hunt, J. R., Trethowan, R. M., and Trevaskis, B. (2020). Phenology and Related Traits for Wheat Adaptation. *Heredity* 125, 417–430. doi:10.1038/s41437-020-0320-1

Iqra, T., Ullah, K. N., Samrin, G., Sheraz, A., Yusra, A., Amina, H., et al. (2022). Late Planting Induced Heat Stress Tolerance in Wheat. *Pak. J. Bot.* 54.

Kaur, S., Dhugga, K. S., Beech, R., and Singh, J. (2017). Genome-wide Analysis of the Cellulose Synthase-like (Csl) Gene Family in Bread Wheat (*Triticum aestivum* L.). *BMC Plant Biol.* 17, 193–217. doi:10.1186/s12870-017-1142-z

Kumar, S., Tamura, K., and Nei, M. (1994). MEGA: Molecular Evolutionary Genetics Analysis Software for Microcomputers. *Bioinformatics* 10, 189–191. doi:10.1093/bioinformatics/10.2.189

Lee, I. C., Hong, S. W., Whang, S. S., Lim, P. O., Nam, H. G., and Koo, J. C. (2011). Age-dependent Action of an ABA-Inducible Receptor Kinase, RPK1, as a Positive Regulator of Senescence in Arabidopsis Leaves. *Plant Cell. physiology* 52, 651–662. doi:10.1093/pcp/pcr026

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

Letunic, I., and Bork, P. (2018). 20 Years of the SMART Protein Domain Annotation Resource. *Nucleic acids Res.* 46, D493–D496. doi:10.1093/nar/gkx922

Li, J., and Tax, F. E. (2013). *Receptor-like Kinases: Key Regulators of Plant Development and Defense*. Wiley Online Library.

Manès, Y., Gomez, H., Puhl, L., Reynolds, M., Braun, H., and Trethowan, R. (2012). Genetic Yield Gains of the CIMMYT International Semi-arid Wheat Yield Trials from 1994 to 2010. *Crop Sci.* 52, 1543–1552.

Marcussen, T., Sandve, S. R., Heier, L., Spannagl, M., Pfeifer, M., Jakobsen, K. S., et al.International Wheat Genome Sequencing Consortium (2014). Ancient Hybridizations Among the Ancestral Genomes of Bread Wheat. *science* 345, 1250092. doi:10.1126/science.1250092

Milner, M. J., Howells, R. M., Craze, M., Bowden, S., Graham, N., and Wallington, E. J. (2018). A PSTOL-like Gene, TaPSTOL, Controls a Number of Agronomically Important Traits in Wheat. *BMC Plant Biol.* 18, 115–214. doi:10.1186/s12870-018-1331-4

Mishra, D., Suri, G. S., Kaur, G., and Tiwari, M. (2021). Comprehensive Analysis of Structural, Functional, and Evolutionary Dynamics of Leucine Rich Repeats-RLKs in Thinopyrum Elongatum. *Int. J. Biol. Macromol.* 183, 513–527. doi:10.1016/j.ijbiomac.2021.04.137

Mondal, S., Singh, R. P., Huerta-Espino, J., Kehel, Z., and Autrique, E. (2015). Characterization of Heat- and Drought-Stress Tolerance in High-Yielding Spring Wheat. *Crop Sci.* 55, 1552–1562. doi:10.2135/cropsci2014.10.0709

Motte, H., Vercauteren, A., Depuydt, S., Landschoot, S., Geelen, D., Werbrouck, S., et al. (2014). Combining Linkage and Association Mapping Identifies RECEPTOR-LIKE PROTEIN KINASE1 as an Essential Arabidopsis Shoot Regeneration Gene. *Proc. Natl. Acad. Sci. U.S.A.* 111, 8305–8310. doi:10.1073/pnas.1404978111

Nejat, N., and Mantri, N. (2017). Plant Immune System: Crosstalk between Responses to Biotic and Abiotic Stresses the Missing Link in Understanding Plant Defence. *Curr. Issues Mol. Biol.* 23, 1–16. doi:10.21775/cimb.023.001

Nodine, M. D., and Tax, F. E. (2008). Two Receptor-like Kinases Required Together for the Establishment of Arabidopsis Cotyledon Primordia. *Dev. Biol.* 314, 161–170. doi:10.1016/j.ydbio.2007.11.021

Nodine, M. D., Yadegari, R., and Tax, F. E. (2007). RPK1 and TOAD2 Are Two Receptor-like Kinases Redundantly Required for Arabidopsis Embryonic Pattern Formation. *Dev. Cell.* 12, 943–956. doi:10.1016/j.devcel.2007.04.003

Odegard, I. Y. R., and Van Der Voet, E. (2014). The Future of Food - Scenarios and the Effect on Natural Resource Use in Agriculture in 2050. *Ecol. Econ.* 97, 51–59. doi:10.1016/j.ecolecon.2013.10.005

Rasool, F., Uzair, M., Naeem, M. K., Rehman, N., Afroz, A., Shah, H., et al. (2021). Phenylalanine Ammonia-Lyase (PAL) Genes Family in Wheat (*Triticum aestivum* L.): Genome-wide Characterization and Expression Profiling. *Agronomy* 11, 2511. doi:10.3390/agronomy11122511

Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., et al. (2019). g:Profiler: a Web Server for Functional Enrichment Analysis and Conversions of Gene Lists (2019 Update). *Nucleic acids Res.* 47, W191–W198. doi:10.1093/nar/gkz369

Rehman, O. U., Uzair, M., Chao, H., Fiaz, S., Khan, M. R., and Chen, M. (2022). Role of the Type-B Authentic Response Regulator Gene Family in Fragrant Rice under Alkaline-salt Stress. *Physiol. Plant.* 174, e13696. doi:10.1111/ppl.13696

Roser, M., and Ortiz-Ospina, E. (2013). *Global Extreme Poverty*. Our World in Data.

Sahu, S. S., Loaiza, C. D., and Kaundal, R. (2020). Plant-mSubP: a Computational Framework for the Prediction of Single- and Multi-Target Protein Subcellular Localization Using Integrated Machine-Learning Approaches. *AoB Plants* 12, plz068. doi:10.1093/aobpla/plz068

Schmittgen, T. D., and Livak, K. J. (2008). Analyzing Real-Time PCR Data by the Comparative CT Method. *Nat. Protoc.* 3, 1101–1108. doi:10.1038/nprot.2008.73

Shabbir, R. N., Ashraf, M. Y., Waraich, E., Ahmad, R., and Shahbaz, M. (2015). Combined Effects of Drought Stress and NPK Foliar Spray on Growth,

Physiological Processes and Nutrient Uptake in Wheat. *Pak. J. Bot.* 47, 1207–1216.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* 13, 2498–2504. doi:10.1101/gr.1239303

Shi, C.-C., Feng, C.-C., Yang, M.-M., Li, J.-L., Li, X.-X., Zhao, B.-C., et al. (2014). Overexpression of the Receptor-like Protein Kinase Genes *AtRPK1* and *OsRPK1* Reduces the Salt Tolerance of *Arabidopsis thaliana*. *Plant Sci.* 217-218, 63–70. doi:10.1016/j.plantsci.2013.12.002

Shumayla, S., Sharma, S., Kumar, R., Mendu, V., Singh, K., and Upadhyay, S. K. (2016). Genomic Dissection and Expression Profiling Revealed Functional Divergence in *Triticum aestivum* Leucine Rich Repeat Receptor like Kinases (TaLRRKs). *Front. Plant Sci.* 7, 1374. doi:10.3389/fpls.2016.01374

Ulfat, A., Majid, S. A., and Hameed, A. (2017). Hormonal Seed Priming Improves Wheat (*Triticum aestivum* L.) Field Performance under Drought and Non-stress Conditions. *Pak. J. Bot.* 49, 1239–1253.

Uzair, M., Long, H., Zafar, S. A., Patil, S. B., Chun, Y., Li, L., et al. (2021). Narrow Leaf21, Encoding Ribosomal Protein RPS3A, Controls Leaf Development in Rice. *Plant Physiol.* 186, 497–518. doi:10.1093/plphys/kiab075

Wang, Y., Tang, H., Debarry, J. D., Tan, X., Li, J., Wang, X., et al. (2012). MCScanX: a Toolkit for Detection and Evolutionary Analysis of Gene Synteny and Collinearity. *Nucleic acids Res.* 40, e49. doi:10.1093/nar/gkr1293

Wasaya, A., Manzoor, S., Yasir, T. A., Sarwar, N., Mubeen, K., Ismail, I. A., et al. (2021). Evaluation of Fourteen Bread Wheat (*Triticum aestivum* L.) Genotypes by Observing Gas Exchange Parameters, Relative Water and Chlorophyll Content, and Yield Attributes under Drought Stress. *Sustainability* 13, 4799. doi:10.3390/su13094799

Xiang, Z. (2006). Advances in Homology Protein Structure Modeling. *Cpps* 7, 217–227. doi:10.2174/138920306777452312

Yan, F., Li, H., and Zhao, P. (2019). Genome-Wide Identification and Transcriptional Expression of the PAL Gene Family in Common Walnut (*Juglans Regia* L.). *Genes.* 10, 46. doi:10.3390/genes10010046

Yi-Qin, H., Wen-Di, H., Lei, Y., Yi-Ting, L., Chen, L., Yong-Xing, Z., et al. (2020). Genome-wide Analysis of Ethylene-insensitive3 (EIN3/EIL) in *Triticum aestivum*. *Crop Sci.* 60, 2019–2037.

Zargar, S. M., Gupta, N., Nazir, M., Mahajan, R., Malik, F. A., Sofi, N. R., et al. (2017). Impact of Drought on Photosynthesis: Molecular Perspective. *Plant gene.* 11, 154–159. doi:10.1016/j.plgene.2017.04.003

Zou, Y., Liu, X., Wang, Q., Chen, Y., Liu, C., Qiu, Y., et al. (2014). OsRPK1, a Novel Leucine-Rich Repeat Receptor-like Kinase, Negatively Regulates Polar Auxin Transport and Root Development in Rice. *Biochimica Biophysica Acta (BBA) - General Subj.* 1840, 1676–1685. doi:10.1016/j.bbagen.2014.01.003

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.

Copyright © 2022 Rahim, Uzair, Rehman, Rehman, Zahra and Khan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 22 July 2022
doi: 10.3389/fgene.2022.894020

Identification and Characterization of DNA Demethylase Genes and Their Association With Thermal Stress in Wheat (*Triticum aestivum* L.)

Vijay Gahlaut[1,2]*, *Harsha Samtani*[1], *Tinku Gautam*[3] and *Paramjit Khurana*[1]

[1]Department of Plant Molecular Biology, University of Delhi, New Delhi, India, [2]CSIR-Institute of Himalayan Bioresource Technology, Palampur, India, [3]Department of Genetics and Plant Breeding, Chaudhary Charan Singh University, Meerut, India

DNA demethylases (dMTases) are essential proteins in plants that regulate DNA methylation levels. The dMTase genes have been explored in a number of plant species, however, members of this family have not been reported in wheat. We identified 12 wheat dMTase genes divided into two subfamilies: repressor of silencing 1 (ROS1) and DEMETER-Like (DML). The TadMTases in the same subfamily or clade in the phylogenetic tree have similar gene structures, protein motifs, and domains. The promoter sequence contains multiple *cis*-regulatory elements (CREs) that respond to abiotic stress, hormones, and light, suggesting that the majority of *TadMTase* genes play a role in wheat growth, development, and stress response. The nuclear localization signals (NLSs), subcellular localization, and SRR motifs were also analyzed. The expression profile analyses revealed that *TadMTase* genes showed differential gene expression patterns in distinct developmental stages and tissues as well as under heat stress (HS). Furthermore, the qRT-PCR analysis revealed that *TadMTase* gene expression differed amongst wheat cultivars with varying degrees of HS tolerance. Overall, this work contributes to the understanding of the biological function of wheat dMTases and lays the foundation for future investigations.

Keywords: DNA demethylation, bread wheat, heat stress response, cis-regulatory elements, simple sequence repeats (SSR)

OPEN ACCESS

Edited by:
Jianping Wang,
University of Florida, United States

Reviewed by:
Sajid Fiaz,
The University of Haripur, Pakistan
Rahul Kumar,
Chaudhary Charan Singh University, India
Wen-Feng Nie,
Yangzhou University, China

***Correspondence:**
Vijay Gahlaut
zone4vijay@gmail.com

Specialty section:
This article was submitted to Plant Genomics, a section of the journal Frontiers in Genetics

Received: *11 March 2022*
Accepted: *13 June 2022*
Published: *22 July 2022*

Citation:
Gahlaut V, Samtani H, Gautam T and Khurana P (2022) Identification and Characterization of DNA Demethylase Genes and Their Association With Thermal Stress in Wheat (Triticum aestivum L.). Front. Genet. 13:894020. doi: 10.3389/fgene.2022.894020

INTRODUCTION

DNA methylation, along with other epigenetic components (histone modifications, chromatin remodeling, and non-coding small RNAs), regulates the transcription dynamics of several downstream genes and affects the development and stress response in plants (Zhang et al., 2018). DNA methylation at cytosine (C) residues in plants is established through the RNA-directed DNA methylation (RdDM) pathway in three sequence contexts: CG, CHG, and CHH (H stands for A/C/T) (Chan et al., 2005). Four different kinds of methyltransferases have established and maintained C methylation, that is, methyltransferase (MET), chromomethylase (CMT), domain rearranged methyltransferase (DRM), and DNA methyltransferase homologue 2 (DNMT2) (Bender, 2004; Chan et al., 2005; Law and Jacobsen, 2010; Gahlaut et al., 2020).

The homeostasis of DNA methylation also needs the active involvement of DNA demethylase (dMTase) enzymes, which remove C methylation *via* the base excision repair (BER) pathway (Gong et al., 2002; Zhu, 2009; Mok et al., 2010). In Arabidopsis, four dMTase enzymes (DNA glycosylase/

lyases), that is, DEMETER (DME), REPRESSOR OF SILENCING 1 (ROS1), DEMETER-LIKE (DML2) and DML3 have been reported (Choi et al., 2002; Gehring et al., 2006; Penterman et al., 2007; Ortega-Galisteo et al., 2008; Zhu, 2009; Zhang et al., 2018). The ROS1 and DML dMTases are expressed in all vegetative tissues, including root and shoot tissues (Penterman et al., 2007; Calarco et al., 2012), whereas DME is mainly expressed in companion cells of the gametes tissues (Penterman et al., 2007; Huh et al., 2008). In Arabidopsis, it was shown that DME is required for sporophyte development and plays an essential role in the maintenance of stem cell activities during the sporophytic life cycle (Kim et al., 2021). Furthermore, it was demonstrated that dMTases (ROS1 and DME) were also involved in regulating seed development in rice and maize (La et al., 2011; Ono et al., 2012; Kapazoglou et al., 2013). The null mutation of *ROS1a* dMTase in rice causes abnormal early-stage endosperm development and affects the pollen gametophytic transmission (Ono et al., 2012; Kim et al., 2019). Another dMTase encoding gene, *DNG701* in rice, demethylates the *Tos17* and plays a critical role in seed development (La et al., 2011). The active DNA demethylation executed by *OsROS1* also regulates the number of aleurone cell layers in rice (Liu et al., 2018b). Three rice dMTase genes (*DNG702, DNG701,* and *DNG704*) have recently been shown to demethylate DNA at different genomic locations in gametes and zygotes. These findings suggest that active DNA demethylation is crucial for zygotic gene expression and reproduction (Zhou et al., 2021). Recently, it was shown that active DNA demethylation of *RESISTANCE METHYLATED GENE 1* (*RMG1*) and *RECEPTOR-LIKE PROTEINS 43* (*RLP43*) promoters by ROS1 regulates transcriptional immune reprogramming and provides disease resistance in plants (Halter et al., 2021). In barley, a DME glycosylase (*HvDME*) was characterized and the transcriptomic studies during seed development and under dehydration stress revealed that *HvDME* is involved in seed maturation and drought stress response (Kapazoglou et al., 2013). Plants' abiotic stress response regulation also requires active DNA demethylation. For instance, in rice, the salt-sensitive variety IR29 showed induced expression of dMTase genes (*DNG701* and *DNG710*) in response to salinity stress (Ferreira et al., 2015). Similarly, in tomatoes, it was observed that expression of the *DML2* gene is downregulated upon cold stress treatment (Zhang et al., 2016). DNA demethylation was also shown to have a role in the regulation of plant heat stress (HS) responses. For example, in Arabidopsis, it was shown that HS causes a reduction in DNA methylation levels in stress responsive genes (*HSP70*, *RPL26A*, and *POX1*). This demonstrated that the active demethylation process was triggered by HS (Korotko et al., 2021). DNA demethylation has also been linked to seed germination regulation during HS (Malabarba et al., 2021). The above results indicate that active DNA demethylation maintained by dMTases is a crucial process involved in various biological processes, including thermal stress regulation in plants. Apart from this, the role of DNA methylation has also been implicated in the stress-induced formation of the quality-related metabolites in tea (Yang et al., 2021a). Interestingly, the DNA methylation levels of key biosynthesis gene of indole, that is, tryptophan synthase β-subunit 2 were found to be reduced during the continuous wounding stress which occurs during oolong tea processing stage. This promoted the binding of CsMYC2a to the promoter of this gene, which caused the increased accumulation of the aromatic indole compound in the tea (Yang et al., 2021b). Similarly, in the case of Arabidopsis, under heat stress conditions, DNA demethylation has been found to be responsible for seed germination (a process where metabolites accumulated during maturation are used) (Malabarba et al., 2021). Thus, it could be speculated that a similar mechanism might also exist in wheat where DNA methylation affects the quality of wheat seeds by regulating the metabolites.

Wheat (*Triticum aestivum* L.) has been one of the world's most important agricultural crops, accounting for around 30% of global grain output (FAO, 2021). Recently, it was projected that global surface temperatures will rise by 1.5°C in the next 20 years (by 2040), resulting in a major decrease in worldwide wheat productivity (IPPC, 2021). Therefore, to ensure wheat yield, researchers must analyze genomic regions that regulate HS tolerance in wheat. The characterization of wheat dMTases becomes important as their roles in plant development and stress responses are still largely unknown.

In the present study, using the most recent wheat genome sequences, we systematically identified 12 dMTase genes. The chromosome localization, evolutionary relationship, gene structures, conserved domains, motifs, SSR motifs, *cis*-regulatory elements, and subcellular localization were analyzed. Finally, the expression profiles of *TadMTases* were investigated in various plant tissues and development stages, as well as in response to heat stress. Altogether, these findings contribute to our understanding of the structure, phylogeny, and significant regulatory functions of wheat dMTase genes. Furthermore, our findings lay the foundation for possible functional investigation of these genes in order to improve wheat heat stress resistance.

MATERIALS AND METHODS

Sequence Retrieval and Identification of dMTases in Wheat

To obtain wheat sequences of dMTases, we constructed a local protein database, which contains all of the wheat protein sequences accessible on the Ensemble database available at http://plants.ensembl.org/index.html (Howe et al., 2020). Then, the database was examined with known Arabidopsis dMTase proteins (retrieved from the TAIR database, https://www.arabidopsis.org) utilizing the $BLAST_P$ program with an e-value of le-5 and an identity of 50%. The dMTase genes were cross-checked and redundant sequences were eradicated. Using the SMART online software program (http://smart.embl-heidelberg.de/) (Schultz et al., 2000) and the Pfam database (http://pfam.xfam.org/) (El-Gebali et al., 2018), all of the detected wheat dMTase protein sequences were also validated for the presence of conserved domains (HhH-GPD glycosylase domain; PF00730 and the RNA-recognition motif in Demeter;

PF15628). The ExPASy server (http://www.expasy.org/) (Artimo et al., 2012) was used to determine the relative molecular weight (MW) and isoelectric point (PI) of TadMTase proteins.

Chromosomal Distribution, Gene Structure, and Motif Analysis

The *TadMTase* genes were assigned to individual chromosomes using information available in the Ensemble database (http://plants.ensembl.org) (Howe et al., 2020). The chromosomal location of TadMTase genes was drawn using TBtools software (Chen et al., 2020). The gene structures were also analyzed using the TBtools software. The motifs of TadMTase were analyzed using the MEME suite (Bailey et al., 2009), and their positions were displayed using TBtools software.

Multiple Sequence Alignment and Phylogenetic Tree Construction

The dMTase protein sequences of wheat (12), *Aegilops tauschii* (4), *Triticum urartu* (3), *Triticum turgidum* (7), *Oryza sativa* (4), *Arabidopsis thaliana* (4), *Gossypium raimondii* (4), *Gossypium arboretum* (5), *Gossypium barbadense* (6), and *Gossypium hirsutum* (10) were aligned using the MUSCLE package available at MEGA X software (Stecher et al., 2020). The phylogenetic tree was created using the neighbor-joining method with the Poisson model, pair-wise deletion, and 1,000 bootstrap values in MEGA X software. The phylogenetic tree was visualized using the iTOL online tool (Letunic and Bork, 2021). Dated phylogeny trees for 31 plant species were retrieved from TimeTree (Kumar et al., 2017).

Analysis of *Cis*-Regulatory Elements

To analyze the *cis*-regulatory elements (CREs) of TadMTase, the upstream sequences (1,500 bp) of the start codon were fetched from the Ensemble database (Howe et al., 2020). The CREs were identified using the PlantCARE online server (Lescot et al., 2002).

Nuclear Localization Signals, Subcellular Localization, and Gene Ontology Analysis

The NLS in TadMTases was predicted using the cNLS Mapper online tool (Kosugi et al., 2009). Using the CELLO online resource (Yu et al., 2014), the subcellular localization of TadMTases was predicted. The corresponding Ensemble database IDs of TadMTase genes were subjected to the ShinyGO v0.61 database (Ge et al., 2020) to obtain gene ontology (GO) annotation.

Identification of Simple Sequence Repeats and Primer Designing

The WebSat online tool (Martins et al., 2009) was used to mine simple sequence repeats (SSRs) in the TadMTase gene sequences. The minimum length criteria of six for di-nucleotide repeats and three for tri-nucleotide repeats, tetra-nucleotide repeats, penta-nucleotide repeats, and hexa-nucleotide repeats. The Primer 3 software (available on the WebSat online tool) was used to design primers. The SSR primers were developed using the following criteria: melting temperature (Tm) of 55°C–60°C, primer length of 20–25 bp, and product size ranging from 110 to 400 bp.

Expression Profiling of DNA Demethylase Genes

The expression levels of *TadMTase* genes in different tissues, developmental stages, and responses to heat stress were analyzed using the Genevestigator tool (Hruz et al., 2008). The Genevestigator has a broad collection of public microarrays and RNA-Seq study data.

Plant Materials and Heat Stress Treatment

Expression profiling of *TadMTase* genes was examined in two wheat cvs. HD2329 (heat-sensitive) and HD2985 (heat-tolerant) under control and heat stress conditions. Seeds of two cultivars were surface-sterilized with 1% hydrogen peroxide, washed gently with distilled water, and germinated for 2 days in Petri plates on water-soaked filter paper at 22°C. Then the seedlings were transferred and cultivated in half-strength Hoagland nutrient solution at 22°C with a 16:8 h light/dark photoperiod. HS was induced by exposing one-week-old seedlings to a temperature of 42°C for 2 h (Meena et al., 2022). The environment for the control sample was set to 22°C. Each treatment (control and HS treated) included three biological replicates. All samples were promptly frozen using liquid nitrogen and preserved at −80°C for RNA extraction.

RNA Isolation and Quantitative Real-Time PCR

Total RNA was extracted from 100 mg tissues using TRizol reagent (Ambion), and DNA was removed by DNaseI enzyme (TaKaRa, United States). cDNA was generated using the RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific). Quantitative expression was performed with a QuantStudio Real-Time PCR system (Applied Biosystems, United States). Three technical replicates for each sample were used. The details of gene-specific primers are provided in **Supplementary Table S1**. The wheat *GAPDH* gene was used as an internal reference gene and the relative gene expression was estimated using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001).

RESULTS

Identification of dMTase Genes in Wheat

After examining the wheat reference genome, a total of 12 full-length dMTase genes were identified. The nomenclature of 12 wheat dMTase genes was based on the corresponding genes reported for rice and Arabidopsis. The 12 *TadMTase* genes were mainly divided into two groups (ROS1 and DML). The *TaROS1a* (5A/5B/5D) genes were homologous to *AtROS1* (AT2G36490)

TABLE 1 | Summary and characteristics of dMTase genes identified in wheat.

Gene name	Gene ID[a]	Chromosome location	Transcript length (bp)	No. of exons	Protein			Subcellular localization
					Length (aa)	pI	Mw (kDa)	
TaROS1a-5A	TraesCS5A02G169000	5A:360071286-360084523	6,607	18	1,975	6.0	217.3	Nuclear
TaROS1a-5B	TraesCS5B02G165800	5B:307397203-307412033	7,176	18	1,982	6.1	218.4	Nuclear
TaROS1a-5D	TraesCS5D02G173300	5D:270261993-270276412	7,285	18	1,981	6.1	218.0	Nuclear
TaROS1c1-1A	TraesCS1A02G349600	1A: 534535188-534547681	5,232	20	1,595	6.8	177.8	Nuclear
TaROS1c1-1B	TraesCS1B02G364100	1B: 593446994-593459819	5,525	19	1,600	6.5	178.0	Nuclear
TaROS1c1-1D	TraesCS1D02G352500	1D: 437848604-437860415	5,770	19	1,586	6.8	176.1	Nuclear
TaROS1c2-1A	TraesCS1A02G278000	1A:473630543-473641488	5,775	21	1,765	7.4	195.4	Nuclear
TaROS1c2-1B	TraesCS1B02G286900	1B:498862239-498872297	5,740	22	1,769	7.0	196.3	Nuclear
TaROS1c2-1D	TraesCS1D02G277100	1D:374039261-374049890	5,497	19	1,734	6.7	192.3	Nuclear
TaDML3a-3A	TraesCS3A02G022500	3A:12776134-12784267	3,429	17	989	6.9	109.6	Nuclear
TaDML3a-3B	TraesCS3B02G023200	3B:9931500-9939414	3,490	17	1,000	6.8	110.7	Nuclear
TaDML3a-3D	TraesCS3D02G024100	3D:7605843-7613892	3,003	17	1,000	6.9	110.6	Nuclear

[a]*ID available at Ensemble plant database (https://plants.ensembl.org/index.html).*

and *OsROS1a* (Os01t0218032) in Arabidopsis and rice respectively; the *TaROS1c1* (1A/1B/1D) and *TaROS1c2* (1A/1B/1D) genes were homologous to *AtROS1* (AT2G36490) and *OsROS1c* (Os05t0445900) in Arabidopsis and rice respectively; the *TaDML3a* (3A/3B/3D) genes were homologous to *AtDML3* (AT4G34060) and *OsDML3a* (Os02t0496500) in Arabidopsis and rice respectively (see **Supplementary Table S2**). The length of the transcript of *TadMTase* genes ranged from 3,003 bp (*TaDML3a-3D*) to 7,285 bp (*TaROS1a-5D*), and the average length was found to be 5,377 bp. The protein length ranged from 989 aa (*TaDML3a-3A*) to 1982 aa (*TaROS1a-5B*), and the average length was 1,581 aa. The estimated molecular weights of each TadMTase protein varied from 109.6 to 218.4 kDa (**Table 1**). The dMTase gene names, gene IDs, chromosomal location and other features are summarized in **Table 1**.

Chromosome Locations of Wheat dMTases

The 12 wheat dMTases were located on nine (1A, 1B, 1D, 3A, 3B, 3D, 5A, 5B, and 5D) of the 21 wheat chromosomes (**Supplementary Figure S1**). Furthermore, a maximum of six genes were located on group 1 chromosomes (1A/1B/1D), whereas a minimum of three genes were mapped onto group 5 and 3 chromosomes. Since bread wheat is a hexaploid species (i.e., it contains three sub-genomes). Three homologues for each gene were found corresponding to chromosomes A, B, and D (**Table 1**; **Supplementary Figure S1**).

Phylogenetic Relationship of dMTases

To determine the phylogenetic relatedness of dMTases in plants, 59 dMTase protein sequences from five monocots (*Triticum aestivum*, *Aegilops tauschii, Triticum urartu, Triticum turgidum, and O. sativa*) and five dicotyledons (*Arabidopsis thaliana*, *Gossypium hirsutum*, *Gossypium arboretum*, *Gossypium raimondii, and Gossypium barbadense*) were used for multiple sequence alignments and to build a phylogenetic tree (**Figure 1A**). All the members of dMTase are grouped into three subfamilies/clades: ROS1, DME, and DML. The ROS1 clade comprised 22 monocots and nine dicots, the DML clade comprised eight members of monocots and 13 dicots, while the DME comprised seven members of dicots only (**Figure 1A**). Furthermore, we also constructed an evolutionary tree of life for 31 plant species belonging to monocot (8) and dicot (23) and showed the total number of dMTase proteins for each species (**Figure 1B**). Our analysis revealed that the DME subfamily of dMTases was absent in monocots and existed only in dicots.

Gene Structure, Motif and Domain Composition of TadMTases

To gain further insights into the *TadMTase* genes, we surveyed the gene structure, conserved domain, and motif components of each *TadMTase* gene. The number of exons in the *TadMTase* genes ranged from 17 to 22 (**Figure 2A**; **Table 1**). We found that the number of exons in *dMTase* genes from the same homoeologous group did not differ much (maximum up to three exons in *TaROS1c2*). Among these, the dMTase gene with the maximum number (22) of exons was found to be *TaROSc2-1B*, while *TaDML3a* had the least number of exons (17).

The distribution of conserved motifs in the TadMTase proteins was predicted using the MEME online suite. Twenty different types of motifs, which were designated 1 to 20, were predicted in each TadMTase protein. The amino acid length of motifs varied from 21 to 50. The number of motifs in each dMTase varied from 13 to 20. Among them, motifs 3 and 9 were part of Domain A, motifs 1, 2, 6, 8, and 13 were part of the HhH-GPD glycosylase domain, and motifs 4, 5, 7, 10, 11, and 19 were part of the RRM_DME domain (**Figure 2B**). Furthermore, the Motifs 1, 2, 3, 4, 5, 10, 11, 13, 16, 18, and 19 were highly conserved and found in every dMTase protein. However, some motifs (9, 15, and 20) were specific to the ROS gene group (**Figure 2B**). It could

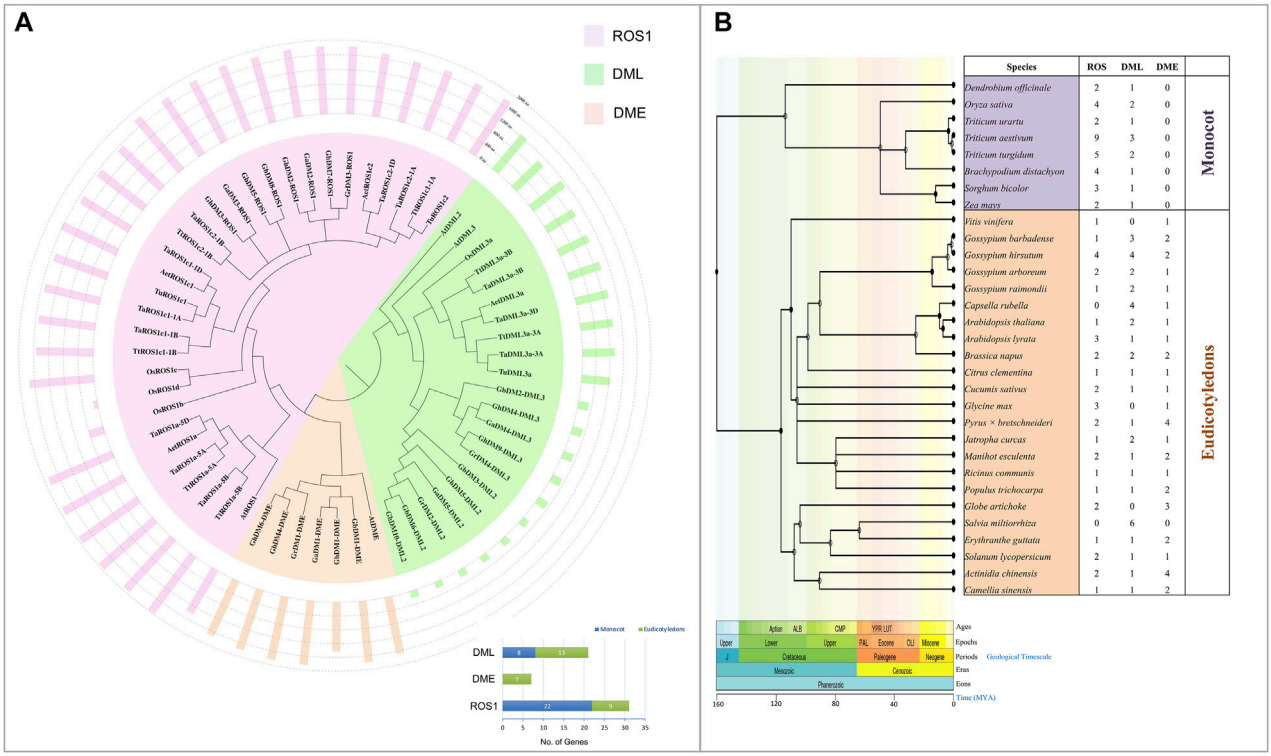


FIGURE 1 | Evolutionary relationship among various dMTases. **(A)** Evolution relationship among the 10 plant species (*Ta*, *Triticum aestivum*; *Aet*, *Aegilops tauschii*; *Tu*, *Triticum Urartu*; *Tt*, *Triticum turgidum*; *At*, *Arabidopsis thaliana*; *Os*, *Oryza sativa*; *Ga*, *Gossypium arboretum*; *Gb*, *Gossypium barbadense*; *Gh*, *Gossypium hirsutum*; *Gr*, *Gossypium raimondii*). Graph (present at the lower right corner) shows the number of dMTases in each sub-family (ROS1, DML, and DME). **(B)** Dated phylogeny trees for 31 plant species belonging to monocot and eudicotyledons. The number of genes present/absent of ROS1, DML, and DME in different plant species is shown in the columns. MYA, a million years ago.

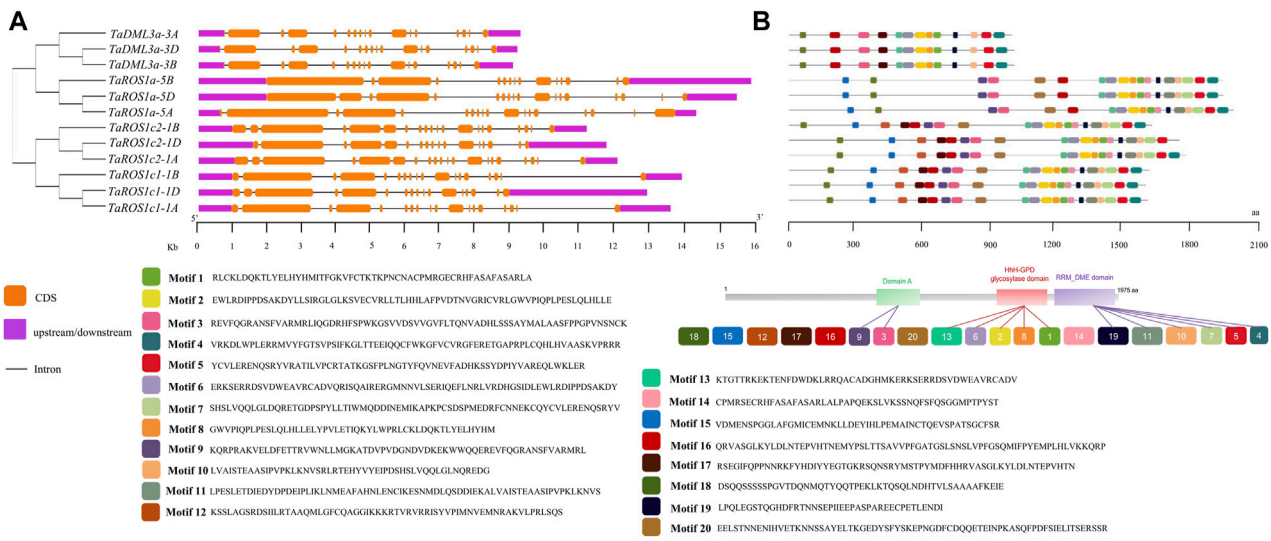


FIGURE 2 | Exon-intron structures and conserved motif compositions of TadMTases. **(A)** Exon–intron structures of *TadMTase* genes. Exons are shown as orange boxes, introns are denoted by thin dark grey lines, and upstream/downstream regions are shown as purple boxes. The lengths of exons and introns can be determined using the scale bar on the bottom. **(B)** Motifs organization of TaDMTase proteins. Color boxes represent the position of different motifs, and box sizes show the length of motifs. Sequences for 20 different motifs were also provided on the right side.



FIGURE 3 | Conserved domain composition and multiple sequence alignments of the conserved domain sequences of *TaDMTase* proteins. **(A)** Three different conserved domains (Domain A, HhH-GPD glycosylase domain and RRM_DME domain) are shown in different colored boxes. **(B)** Multiple sequence alignment analysis of the conserved domain sequences of 12 wheat dMTase protein sequences and an Arabidopsis dMTase protein. The key conserved domains are highlighted by double lines (Domain A by green color, HhH-GPD glycosylase domain by red color and RRM_DME domain by brown color) and the domain names are shown at the top of the sequence. Gene names are shown on the left.

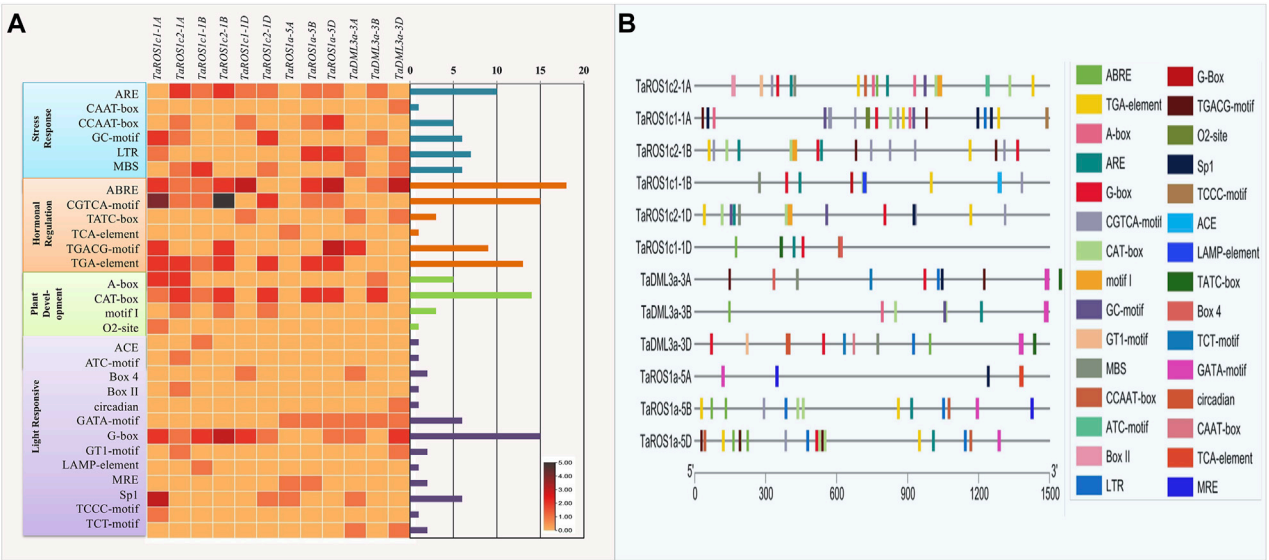


FIGURE 4 | Cis-regulatory elements (CREs) of wheat dMTases genes. **(A)** Showing the number of CREs belonging to the following four categories (stress-responsive, hormonal regulation, plant development, and light-responsive) per dMTase gene as a heatmap. The bar plots on the right represent the number of CREs for each dMTase gene. **(B)** The distribution of different CREs in the promoter region (1,500 base pairs) of wheat dMTase genes. Different CREs are represented by distinct colors as indicated.

be inferred that motifs found in all dMTase proteins are possibly associated with conserved functions, but those specific to the ROS gene group may be involved in gene-specific functions.

To further determine the structural characteristics of the wheat dMTase family members, conserved domains were also identified. All the wheat dMTases contained three different domains, including 1) Domain A, 2) HhH-GPD glycosylase domain (PF00730) and 3) Domain B or RRM_DME (RNA-recognition motif in Demeter; PF15628) domain (**Figure 3**). The domain A contains the stretch of basic amino acids (K and R) that is required for DNA binding activity. The HhH-GPD glycosylase domain contains a helix–hairpin–helix (HhH) motif, a glycine/proline-rich loop with a conserved aspartic acid (GPD), and four cysteine residues in the Fe-S cluster, these are involved in 5-methylcytosine excision activity. The RRM_DME domain facilitates the interaction of the catalytic domain with ssDNA or regulatory RNA. The multiple sequence alignments of wheat dMTases showed that these three domains are conserved (**Supplementary Figure S2**).

SSR Motifs and Markers' Development

Out of 12 wheat *dMTase* gene sequences screened for SSRs, ten genes have 17 SSR motifs. The tri-nucleotide repeats motif were found most abundant, comprising about nine (53%) followed by the tetra-nucleotide repeat motifs comprising four (23.5%), whereas, di- and penta-nucleotides SSR formed a small share having 2 (11%) each (**Supplementary Table S3**). The number of repetitions of a motif ranged from 3 to 19. For each of the 17 SSR motifs, primer pairs were also designed. The following details, such as nucleotide sequence, melting temperature, and product size related to SSR motif primers, are provided in **Supplementary Table S3**.

NLS Prediction and Sub-Cellular Localization

All TadMTase members contain both monopartite and bipartite NLSs (except *TaROS1a-5A*). In total, 12 dMTases were predicted with 35 NLS (21 mono and 14 bipartite NLSs) (**Supplementary Table S4**). Subcellular localization prediction indicated that all the 12 wheat dMTase proteins are localized in the nucleus (**Table 1**).

Cis-Regulatory Elements in the Promoter of TadMTases

To further understand the possible regulatory function of dMTases in plant growth and stress response, we searched for CREs in the 1,500 bp promoter region of wheat *dMTase* genes. A total of 158, representing 29 types of CREs, were predicted (**Figure 4**). Most of the identified CREs were hormone response factors (37.34%), followed by light-responsive elements (25.94%), stress response elements (22.15%), and plant development-related elements (14.55%). The CREs involved in hormonal signaling comprised auxin-responsive (TGA-element), abscisic acid-responsive (ABRE), methyl jasmonate responsive (CGTCA and TGACG motif), gibberellins responsive (TATC-motif) and salicylic acid-responsive (TCA-element) elements. Among these CREs, the CGTCA, ABRE, TGACG, and TGA motifs appeared to be the most prevalent and were detected in all TadMTase genes. In

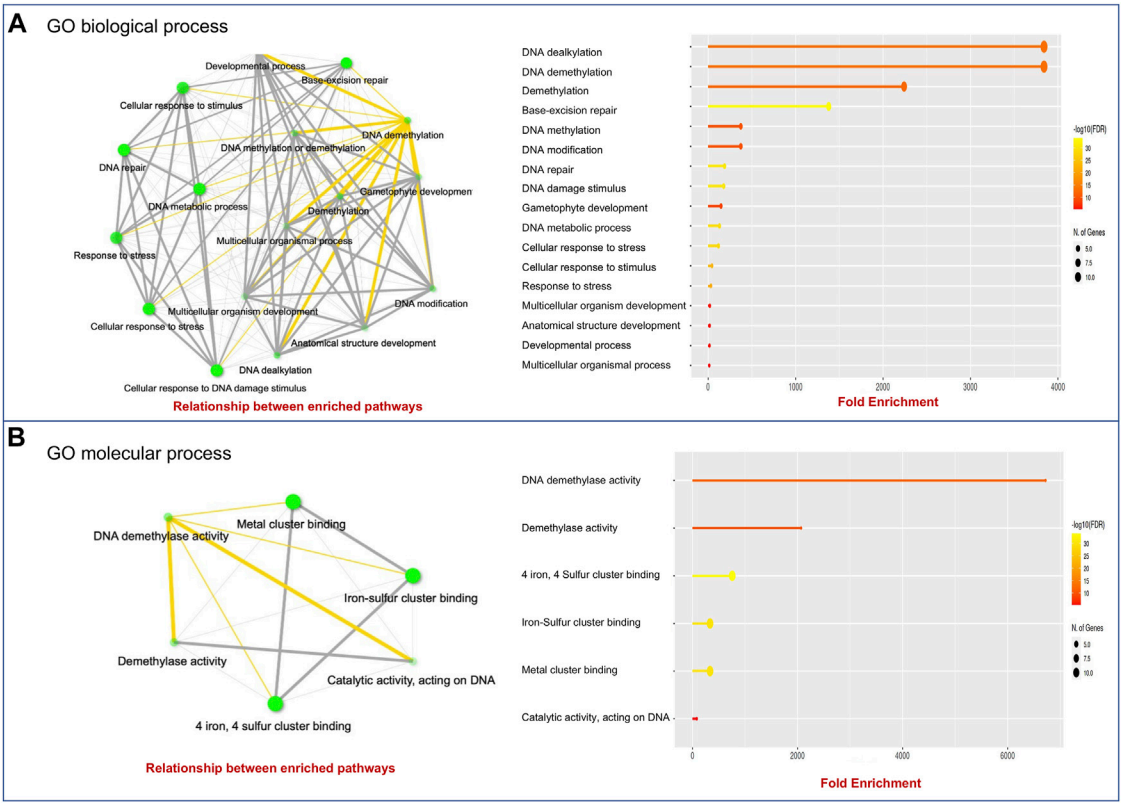


FIGURE 5 | GO analysis using ShinyGo identified enriched biological processes **(A)** and molecular processes **(B)** of dMTase genes. The green dots represent the nodes for each GO biological process or molecular process, while the lines (yellow and grey) represent the interaction between the nodes (minimum of 20% genes common between two connected) GO processes.

TadMTase promoters, all 13 kinds of light-responsive CREs were also found. Out of those, G-box CREs appeared to be the most prominent. Each TadMTase gene had one to three G-box copies (**Figure 4A**).

Thirty-five stress-responsive CREs were found in *TadMTase* genes and out of those, ARE (anaerobic responsive elements) elements comprised 28% of them (10). AREs were most commonly detected at the TadMTase gene promoters. The CAAT-box, CCAAT-box, GC-motif, LTR (low temperature responsive) and MBS were among the other stress-related CREs detected. The GC-motif and CAAT-box have been linked to a variety of abiotic stressors. MYB binding sites (MBS) are known to control drought stress, ARE are known to regulate anaerobic induction under oxidative stress, and LTRs regulate the temperature response.

Furthermore, 24 CREs were discovered that were involved in the growth and development of plants. These included A-Box (5), CAT-box (14), motif I (3), and O2 site (1) (**Figure 4A**). The CAT-box element appeared to be the most frequent and was detected in ROS and DML dMTase genes. We also found Motif-I, a root-specific regulatory element only in the ROS sub-group genes. Overall, these findings reveal that TadMTase genes may be involved in a variety of abiotic stress responses as well as plant growth regulation.

Gene Ontology Annotation of TadMTases

To further characterize the function of the wheat *dMTase* genes, we performed GO enrichment analysis using the ShinyGO tool (Ge et al., 2020). GO enrichment analysis showed that the *TadMTase* genes were involved in DNA demethylation, the BER pathway, DNA modification, developmental processes, and stress responses in GO biological processes (**Figure 5A**). The GO molecular process showed that *TadMTase* genes were involved in DNA demethylase activity, metal/Iron-sulphur cluster binding and catalytic activity (**Figure 5B**). This annotation indicates that the *TadMTase* genes may be involved in stress regulation in plants *via* DNA demethylation.

Developmental Stage and Tissue-Specific Expression Patterns of TadMTases

To decipher the function of wheat *dMTase* genes, the expression pattern in ten developmental stages and 21 tissues was analyzed (**Figure 6**). The expression levels indicate that all the wheat *dMTase* genes were expressed in different developmental stages/tissues but varied significantly. During various developmental stages, the expression of dMTases gradually increased from the seedling to the booting stage and declined continuously till the ripening stages. However, *TaROS1a-5A/5B/*

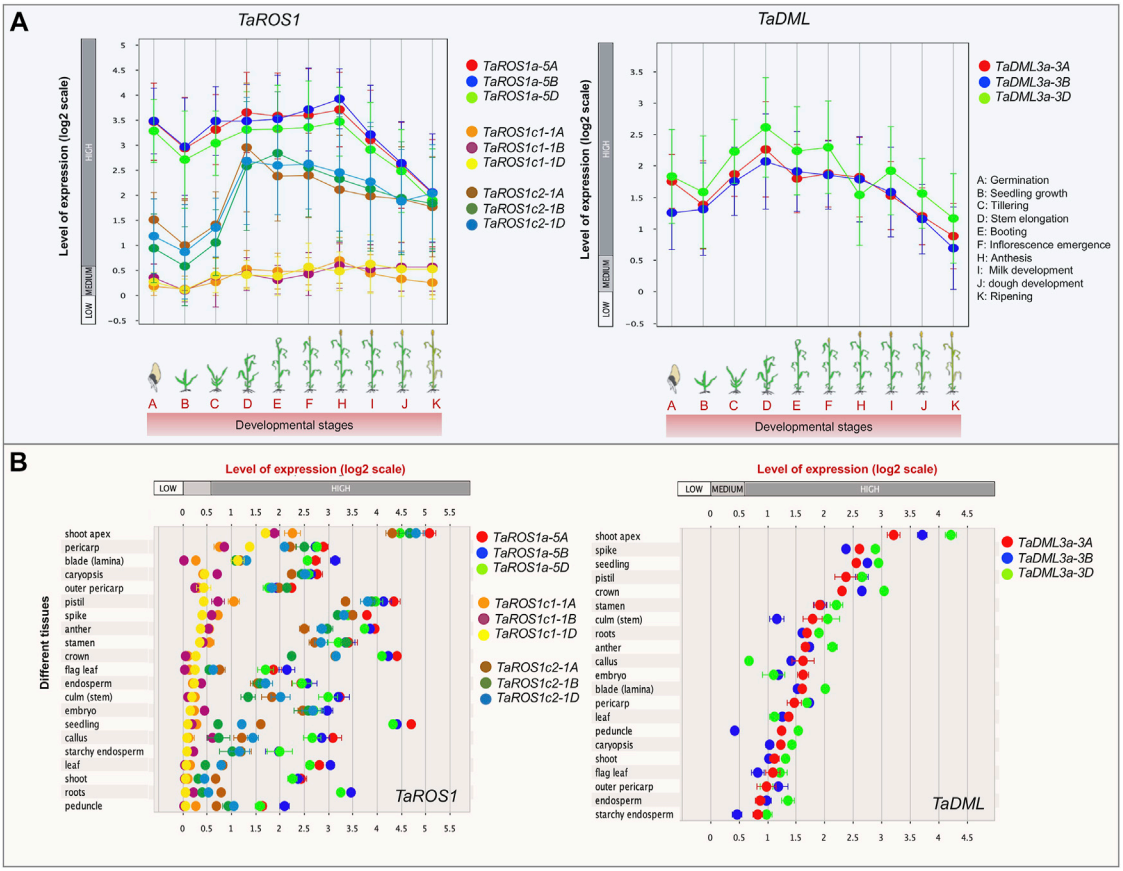


FIGURE 6 | Expression profiles of wheat dMTase in the different developmental stages **(A)** and wheat tissues **(B)**. Showing the scatterplots for the 12 *TadMTase* genes, the left side in **(A)** and the topside in **(B)** represent the change of gene expression using the base 2 logarithm scale. Different dMTase genes are represented by distinct colors as indicated. Error bars represent standard errors. Data were analyzed with the Genevestigator tool.

5D had the highest expression at anthesis stages, and *TaROS1c1/2 1A/1B/1D* and *TaDML3a 3A/3B/3D* had the highest expression at the stem elongation stage (**Figure 6A**). In the case of tissue-specific expression, *TaROS1a 5A/5B/5D*, *TaROS1c2 1A/1B/1D* and *TaDML3a 3A/3B/3D* exhibited relatively higher expression in all tissues. However, *TaROS1c1 1A/1B/1D* had tissue-specific expression and was expressed especially in the shoot apex, pericarp, and blade (lamina) tissues (**Figure 6B**). Furthermore, we also studied the expression of *TadMTase* genes in hormone-treated [abscisic acid (ABA) and gibberellic acid (GA3)] wheat tissues. *TaROS1a-5A/5B/5D* were downregulated (up to 1.96 fold) in root tissues at the seedling stage upon ABA treatment. Other genes do not show any significant change in their expression pattern (**Supplementary Figure S3**).

Expression Patterns of TadMTases Under Heat Stress

Heat stress is one of the major stresses that severely affects wheat productivity. Therefore, there has been recent research to understand the mechanism by which plants respond to HS and identify the genes associated with HS tolerance (Samtani et al., 2022). To investigate the putative biological role of *TadMTase* under HS, the expression pattern of these genes under HS was determined using wheat expression data available in the Genevestigator tool and by performing the qRT-PCR experiments. The wheat RNA-seq data (available on the Genevestigator tool) of the four different experiments when were subjected to ambient (22°C) and heat stress (37°C–40°C) treatment for 5 min to 5 days (as shown in **Figure 7A**) were evaluated. Results showed that there were significant differential expression patterns and most of the *TadMTase* genes were suppressed (up to 2.24 folds) in response to HS as compared to controls (**Figure 7A**). For instance, in response to HS, *TaROS1a-5A, -5B,* and *-5D*, and *TaROS1c2-1A*, *-1B,* and *-1D* were downregulated in the seedling stage (leaf and caryopsis tissue; up to ~2.3 fold) and *TaDML3a-3A*, *-3B,* and *-3D* were downregulated in the anthesis stage (flag leaf tissue; up to ~2 fold).

Furthermore, we selected five genes (*TaROS1a-5A*, *TaROS1c1-1B*, *TaROS1c2-1A*, *TaDML3a-3B*, *TaDML3a-3D*) out of the 12 *TadMTase* genes for quantitative real-time PCR analysis. For this purpose, contrasting wheat genotypes (sensitive cv. HD2329 and tolerant cv. HD2985) were used. All the five

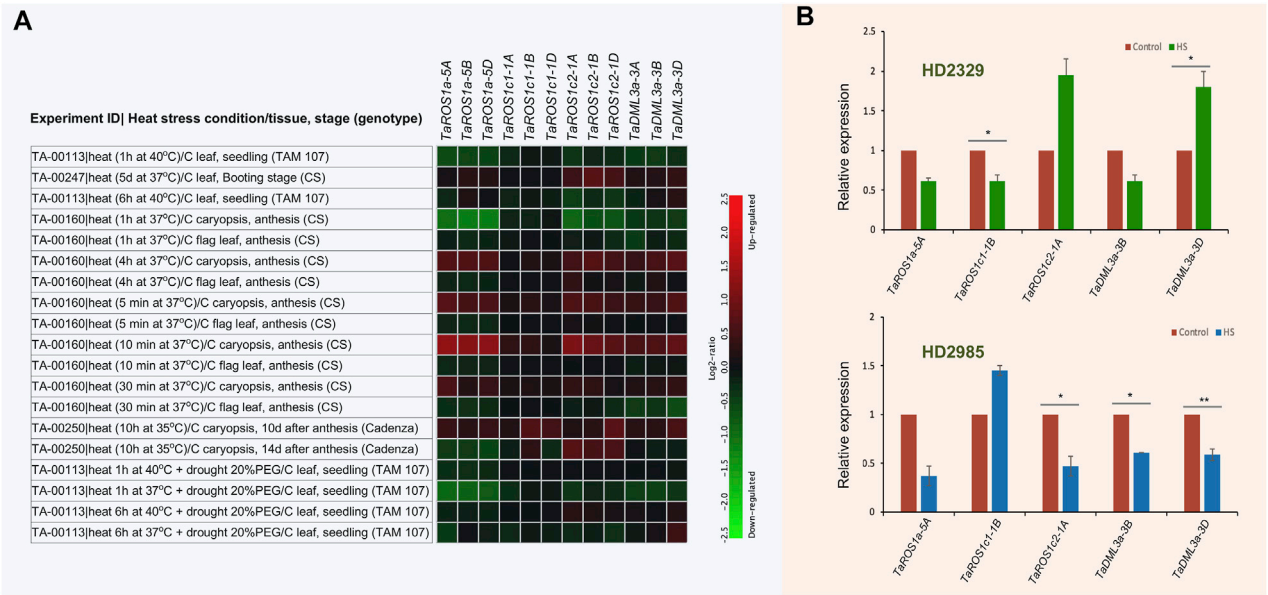


FIGURE 7 | Expression profiles of wheat *dMTase* genes under heat stress (HS) conditions. **(A)** Heatmap showing the expression of 12 *TadMTase* genes in various tissues and developmental stages during heat stress conditions. The "heatmap" color represents relative expression values, calculated as log2 ratios between the signal intensities from HS treated genotypes vs. controls (C). The red color indicates up-regulation and the green color indicates downregulation. **(B)** The relative expression level of five *TadMTase* (*TaROS1a-5A, TaROS1c1-1B, TaROS1c2-1A, TaDML3a-3B, TaDML3a-3D*) genes was analyzed by quantitative RT-PCR in wheat seedling tissue under HS treatments (42°C for 2 h). Each bar value represent the means (±SE) derived from three biological. Asterisks on the bar showed significant differences between HS treated and control samples (*$p < 0.05$, **$p < 0.01$, Student's *t*-test). CS, Chinese spring; PEG, poly (ethylene glycol).

TadMTase genes showed considerable differential expression after being stimulated by HS (**Figure 7B**). The expression levels of *TaROS1a-5A* and *TaDML3a-3B* were repressed (up to 2.7-fold) in both the genotypes. The expression levels of *TaROS1c2-1A* and *TaDML3a-3D* were repressed (up to 2.12-fold) in HS tolerant cv. HD2985 and upregulated in HS sensitive cv. HD2329. The expression level of *TaROS1c1-1B* was upregulated in HS tolerant cv. HD2985 and repressed in HS sensitive cv. HD2329 (**Figure 7B**). The above results indicate that active demethylation is possibly involved in wheat's responses to thermal stress. Furthermore, the differentially regulated *TadMTase* might regulate thermotolerance in crops such as wheat.

DISCUSSION

DNA methylation is an important epigenetic mark and assists plants to adapt to different abiotic stress conditions (Lämke and Bäurle, 2017; Korotko et al., 2021). DNA methylation negatively controls the transcription of genes and the transposition of transposable elements (TEs) in plants (Matzke and Mosher, 2014). Furthermore, the DNA methylation state on specific sites depends on the following three activities: 1) establishment, 2) maintenance, and 3) active DNA demethylation. In plants, active DNA demethylation is performed by DNA demethylases genes (*ROS1, DML, and DME*) *via* the BER pathway (Zhu, 2009; Mok et al., 2010; Zhang et al., 2018). The model plant Arabidopsis encodes four different dMTases, namely *ROS1*, *DME*, *DML2*, and *DML3* (Zhang et al., 2018). The *dMTase* gene family has been detected and characterized in the following plant species, such as *Solanum lycopersicum* (Cao et al., 2014), *Arachis hypogaea* (Wang et al., 2016), *Fragaria vesca* (Gu et al., 2016), *Cynara cardunculus* var. *scolymus* (Gianoglio et al., 2017), *Pyrus bretschneideri* (Liu et al., 2018a), *Ricinus communis* (Victoria et al., 2018), cotton (*Gossypium hirsutum, Gossypium arboretum, Gossypium raimondii,* and *Gossypium barbadebse*) (Yang et al., 2019) *Actinidia chinensis* (Zhang et al., 2020), *Camellia Sinensis* (Zhu et al., 2020), *Dendrobium officinale* (Yu et al., 2021). However, no comprehensive identification and characterization of these genes in wheat has been reported. Here, by utilizing the wheat reference genome sequence data (available at https://plants.ensembl.org/index.html), we have identified 12 *dMTases* (9 *ROS1* and 3 *DML*) genes in the wheat (**Table 1**). The phylogenetic relationships, gene structures, CREs, conserved domains and motifs, SSR motifs, and expression profiling in various development stages/tissues and during HS for TadMTases were also analyzed.

As per our findings, wheat has a higher number of dMTases than some other plant species. In different plant species, it varies from 2 (*Vitis vinifera*) to 10 (*Gossypium hirsutum*). This disparity might be attributable to wheat's allohexaploid nature. Based on the corresponding genes identified for Arabidopsis and rice, as well as phylogenetic analyses, these dMTase genes were divided into two groups (ROS1 and DML). The DME group was not found in wheat and other monocots (**Figure 1B**), and earlier studies have also discovered that DME was only detected in eudicots and absent in monocots (Choi et al., 2002; Zemach et al., 2010; Yu et al., 2021). The role of DME genes in monocots must be supplemented by ROS1 or DML orthologs. Furthermore, the

absence of DME genes in wheat and in other monocots suggests that DMEs are a recently developed type of DNA demethylase gene in dicots.

The phylogenetic analysis showed that 59 dMTase proteins belonging to 10 species were grouped into three distinct subfamilies, that is, ROS1, DME, and DML (**Figure 1A**). Within each subfamily, wheat dMTases showed a closer phylogenetic relatedness to a tier of wild relative species. For example, *TaROS1a-5a* and *TtROS1-5A* (*Triticum turgidum*) and *TaDML3a-3A* and *TuDML3a* (*Triticum Urartu*) had a close evolutionary relationship (**Figure 1A**). We also observed that in the ROS1 and DML subfamilies, monocotyledons and dicotyledons formed distinct subgroups. The DME subfamily was only found in dicots and absent in monocots, which suggests the loss of DME genes during monocot evolution (**Figure 1B**). Furthermore, wheat dMTases had similar exon-intron, motif, and domain organization throughout each subfamily (**Figures 2**, **3**). Exon-intron, motif, and domain organizations were discovered to be varied amongst the sub-families. Similar observations for dMTase genes were found in several other plant species (Wang et al., 2016; Zhang et al., 2020; Zhu et al., 2020; Yu et al., 2021).

NLS analysis has shown that most of the TadMTases have NLS domains (**Supplementary Table S4**), which are required for the precise targeting of these proteins to the nucleus. The sub-cellular localization studies also found that wheat dMTases were positioned in the nucleus (**Table 1**). The dMTases in the following plant species, that is, Arabidopsis (Gong et al., 2002), tomato (Cao et al., 2014), peanut (Wang et al., 2016), globe artichoke (Gianoglio et al., 2017), castor bean (Victoria et al., 2018), eggplant (Moglia et al., 2019), tea plant (Zhu et al., 2020), and orchid (Yu et al., 2021) were also shown to be localized in the nucleus. Considering the importance of subcellular location in defining a protein's function, these results indicate that the function of *dMTase* genes in plants may be conserved.

Gene-based SSR markers are a highly valuable resource and could be used for functional diversity analysis, comparative mapping, evolutionary analysis, and molecular plant breeding (Varshney et al., 2005). Gene-based SSR markers were developed in several crop plants like rice (Molla et al., 2015), barley (Zhang et al., 2014), wheat (Kumar et al., 2018; Singh et al., 2018; Mehta et al., 2021), millet (Desai et al., 2021) and many others. For wheat *dMTase* genes, we also identified 17 SSR motifs and designed primer pairs (**Supplementary Table S3**). These SSR markers may be examined in different wheat genotypes for polymorphism, and the polymorphic SSRs may be utilized for molecular breeding (i.e., marker-assisted selection, marker-assisted recurrent selection) in wheat breeding program for the improvement of different agronomic traits and for heat stress tolerance.

Cis-regulatory elements (CREs) are known to play a role in the modulation of a variety of biological processes, including abiotic stresses (Azodi et al., 2020). We also found abiotic stress-responsive, hormonal regulation related, light-responsive, and plant development-related CREs in the promoter of *TadMTase* genes in this investigation (**Figure 4**). Similar types of CREs have been reported in the promoter regions of dMTase genes in other plants. For instance, CAAT, G-box, TCA, ABRE elements, etc., were reported in the promoter of dMTase genes belonging to *Arachis hypogaea* (Wang et al., 2016), *Actinidia chinensis* (Zhang et al., 2020), *Camellia Sinensis* (Zhu et al., 2020), and *Dendrobium officinale* (Yu et al., 2021). According to these findings, wheat dMTase genes may have a role in the development and stress response regulation in plants. Furthermore, in wheat, we also observed that the number of CREs and their distribution patterns differ between the promoters of three wheat homoeologous genes (**Figure 4**). On the other hand, their encoding proteins had similar domains and amino acid sequences. This highlighted the homologous genes' distinct regulatory activities during wheat polyploidization.

DNA methylation regulates an array of biological processes that control plant growth and development as well as biotic and abiotic stress responses (Bender, 2004; Zhang et al., 2018). DNA methylation maintenance also requires the active participation of dMTases, which remove C methylation (Mok et al., 2010). Furthermore, several studies have revealed that dMTases are intricate in the regulation of growth and development, either directly or indirectly. For instance, *DML3* was reported to be expressed during leaf senescence in Arabidopsis and regulates the leaf senescence process (Yuan et al., 2020). The *OsROS1a* knock-in null mutation affects the endosperm development and causes the formation of irregular embryos (Ono et al., 2012). DNA demethylation mediated by *OsROS1a* can alter the number of aleurone layers (Liu et al., 2018b). Herein, in wheat, we also found that dMTase genes exhibited significantly differential expression in different tissues and developmental stages (**Figure 6**). For instance, TaROS1a had the highest expression at the anthesis stages, and TaROS1c and *TaDML3a* had the highest expression at the stem elongation stage (**Figure 6A**). In tissue-specific expression analysis, it was observed that *TaROS1a* and *TaDML3a* had relatively higher expression in all tissues. Similar results were also noticed in other plant species. For instance, ROS2 like in peanut (Wang et al., 2016), *SmelDemethylase_5* in eggplant (Moglia et al., 2019) and *DoDML3* in *Dendrobium officinale* (Yu et al., 2021), had relatively higher expression in all the tissues. However, *TaROS1c1* had tissue-specific expression and was expressed especially in the shoot apex, pericarp, and blade (lamina) tissues (**Figure 6B**).

It was reported that the active demethylation phenomenon was involved in HS regulation in plants (Naydenov et al., 2015; Korotko et al., 2021; Malabarba et al., 2021). In previous studies on strawberry (Gu et al., 2016) and cotton (Yang et al., 2019), demethylase gene expression was found to be modulated in response to HS, implying that it may have a role in the plant's response to HS. We analyzed the expression profiles of wheat *dMTase* genes under HS conditions. Our results revealed that there are significant differential expression patterns of *TadMTase* genes and most of them were suppressed in response to HS as compared to controls. These findings were then confirmed by using the qRT-PCR method for the following five genes (*TaROS1a-5A, TaROS1c1-1B, TaROS1c2-1A, TaDML3a-3B, TaDML3a-3D*). During the HS condition, all the five *TadMTase* genes showed considerable differential expression in response to HS (**Figure 7B**). For example, *TaROS1a* and *TaDML3a* expression levels were suppressed in both genotypes. The expression levels of *TaROS1c2 and TaDML3a* were reduced in HS tolerant cv. HD2985 and induced in HS sensitive cv. HD2329 (**Figure 7B**). Interestingly, these genes also

contain stress-responsive CREs in their promoter region, that is, ARE, CAAT-box, CCAAT-box, GC-motif, LTR, and MBS, which are associated with various abiotic stresses, including HS (**Figure 4**). These findings suggest that *TadMTase* genes may have a variety of roles in plants, including plant growth and development as well as HS control. However, further functional research is required to corroborate the findings of this study.

CONCLUSION

A comprehensive and systematic analysis was conducted, and 12 dMTase genes were identified in the wheat genome. Each *TadMTase* member is comprised of three conserved domains (Domain A, HhH-GPD glycosylase domain, and RRM_DME domain). The study also analyzed chromosomal distributions, evolutionary relationships, gene and protein structures, NLS and subcellular localization, CREs, and SRR motifs. Differential expression profiles of TadMTases in different tissues and developmental stages suggest that they are essential in plant growth and development. Furthermore, differential gene expression in response to HS conditions revealed a role for these genes in the wheat's stress response. Future studies on the role of dMTases in plant development and abiotic stresses *via* DNA methylation might add to our understanding and help us improve wheat.

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

Conceptualization, PK and VG; Methodology, VG and HS; Formal analysis, VG and HS; Data curation, TG; Writing—original draft preparation, VG, HS, and TG; Writing—review and editing, PK and VG; Supervision, PK. All authors have read and agreed to the published version of the manuscript.

FUNDING

The study was supported by the DST-INSPIRE Faculty fellowship (Department of Science and Technology, Ministry of Science and Technology, India) awarded to VG.

ACKNOWLEDGMENTS

HS acknowledges the University Grant Commission (UGC), India, for research fellowships.

PUBLISHER ACKNOWLEDGMENT

We would like to thank and acknowledge the contribution of Jitendra Kumar as the initial handling editor of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.894020/full#supplementary-material

REFERENCES

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

Azodi, C. B., Lloyd, J. P., and Shiu, S.-H. (2020). The Cis-Regulatory Codes of Response to Combined Heat and Drought Stress in *Arabidopsis thaliana*. *Nar. Genomics Bioinforma.* 2, lqaa049. doi:10.1093/nargab/lqaa049

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

Bender, J. (2004). DNA Methylation and Epigenetics. *Annu. Rev. Plant Biol.* 55, 41–68. doi:10.1146/annurev.arplant.55.031903.141641

Calarco, J. P., Borges, F., Donoghue, M. T. A., Van Ex, F., Jullien, P. E., Lopes, T., et al. (2012). Reprogramming of DNA Methylation in Pollen Guides Epigenetic Inheritance via Small RNA. *Cell.* 151, 194–205. doi:10.1016/j.cell.2012.09.001

Cao, D., Ju, Z., Gao, C., Mei, X., Fu, D., Zhu, H., et al. (2014). Genome-wide Identification of Cytosine-5 DNA Methyltransferases and Demethylases in *Solanum lycopersicum*. *Gene* 550, 230–237. doi:10.1016/j.gene.2014.08.034

Chan, S. W.-L., Henderson, I. R., and Jacobsen, S. E. (2005). Gardening the Genome: DNA Methylation in *Arabidopsis thaliana*. *Nat. Rev. Genet.* 6, 351–360. doi:10.1038/nrg1601

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

Choi, Y., Gehring, M., Johnson, L., Hannon, M., Harada, J. J., Goldberg, R. B., et al. (2002). DEMETER, a DNA Glycosylase Domain Protein, Is Required for Endosperm Gene Imprinting and Seed Viability in Arabidopsis. *Cell.* 110, 33–42. doi:10.1016/S0092-8674(02)00807-3

Desai, H., Hamid, R., Ghorbanzadeh, Z., Bhut, N., Padhiyar, S. M., Kheni, J., et al. (2021). Genic Microsatellite Marker Characterization and Development in Little Millet (*Panicum Sumatrense*) Using Transcriptome Sequencing. *Sci. Rep.* 11, 20620. doi:10.1038/s41598-021-00100-4

El-Gebali, S., Mistry, J., Bateman, A., Eddy, S. R., Luciani, A., Potter, S. C., et al. (2018). The Pfam Protein Families Database in 2019. *Nucleic Acids Res.* 47, D427–D432. doi:10.1093/nar/gky995

FAO (2021). *World Food and Agriculture - Statistical Yearbook 2021*. Rome: FAO. doi:10.4060/cb4477en

Ferreira, L. J., Azevedo, V., Maroco, J., Oliveira, M. M., and Santos, A. P. (2015). Salt Tolerant and Sensitive Rice Varieties Display Differential Methylome Flexibility under Salt Stress. *PLoS One* 10, e0124060. doi:10.1371/journal.pone.0124060

Gahlaut, V., Samtani, H., and Khurana, P. (2020). Genome-wide Identification and Expression Profiling of Cytosine-5 DNA Methyltransferases during Drought and Heat Stress in Wheat (*Triticum aestivum*). *Genomics* 112, 4796–4807. doi:10.1016/j.ygeno.2020.08.031

Ge, S. X., Jung, D., and Yao, R. (2020). ShinyGO: a Graphical Gene-Set Enrichment Tool for Animals and Plants. *Bioinformatics* 36, 2628–2629. doi:10.1093/bioinformatics/btz931

Gehring, M., Huh, J. H., Hsieh, T.-F., Penterman, J., Choi, Y., Harada, J. J., et al. (2006). DEMETER DNA Glycosylase Establishes MEDEA Polycomb Gene Self-

Imprinting by Allele-specific Demethylation. *Cell.* 124, 495–506. doi:10.1016/j.cell.2005.12.034

Gianoglio, S., Moglia, A., Acquadro, A., Comino, C., and Portis, E. (2017). The Genome-wide Identification and Transcriptional Levels of DNA Methyltransferases and Demethylases in Globe Artichoke. *PLoS One* 12, e0181669–17. doi:10.1371/journal.pone.0181669

Gong, Z., Morales-Ruiz, T., Ariza, R. R., Roldán-Arjona, T., David, L., and Zhu, J.-K. (2002). ROS1, a Repressor of Transcriptional Gene Silencing in Arabidopsis, Encodes a DNA Glycosylase/Lyase. *Cell.* 111, 803–814. doi:10.1016/S0092-8674(02)01133-9

Gu, T., Ren, S., Wang, Y., Han, Y., and Li, Y. (2016). Characterization of DNA Methyltransferase and Demethylase Genes in *Fragaria Vesca*. *Mol. Genet. Genomics* 291, 1333–1345. doi:10.1007/s00438-016-1187-y

Halter, T., Wang, J., Amesefe, D., Lastrucci, E., Charvin, M., Singla Rastogi, M., et al. (2021). The Arabidopsis Active Demethylase ROS1 Cis-Regulates Defence Genes by Erasing DNA Methylation at Promoter-Regulatory Regions. *eLife* 10, e62994. doi:10.7554/eLife.62994

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

Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., et al. (2008). Genevestigator V3: a Reference Expression Database for the Meta-Analysis of Transcriptomes. *Adv. Bioinforma.* 2008, 420747. doi:10.1155/2008/420747

Huh, J. H., Bauer, M. J., Hsieh, T.-F., and Fischer, R. L. (2008). Cellular Programming of Plant Gene Imprinting. *Cell.* 132, 735–744. doi:10.1016/j.cell.2008.02.018

IPPC (2021). in *The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, et al. (Cambridge, United Kingdom: Cambridge University Press).

Kapazoglou, A., Drosou, V., Argiriou, A., and Tsaftaris, A. S. (2013). The Study of a Barley Epigenetic Regulator, *HvDME*, in Seed Development and under Drought. *BMC Plant Biol.* 13, 172. doi:10.1186/1471-2229-13-172

Kim, J.-S., Lim, J. Y., Shin, H., Kim, B.-G., Yoo, S.-D., Kim, W. T., et al. (2019). ROS1-Dependent DNA Demethylation Is Required for ABA-Inducible NIC3 Expression. *Plant Physiol.* 179, 1810–1821. doi:10.1104/pp.18.01471

Kim, S., Park, J.-S., Lee, J., Lee, K. K., Park, O.-S., Choi, H.-S., et al. (2021). The DME Demethylase Regulates Sporophyte Gene Expression, Cell Proliferation, Differentiation, and Meristem Resurrection. *Proc. Natl. Acad. Sci. U.S.A.* 118, e2026806118. doi:10.1073/pnas.2026806118

Korotko, U., Chwiałkowska, K., Sańko-Sawczenko, I., and Kwasniewski, M. (2021). DNA Demethylation in Response to Heat Stress in *Arabidopsis thaliana*. *Ijms* 22, 1555. doi:10.3390/ijms22041555

Kosugi, S., Hasebe, M., Tomita, M., and Yanagawa, H. (2009). Systematic Identification of Cell Cycle-dependent Yeast Nucleocytoplasmic Shuttling Proteins by Prediction of Composite Motifs. *Proc. Natl. Acad. Sci. U.S.A.* 106, 10171–10176. doi:10.1073/pnas.0900604106

Kumar, A., Batra, R., Gahlaut, V., Gautam, T., Kumar, S., Sharma, M., et al. (2018). Genome-wide Identification and Characterization of Gene Family for RWP-RK Transcription Factors in Wheat (*Triticum aestivum* L.). *PLoS One* 13, e0208409. doi:10.1371/journal.pone.0208409

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

La, H., Ding, B., Mishra, G. P., Zhou, B., Yang, H., Bellizzi, M. d. R., et al. (2011). A 5-methylcytosine DNA Glycosylase/lyase Demethylates the Retrotransposon Tos17 and Promotes its Transposition in Rice. *Proc. Natl. Acad. Sci. U.S.A.* 108, 15498–15503. doi:10.1073/pnas.1112704108

Lämke, J., and Bäurle, I. (2017). Epigenetic and Chromatin-Based Mechanisms in Environmental Stress Adaptation and Stress Memory in Plants. *Genome Biol.* 18, 124. doi:10.1186/s13059-017-1263-6

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

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

Letunic, I., and Bork, P. (2021). Interactive Tree of Life (iTOL) V5: an Online Tool for Phylogenetic Tree Display and Annotation. *Nucleic Acids Res.* 49, W293–W296. doi:10.1093/nar/gkab301

Liu, C., Li, H., Lin, J., Wang, Y., Xu, X., Cheng, Z.-M., et al. (2018a). Genome-wide Characterization of DNA Demethylase Genes and Their Association with Salt Response in Pyrus. *Genes* 9, 398. doi:10.3390/genes9080398

Liu, J., Wu, X., Yao, X., Yu, R., Larkin, P. J., and Liu, C.-M. (2018b). Mutations in the DNA Demethylase *OsROS1* Result in a Thickened Aleurone and Improved Nutritional Value in Rice Grains. *Proc. Natl. Acad. Sci. U.S.A.* 115, 11327–11332. doi:10.1073/pnas.1806304115

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

Malabarba, J., Windels, D., Xu, W., and Verdier, J. (2021). Regulation of DNA (de) Methylation Positively Impacts Seed Germination during Seed Development under Heat Stress. *Genes* 12, 457. doi:10.3390/genes12030457

Martins, W. S., Lucas, D. C. S., Neves, K. F. S., and Bertioli, D. J. (2009). WebSat - A Web Software for MicroSatellite Marker Development. *Bioinformation* 3, 282–283. doi:10.6026/97320630003282

Matzke, M. A., and Mosher, R. A. (2014). RNA-directed DNA Methylation: an Epigenetic Pathway of Increasing Complexity. *Nat. Rev. Genet.* 15, 394–408. doi:10.1038/nrg3683

Meena, S., Samtani, H., and Khurana, P. (2022). Elucidating the Functional Role of Heat Stress Transcription Factor A6b (TaHsfA6b) in Linking Heat Stress Response and the Unfolded Protein Response in Wheat. *Plant Mol. Biol.* 108, 621–634. doi:10.1007/s11103-022-01252-1

Mehta, G., Muthusamy, S. K., Singh, G. P., and Sharma, P. (2021). Identification and Development of Novel Salt-Responsive Candidate Gene Based SSRs (Cg-SSRs) and MIR Gene Based SSRs (Mir-SSRs) in Bread Wheat (*Triticum aestivum*). *Sci. Rep.* 11, 2210. doi:10.1038/s41598-021-81698-3

Moglia, A., Gianoglio, S., Acquadro, A., Valentino, D., Milani, A. M., Lanteri, S., et al. (2019). Identification of DNA Methyltransferases and Demethylases in *Solanum Melongena* L., and Their Transcription Dynamics during Fruit Development and after Salt and Drought Stresses. *PLoS One* 14, e0223581–18. doi:10.1371/journal.pone.0223581

Mok, Y. G., Lee, R. Y. J., Weiner, G. M., Eichman, B. F., Fischer, R. L., and Huh, J. H. (2010). Domain Structure of the DEMETER 5-methylcytosine DNA Glycosylase. *Proc. Natl. Acad. Sci. U.S.A.* 107, 19225–19230. doi:10.1073/pnas.1014348107

Molla, K. A., Debnath, A. B., Ganie, S. A., and Mondal, T. K. (2015). Identification and Analysis of Novel Salt Responsive Candidate Gene Based SSRs (cgSSRs) from Rice (*Oryza Sativa* L.). *BMC Plant Biol.* 15, 122. doi:10.1186/s12870-015-0498-1

Naydenov, M., Baev, V., Apostolova, E., Gospodinova, N., Sablok, G., Gozmanova, M., et al. (2015). High-temperature Effect on Genes Engaged in DNA Methylation and Affected by DNA Methylation in Arabidopsis. *Plant Physiology Biochem.* 87, 102–108. doi:10.1016/j.plaphy.2014.12.022

Ono, A., Yamaguchi, K., Fukada-Tanaka, S., Terada, R., Mitsui, T., and Iida, S. (2012). A Null Mutation of ROS1a for DNA Demethylation in Rice Is Not Transmittable to Progeny. *Plant J.* 71, 564–574. doi:10.1111/j.1365-313X.2012.05009.x

Ortega-Galisteo, A. P., Morales-Ruiz, T., Ariza, R. R., and Roldán-Arjona, T. (2008). Arabidopsis DEMETER-LIKE Proteins DML2 and DML3 Are Required for Appropriate Distribution of DNA Methylation Marks. *Plant Mol. Biol.* 67, 671–681. doi:10.1007/s11103-008-9346-0

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

Samtani, H., Sharma, A., and Khurana, P. (2022). Overexpression of *HVA1* Enhances Drought and Heat Stress Tolerance in *Triticum aestivum* Doubled Haploid Plants. *Cells* 11, 912. doi:10.3390/cells11050912

Schultz, J., Copley, R. R., Doerks, T., Ponting, C. P., and Bork, P. (2000). SMART: a Web-Based Tool for the Study of Genetically Mobile Domains. *Nucleic Acids Res.* 28, 231–234. doi:10.1093/nar/28.1.231

Singh, A. K., Chaurasia, S., Kumar, S., Singh, R., Kumari, J., Yadav, M. C., et al. (2018). Identification, Analysis and Development of Salt Responsive Candidate Gene Based SSR Markers in Wheat. *BMC Plant Biol.* 18, 249. doi:10.1186/s12870-018-1476-1

Stecher, G., Tamura, K., and Kumar, S. (2020). Molecular Evolutionary Genetics Analysis (MEGA) for macOS. *Mol. Biol. Evol.* 37, 1237–1239. doi:10.1093/molbev/msz312

Varshney, R. K., Graner, A., and Sorrells, M. E. (2005). Genic Microsatellite Markers in Plants: Features and Applications. *Trends Biotechnol.* 23, 48–55. doi:10.1016/j.tibtech.2004.11.005

Victoria, D., Aliki, K., Venetia, K., Georgios, M., and Zoe, H. (2018). Spatial and Temporal Expression of Cytosine-5 DNA Methyltransferase and DNA Demethylase Gene Families of the *Ricinus communis* during Seed Development and Drought Stress. *Plant Growth Regul.* 84, 81–94. doi:10.1007/s10725-017-0323-y

Wang, P., Gao, C., Bian, X., Zhao, S., Zhao, C., Xia, H., et al. (2016). Genome-wide Identification and Comparative Analysis of Cytosine-5 DNA Methyltransferase and Demethylase Families in Wild and Cultivated Peanut. *Front. Plant Sci.* 7, 1–14. doi:10.3389/fpls.2016.00007

Yang, J., Gu, D., Wu, S., Zhou, X., Chen, J., Liao, Y., et al. (2021b). Feasible Strategies for Studying the Involvement of DNA Methylation and Histone Acetylation in the Stress-Induced Formation of Quality-Related Metabolites in Tea *(Camellia sinensis)*. *Hortic. Res.* 8, 253. doi:10.1038/s41438-021-00679-9

Yang, J., Zhou, X., Wu, S., Gu, D., Zeng, L., and Yang, Z. (2021a). Involvement of DNA Methylation in Regulating the Accumulation of the Aroma Compound Indole in Tea *(Camellia sinensis)* Leaves during Postharvest Processing. *Food Res. Int.* 142, 110183. doi:10.1016/j.foodres.2021.110183

Yang, X., Lu, X., Chen, X., Wang, D., Wang, J., Wang, S., et al. (2019). Genome-wide Identification and Expression Analysis of DNA Demethylase Family in Cotton. *J. Cotton Res.* 2, 16. doi:10.1186/s42397-019-0033-2

Yu, C.-S., Cheng, C.-W., Su, W.-C., Chang, K.-C., Huang, S.-W., Hwang, J.-K., et al. (2014). CELLO2GO: A Web Server for Protein subCELlular LOcalization Prediction with Functional Gene Ontology Annotation. *PLoS One* 9, e99368. doi:10.1371/journal.pone.0099368

Yu, Z., Zhang, G., Teixeira da Silva, J. A., Li, M., Zhao, C., He, C., et al. (2021). Genome-wide Identification and Analysis of DNA Methyltransferase and Demethylase Gene Families in *Dendrobium officinale* Reveal Their Potential Functions in Polysaccharide Accumulation. *BMC Plant Biol.* 21, 1–17. doi:10.1186/s12870-020-02811-8

Yuan, L., Wang, D., Cao, L., Yu, N., Liu, K., Guo, Y., et al. (2020). Regulation of Leaf Longevity by DML3-Mediated DNA Demethylation. *Mol. Plant* 13, 1149–1161. doi:10.1016/j.molp.2020.06.006

Zemach, A., McDaniel, I. E., Silva, P., and Zilberman, D. (2010). Genome-Wide Evolutionary Analysis of Eukaryotic DNA Methylation. *Science* 328, 916–919. doi:10.1126/science.1186366

Zhang, B., Tieman, D. M., Jiao, C., Xu, Y., Chen, K., Fei, Z., et al. (2016). Chilling-induced Tomato Flavor Loss Is Associated with Altered Volatile Synthesis and Transient Changes in DNA Methylation. *Proc. Natl. Acad. Sci. U.S.A.* 113, 12580–12585. doi:10.1073/pnas.1613910113

Zhang, H., Lang, Z., and Zhu, J.-K. (2018). Dynamics and Function of DNA Methylation in Plants. *Nat. Rev. Mol. Cell. Biol.* 19, 489–506. doi:10.1038/s41580-018-0016-z

Zhang, M., Mao, W., Zhang, G., and Wu, F. (2014). Development and Characterization of Polymorphic EST-SSR and Genomic SSR Markers for Tibetan Annual Wild Barley. *PLoS One* 9, e94881. doi:10.1371/journal.pone.0094881

Zhang, Y., He, X., Zhao, H., Xu, W., Deng, H., Wang, H., et al. (2020). Genome-Wide Identification of DNA Methylases and Demethylases in Kiwifruit (*Actinidia Chinensis*). *Front. Plant Sci.* 11, 1–11. doi:10.3389/fpls.2020.514993

Zhou, S., Li, X., Liu, Q., Zhao, Y., Jiang, W., Wu, A., et al. (2021). DNA Demethylases Remodel DNA Methylation in Rice Gametes and Zygote and Are Required for Reproduction. *Mol. Plant* 14, 1569–1583. doi:10.1016/j.molp.2021.06.006

Zhu, C., Zhang, S., Zhou, C., Chen, L., Fu, H., Li, X., et al. (2020). Genome-wide Investigation and Transcriptional Analysis of Cytosine-5 DNA Methyltransferase and DNA Demethylase Gene Families in Tea Plant (*Camellia Sinensis*) under Abiotic Stress and Withering Processing. *PeerJ* 8, e8432. doi:10.7717/peerj.8432

Zhu, J.-K. (2009). Active DNA Demethylation Mediated by DNA Glycosylases. *Annu. Rev. Genet.* 43, 143–166. doi:10.1146/annurev-genet-102108-134205

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.

The reviewer RK declared a shared affiliation with one of the authors TG to the handling editor at the time of the review.

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.

Copyright © 2022 Gahlaut, Samtani, Gautam and Khurana. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Genetics

SYSTEMATIC REVIEW
published: 22 July 2022
doi: 10.3389/fgene.2022.900253

Epigenomics as Potential Tools for Enhancing Magnitude of Breeding Approaches for Developing Climate Resilient Chickpea

*B. S. Chandana[1], Rohit Kumar Mahto[1], Rajesh Kumar Singh[1], Rebecca Ford[2], Niloofar Vaghefi[3], Santosh Kumar Gupta[4], Hemant Kumar Yadav[5], Murli Manohar[6] and Rajendra Kumar[1]**

[1]Indian Agricultural Research Institute (ICAR), New Delhi, India, [2]Center for Planetary Health and Food Security, Griffith University, Brisbane, QLD, Australia, [3]School of Agriculture and Food, University of Melbourne, Parkville, VIC, Australia, [4]National Institute of Plant Genome Research (NIPGR), New Delhi, India, [5]National Botanical Research Institute (CSIR), Lucknow, India, [6]Boyce Thompson Institute, Cornell University, Ithaca, NY, United States

OPEN ACCESS

Edited by:
Abhishek Bohra, Indian Institute of Pulses Research (ICAR), India

Reviewed by:
Ali Raza, Fujian Agriculture and Forestry University, China
Aamir W. Khan, University of Missouri, United States
Javaid Akhter Bhat, Nanjing Agricultural University, China

***Correspondence:**
Rajendra Kumar
rajendrak64@yahoo.co.in

Specialty section:
This article was submitted to Plant Genomics, a section of the journal Frontiers in Genetics

Received: *20 March 2022*
Accepted: *10 June 2022*
Published: *22 July 2022*

Citation:
Chandana BS, Mahto RK, Singh RK, Ford R, Vaghefi N, Gupta SK, Yadav HK, Manohar M and Kumar R (2022) Epigenomics as Potential Tools for Enhancing Magnitude of Breeding Approaches for Developing Climate Resilient Chickpea. Front. Genet. 13:900253. doi: 10.3389/fgene.2022.900253

Epigenomics has become a significant research interest at a time when rapid environmental changes are occurring. Epigenetic mechanisms mainly result from systems like DNA methylation, histone modification, and RNA interference. Epigenetic mechanisms are gaining importance in classical genetics, developmental biology, molecular biology, cancer biology, epidemiology, and evolution. Epigenetic mechanisms play important role in the action and interaction of plant genes during development, and also have an impact on classical plant breeding programs, inclusive of novel variation, single plant heritability, hybrid vigor, plant-environment interactions, stress tolerance, and performance stability. The epigenetics and epigenomics may be significant for crop adaptability and pliability to ambient alterations, directing to the creation of stout climate-resilient elegant crop cultivars. In this review, we have summarized recent progress made in understanding the epigenetic mechanisms in plant responses to biotic and abiotic stresses and have also tried to provide the ways for the efficient utilization of epigenomic mechanisms in developing climate-resilient crop cultivars, especially in chickpea, and other legume crops.

Keywords: chickpea (*Cicer arietinum* L.), epigenomics, DNA methylation, RNAi- RNA interference, mutagenesis, climate resilience, sustainability

INTRODUCTION

Chickpea is the world's second most significant grain legume, produced primarily in the tropics, subtropics, and temperate zones. It is a self-pollinated annual crop with a genomic size of 740 Mbp (Arumuganathan and Earle, 1991). Chickpeas are prized for their high levels of dietary proteins (20–30%), carbohydrates (40%), fibers (3–6%), and lipids (3–6%) (Pushpavalli et al., 2015). In addition, it is also a good source of fiber, minerals, vitamins, lysine, and sulfur-containing key amino acids. It is a resilient crop that is well adapted to stressful situations and is a god's gift to tropical farmers. Chickpea yields on average around 780 kg and can reach up to 2.5 tons per hectare. Various

biotic and abiotic stresses have a negative impact on chickpea yield and productivity. The principal factors that limit chickpea production in farmers' fields include abiotic (drought, salinity, heat, and cold stresses), biotic (insect pests including pod borers, aphids—*Aphis craccivora*, leaf miner, bruchid, etc. and diseases like Fusarium wilt, Ascochyta blight, grey mold, and root rots) stresses. Every year, abiotic stresses cause roughly 6.4 million tons of crop output losses, with soil salinity being primary environmental stress (Jha et al., 2014). Soil salinity is a serious barrier to crop output, and it affects almost 80 million hectares of arable land worldwide (Flowers et al., 2010). Since chickpea is a winter crop, it is subjected to low-temperature stress (0–15°C) during the reproductive stage, which results in a significant loss of flowers and hence pods, reducing output potential by 30–40%. High temperature stress chickpea in late-sown crops, primarily during reproductive and pod filling stages and drought stress at several stages of development; terminal dryness, combined with heat stress during blooming and seed filling can reduce yield up to 70% due to drought and heat stress (Kudapa et al., 2014). Climate change is expected to increase the frequency of temperature extremes (cold and heat), as well as inconsistency in rainfall patterns, necessitating the development of stress-tolerant and climate-resilient chickpea cultivars with region-specific traits that perform well under drought, heat, and/or low-temperature stress. Chickpea production in harsh settings has been improved through a variety of methods, including genetic variability, genomic selection, molecular markers involving quantitative trait loci (QTLs), whole-genome sequencing, and transcriptomics study. Biotechnological technologies have improved our understanding of the genetic basis of chickpea stress tolerance as well as plant responses to abiotic challenges, allowing us to build stress-tolerant chickpeas. The immensity of the current task of maintaining or improving productivity in the face of growing salinity to fulfill yield demands has been clearly recognized and leads to nearly 70% increase in crop production as a top priority (Amin et al., 2016; Hong et al., 2016). So far, Mendelian-based genetic approaches and the selection of heritable target DNA sequences have provided significant genetic improvements in many crop species. In addition, a greater understanding and ability to select beneficial epigenetic and epigenomic changes are proposed to encompass a more efficient and holistic strategy for crop improvement. This is because epigenomic mechanisms are central to governing many plant stress responses, including through cell-autonomous epigenetic switching. This enables the registration and memory of unpredictable genetic signals. The term epigenetics was coined by Waddington (1942) that is used as an intermediate factor between the genotype and phenotype. During gene expression studies, there are various heritable changes occurred due to mitotic and meiotic divisions and are not coded in the DNA sequence itself (Tsaftaris and Polidoros, 2000). Heritable changes in gene expressions are independent of DNA sequence variation and steadily congenital from one generation to another (Berger et al., 2009; Chen et al., 2010a). Variations in the heritability of epigenetic marks (changes) occur during mitotic and meiotic cell divisions. Transient epigenetic changes are not heritable, stout ones, and mitotically transmitted with genome imprinting (Spillane et al., 2001). The meiotically generated epigenetic changes are heritable across the generations without the need for the original stimulus until they are lost or erased. Epigenetic repression is limited to one locus—the genes next to *flowering locus C* (*FLC*) are affected by the cold temperature, for example, which is common for many genes in *Arabidopsis* (Kilian et al., 2007). The loss may be because of genetic change, may be spontaneous (unknown reason), or submission to ambient. Those differ from those that induced the initial epigenetic alterations. The mitotically heritable changes that are not kept through meiosis (epigenetic variation in somatic cells) are lost irrespective of the certitude that mitosis usually perpetuates genetic constitution, such as heterosis is explicated as any edge observed in hybrids. The reverberations of heterosis appear to follow a preferably uncomplicated epigenetic presumption in plants. In hybrids, if the gene is entangled in growth, such as photosynthesis, the plant expressed enhanced vitality (Ni et al., 2008). Heritable epigenetic changes are also referred to as "epialleles," where the epialleles of a locus are identical DNA sequences but display different epigenetic states and hence have an influence on a range of phenotypes (Richards, 2006). These may be classified into three categories based on relative dependence on the genotype: 1) **Pure epialleles** that are solely epigenetic and independent of the genetic variations; 2) **Facilitated epialleles** that partially depend on genetic variation. An example of epiallele transposon is that undergoes DNA methylation spreading into a gene after the insertion of an adjoining transposon. That will be passed across the generations and the changes include both genetic and epigenetic differences; and 3) **Obligate epialleles,** which are directly influenced by the genetic variants and co-segregate with the methylation variants (Woo et al., 2007). Epigenetic variations include various post-transcriptional histone modifications. These modifications include the activity of non-coding RNAs, histone variants, and DNA methylation, which showed drastic changes in plants' response to biotic or abiotic stimuli by changing the transcriptional profile. The memory-directed modifications lead to improved capacity to withstand future stresses (Berr et al., 2011). Thus, epigenetic mechanisms influence the accessibility of DNA to enzymes, resulting in a wide range of gene expression and mRNA splicing. These processes add complexity to the traditional genotype-environment interaction for understanding phenotype expression and development since they are differentially sensitive to the environment (Burggren, 2020). Epigenetic pathways have recently been found as a mediator of this interaction, allowing fast phenotypic diversity in a variety of settings (Burggren, 2016). Under the ongoing climatic changes in agricultural adaptability and resilience to environmental changes, epigenetics has emerged as a major crop development method, ultimately leading to the generation of stable climate-smart crops. This has paved the path for crop breeding to take advantage of epigenetic diversity. Even though epigenetics mechanisms have not been demonstrated in many crop species, most mechanistic investigations are from model plant species. Thus, there is a need to understand how epigenetic mechanisms are linked to mortality predictions as a result of climate change, which affects a wide range of fields, from

environmental conservation to climate change mitigation efforts, and is expected to be more frequent under a climate-change scenario (Amaral et al., 2020; González-Benito et al., 2020). Understanding the roles of epigenetics inducing stresses including histone modifications and DNA methylations helps to uncover the mechanisms that regulate plant-stress interactions and conditions (Chinnusamy and Zhu, 2009). In this review, we have attempted to summarize epigenetic contribution to agricultural adaptation in response to climate change, epigenomic mechanisms, and describe several characteristics in plants, problems in utilization and hypothesize with an objective of the future potential use of epigenetic variations in developing more resilient chickpea crop the staple food legume.

UNDERSTANDING EPIGENETICS AND EPIGENOMICS IN PLANTS AS A MODEL SYSTEM

Understanding of the epigenetic regulatory machinery and mechanism in plants has most notably been achieved in *Arabidopsis thaliana*, (http://www.arabidopsis.org) a model species. Studies on crops, particularly maize, have led to a better understanding and more deep insight into the epigenetic phenomenon (Brink 1956). The implications for maize epigenetic research in the post-genomics era are manifold and it would be difficult to expect future discoveries that are yet to come. However, explanations for an epigenetic regulatory mechanism to stress response in staple crops are still not explored (Arefian et al., 2019; Khan et al., 2019). Along with diverse and overlapping epigenetic regulatory pathways in the maize genome, it could be the most important area of paramutation research for exploration (Chandler et al., 2010). Various reports indicate the novel contributions of the model plant, which have generally been provided in epigenetics and epigenomics. It was the first report in which the discrimination between euchromatin and heterochromatin was explained well based on cytological analyses (Heitz, 1929). Studies in tomato and maize gave heritable changes in expressions related to individual alleles with alternative states, a phenomenon known as paramutation, an inter-allelic interaction that leads to heritable changes in gene expression through mitotic and meiotic routes (Pikaard and Mittelsten, 2014). Paramutation was first explained in maize (Brink, 1956) and furnished the proof for non-Mendelian epigenetic transmissions in plants. The frequent eventuality of entities with changed flower conformity was first explained in the 18th century by Carl von Linne. It was recently observed and reported that a silenced epiallele handled these changes, which contain DNA identical to an expressed allele (Cubas et al., 1999). The innovative exertion of interchangeable ingredients in maize was reported by Barbara McClintock and others (1940), which disclosed the innumerable connections between genetic conduct and epigenetic rules. McClintock (1950) worked and published various reports on the transposable element in maize.

Although various ways regulate the gene expression process in the eukaryotes, DNA methylation is a usual epigenetic process by which cells are used to have switching control of genes in the "off" mode. In the last few years, researchers have unfolded the mechanism of DNA methylation, which led to the fact that methylation is a significant constituent in several cellular mechanisms, including embryonic growth, genomic imprinting, X-chromosome inactivation, and preservation of chromosome stoutness (Phillips, 2008). There are several reported mechanisms in which methylation plays a critical role. Researchers have also connected faults in the methylation process to a series of catastrophic outcomes, including innumerable human diseases (Kelkar et al., 2009). So, we can conclude that these studies help in developing bioinformatics tools that have wide applications across species kingdoms, including chickpeas.

In view of distinguishing differentiation between mammals and plants, it is important to consider the life cycle of plants. In mammals, fertilization is achieved by the fusion of two haploid cells produced by meiosis. However, in the case of plants, haploid (gametophyte) growth takes place that follows meiosis and precedes fertilization as presented in **Figure 1**. The male and female gametophytes are produced by mitotic divisions of the initial haploid meiotic products. In haploid gametophytes, loss of genetic or epigenetic information cannot be compensated for by information on homologous chromosomes. Unlike mammals, there is no evidence for a massive erasure of epigenetic marks during plant gametogenesis. Instead, repressive epigenetic marks in plant sperm and egg cells appear to be reinforced by specific trans-silencing RNAs produced in neighboring nuclei.

The first distinguishing feature to notice is that haploid (gametophyte) development begins with meiosis and continues after conception. There is a loss of genetic or epigenetic information in genetically and metabolically active gametophytes (haploids) that cannot be replaced by homologous chromosomes. In required genes, harmful mutation events are chosen in contrast. Plant gametogenesis, in contrast to mammalian gametogenesis, lacks proof for genome imprinting. Rather, certain trans-silencing RNAs originating in neighboring nuclei appear to be equipped in plant sperm and egg cells to suppress epigenetic changes. That may reflect the process of epigenetic changes that occur during meiosis in plants (Pikaard and Mittelsten, 2014). A second distinguishing trait of plants is the lack of a clearly defined germ line during early embryogenesis. Germ cells are formed in the later stages of plant development. Floral organisms emerge at this stage as a result of transformations from vegetative organs to progeny cells. Meiosis and gametogenesis take place in these cells. Thus, epigenetic changes acquired by meristematic cells in response to plant interactions in the presence of a specific environment have the potential to be passed to germ line cells, reviewed by Pikaard and Mittelsten (2014).

Genetic Modifications in Model Plants for Epigenetics and Epigenomics

The arbitrary introduction of transgenes or transposable elements can be achieved in plant genomes by the process of chemical and or physical mutagenesis (Pikaard and Mittelsten, 2014). In the

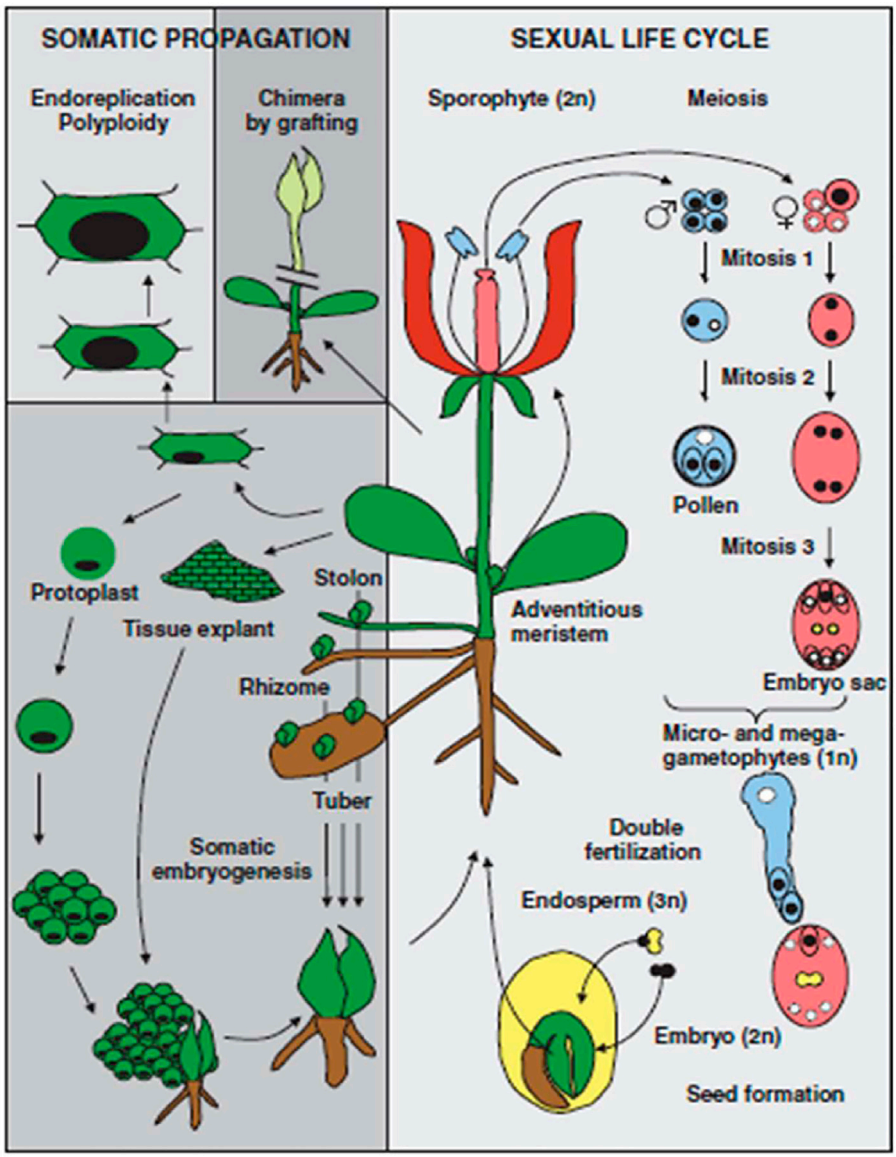


FIGURE 1 | Unique aspects of the plant life cycle (Reproduced from Pikaard and Mittelsten, 2014).

case of *A. thaliana*, it is very easy to identify homozygous mutants, which were identified amongst thousands of progenies of a single mutagen-treated plant. Marker gene is the basis to screen the putative mutants in epigenetic regulators. The promoter region of the *OsMYB91* gene was demethylated and rapid histone modifications at the *OsMYB9* locus in rice account for salinity resistance (Zhu et al., 2015). Increased *Asr1* and *Asr2* gene expressions have been observed during drought-resistance in tomato plants. The expression was enhanced because of the demethylation of putative regulatory and transcribed regions (González and Álvarez, 2013).

The advances in the production of transgenic plants have, therefore, adequately supported epigenetic and epigenomic research. In collaboration with forwarding genetics, another approach, i.e., reverse genetics emphasizing changing gene functions is also feasible. The development of mutants or utilizing transgene-starting RNAi has made an easy way to knockout or knockdown the expression of candidate epigenetic regulator homologs that were previously identified in other organisms. Once, a particular epigenetic mutant is characterized, restrained screenings are usually fortuitous for recognizing interacting constituents or alternative pathways, as observed with *Drosophila* (Elgin and Reuter, 2013) as well as with mouse (Blewitt and Whitelaw, 2013). However, because of the availability of components of epigenetic regulation and exhaustive assemblage of introduced mutations in almost every gene, schematic mutagenesis, and comprehensive instinctive dissimilitude, *A. thaliana* has developed as a model plant used for epigenetic studies based on various researches as presented in **Table 1** (Pikaard and Mittelsten, 2014).

TABLE 1 | Model plant *Arabidopsis thaliana* and their epigenetic regulation (Pikaard and Mittelsten, 2014)

DNA Modification	Mutant name or gene	The Putative or Confirmed Function of Protein
CMT3	Chromomethyl transferase	DNA methyltransferase (mainly CHG and CHH)
ROS3	Silencing repressor	DNA glycosylase-domain protein, cytosine demethylation
MBD10 Methylcytosine-binding protein	Methylcytosine-binding domain protein	Methylcytosine-binding protein

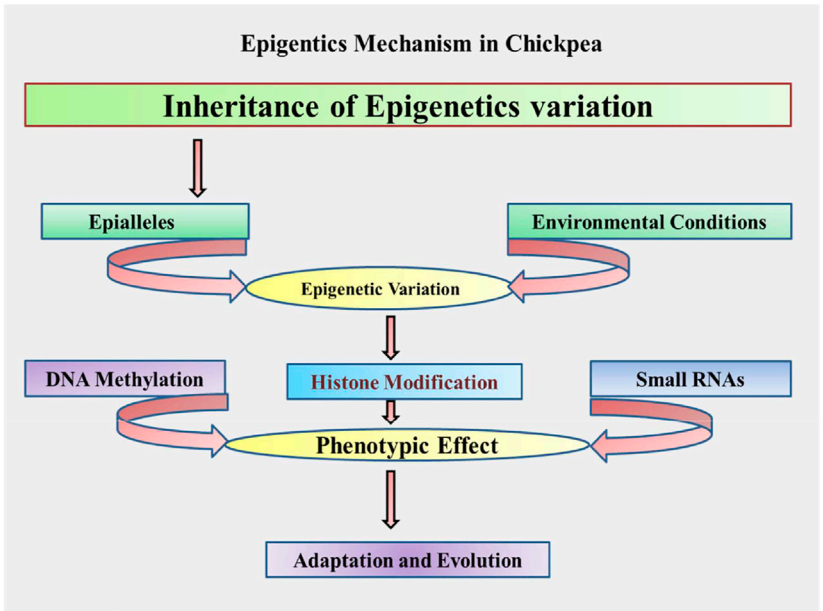


FIGURE 2 | Proposed schematic mechanism/process of epigenetics in chickpea.

MECHANISMS OF EPIGENETICS AND EPIGENOMICS

DNA methylation, histone/nonhistone alterations, and small RNA-mediated interference are the major mechanisms depicted in **Figure 2**, and explained as further in the following sections.

DNA Methylation

DNA methylation is the addition of methyl group at the fifth position of carbon in the DNA molecule of cytosine. DNA methyltransferase carries out the post-replicative modification of DNA known as methylation using S-adenosyl-L-methionine (SAMe) as a methyl donor. Due to this modification, DNA conformation, protein interactions, and chromatin structure were changed, conclusively changing their functional states. Various important biological processes were driven by DNA methylation. These are cell developmental stages, X-chromosome inactivation, transposon tagging, genomic imprinting, and gene silencing. In the case of plants, DNA methylation usually occurs at the positions of CG, CHG, and CHH (H = A, C, or T) and intricate unique DNA methyltransferase, viz, DNMT3B, DNMT1, DNMT2, and DNMT3A (Jurkowska and Jeltsch, 2016). In the case of animals, DNA methylation patterns are associated with the origin, growth, developmental pattern, and progression of cancer (Jones and Baylin, 2007). DNA methylation patterns vary throughout the developmental differentiation in cells and tissues (Barlow and Bartolomei, 2014). This process is carried out by a family of DNA methyltransferase enzymes in eukaryotes. This helps in transferring the methyl groups from the methyl donor SAMe to the cytosine. The resulting 5-methylcytosine (5 mC) is often repressive and leads to gene silencing.

In plants including chickpeas, cytosine methylation is non-randomly distributed mostly to the repetitive regions that abundantly comprise transposable elements of the genome, centromeric frequencies, or multitudes of mute 45S or 5S rRNA gene recurrences. In addition, it further takes place by a handful of divergently controlled enhancers and within the protein-coding domains of tremendously conveyed genes in chickpeas (Zilberman et al., 2007). The eventual gene frame methylation is transformational preservation that plays a role during pre-mRNA splicing (Feng et al., 2010). DNA methylation plays immense contributions to various plant mechanisms, viz,

gene silencing, imprinting of genes, plant immunity, escape from restriction enzymes, apomixis, etc. explained in the following sections.

Long-Term Gene Silencing

The long-term silencing of the gene has been reported in DNA methylation patterns during reprogramming of seed differentiation and its functional relevancy in seed size and seed weight measurements in a large-seeded chickpea cultivar (JGK 3). The identified candidate genes involved in seed size/weight determination exhibited CG context hyper-methylation within the gene and manifold expression in JGK 3 provided insights into the role of DNA methylation in determining size, development, and weight of the seeds. The role of the RNA-dependent DNA methylation pathway has been shown by the gradual achievement of CHH-related DNA methylation in transposable elements (TEs) and by the elevated frequency of small RNAs in hyper-methylated TEs during the development of seed (Rajkumar et al., 2020).

Imprinting of Genes

An epigenetic phenomenon that causes genes to be expressed in a parent-of-origin-specific manner is known as genomic imprinting. This is an inheritance process that is independent of the classical Mendelian inheritance and involves DNA methylation along with histone methylation without altering the genetic sequences. Genome imprinting has been reported with *Arabidopsis* that pushes forward to the rapid escalation of the endosperm, a desirable characteristic (Berger, 2006). This mechanism can apply to commercial crops and hybrids for their regeneration so that we can overcome the current limitations of plant breeding for the perpetual maintenance of hybrid vigor for generations. In the central cell of the female gametophyte, the Demeter molecule-DNA glycosylase domain quantity increases, causing DNA demethylation of the transposons because of a reduction in MET1 methyltransferase. Upon transcription of the transposons, a 24-nucleotide small interfering RNA (siRNA) is produced, which moves to the egg cell and causes imprinting of genes. Similarly, in the vegetative cells of the male gametophyte, DNA demethylation of the transposons occurs because of an increase in the diameter molecule, and upon transcriptions of these transposons, a 21-nucleotide small interfering RNA is produced that moves to the sperm cell and causes imprinting of genes (Bratzel et al., 2012).

Another study has dissected five legumes namely chickpea, soybean, alfalfa, pigeon pea, and lotus indicating the putative role of DNA methylation in the development of inheritable gene silencing and recognized potential DNA MTases (Garg et al., 2014). Based on the domain organization, MTases have been categorized into four subfamilies in legumes, viz, MET, CMT, DRM, and DNA nucleotide methyltransferases (DNMT2). The DNMT2 is a transfer RNA (tRNA) MTase, whereas the first three MTases are a class of DNA MTases. Structural comparative studies of all the known MTases in mammals and plants have assigned biological functions to these MTases (Jurkowski and Jeltsch, 2011). There are various reports in legumes related to the exhaustive gene expression assays of MTases that provide pieces of evidence of their important role in various developmental processes. During the plant life cycle and response to various abiotic stresses, the critical roles of MTases are continued until their survival (Benedito et al., 2008).

Plant Immunity

Plant immunity is the built-in or catalyzed capability of plants to resist or avert a biological strike by pathogens. Molecules emancipated from pathogens are recognized by plant cell exterior sense organs; these receptors stimulate the specific showing cascades that facilitate to withstand the plant against pathogen infection. Roots activate specific tolerance mechanisms in response to elicitors such as molecular/pathogen-associated molecular patterns, (MAMPs/PAMPs), showing compounds (e.g., hormones), and plant defense activators (e.g., β-aminobutyric acid, BABA) (Zhang and Zhou, 2010). DNA methylation can have a critical role in plant immune responses to pathogens; for example, many defense genes in *A. thaliana are* modulated by DNA methylation starting defense reactions against *Pseudomonas syringae* pv *tomato* (Dowen et al., 2012).

Escape from Restriction Enzymes: Methylation-based modification interferes with the restriction process. It serves to protect bacterial chromosomal DNA against "self" restriction and is also responsible for the transient modification of those phages that escape restriction. It was reported in earlier studies that in bacteria, DNA methyltransferases provide a unique mechanism that handles the methylation of specific DNA sequences and is finally linked with epigenetic inheritance (Joseph and David, 2006). The restriction endonuclease recognizes a specifically marked DNA motif those are methylated by an analogs DNA methyltransferase when DNA methylation was primarily unzipped because of restriction-modification (R-M) forms. The R-M systems have been instigated as cellular protection, identifying incoming foreign DNA sequences (viral and another alien) for degradation. The methylation of foreign DNAs was based on specific recognition with related methylase of the same specification. The absolute methylation of the genome is enough to block double-strand DNA cut by the restriction enzyme when the restriction enzyme and its associated methylase both are expressed at levels in R-M systems. Owing to the cell demise that depreciated plasmid containing the EcoRV as post segregationally killing a plasmid comprising the type II R-M EcoRV pattern could not be substituted from the cells by a matchless plasmid (Nakayama and Kobayashi, 1998). The previous accomplishments have advised that R-M systems drive the features of selfish genes (Kobayashi et al., 1999; Kobayashi, 2001).

Apomixis

Plants reproduce by sexual or asexual means and asexual reproduction in plants is carried out by cloning apomixes for sexually reproducing plants including chickpea, fertilization-independent seed formation is not possible because fertilization is a prerequisite for the embryo sac to develop into seeds. For apomictic plants, fertilization is unnecessary because the genes responsible for the fertilization-independent

seed formation produce the apomictic seeds (Hand and Koltunow, 2014). Apomixis was proposed to have developed to enable plant species to propagate under adverse environmental conditions. The previous study has found that reactive oxygen species (ROS) are produced in excess in plants under stress and they possess innate systems for scavenging/detoxifying after they have done their job. Polyamines (putrescine, spermidine, and spermine) are low molecular weight, polycationic aliphatic molecules that are well known for anti-senescence and anti-stress effects because of their antioxidant properties (Kumar and Singh, 2016).

Histone Modifications

DNA methylation, along with histone modifications that include histone acetylation, phosphorylation, methylation, and ubiquitination, contains the most identified epigenetic-associated mechanisms. Similar to DNA methylation, post-translational histone modifications do not possess the ability to influence the DNA nucleotide motif but might change its supply to the transcriptional system. Histone phosphorylation is another mechanism that is performed through histone modifications, often known for its creditability to DNA impairment in reaction to cell injury. The huge gene families in crop plants usually encipher histone-change enzymes by Berr et al. (2011), Deal and Henikoff (2011), and Lauria and Rossi (2011) and are explained in the following paragraphs.

Histone acetylation is carried out by histone acetyltransferases (HATs) enzymes, which mainly add acetyl groups to the histone tails and reduce positive charges and decrease the interaction of histones with DNA. HATs also facilitate transcription by enabling the DNA molecule more accessible to RNA polymerase II. Histone deacetylation is the reverse of histone acetylation and is carried out by Histone deacetylases (HDACs) enzymes. These HDACs remove acetyl groups from histone tails, increase the interaction of DNA, and repress transcription (Bird, 2007). Salinity and drought both are major environmental abiotic obstacles that cruelly affect overall global crop productivity and their nutritional quality (Sen et al., 2017). Plant-specific HD-Zip transcription agents are intricate in plant growth, development, and stresses. A novel HD-Zip (I) gene in chickpea, i.e. *CaHDZ12*, is expressed under water-deficit and salt-distress conditions. An improvement in the tolerance to osmotic stresses was observed in transgenic tobacco genotypes with over-expression *of CaHDZ12*. Silencing of *CaHDZ12* resulted in escalated sensitivity to salt and drought-distresses in chickpeas. Epigenetic changes like histone acetylation at the *CaHDZ12* promoting region have a critical role in distress-induced activation of this gene.

Histone methylation may cause the direct activation or repression of gene expressions and is carried out by histone methyltransferases (HMTs) having several classes including histone lysine methyltransferases (HKMTs), methylated lysine (k) residues, protein/arginine methyltransferase (PRMTs), and methylated arginine (R) residues. The trimethylation of histone H3 at lysine 4 (H3K4) is a specific region for transcription and demethylation to be carried on histone protein H3 to repress the transcription (Collins et al., 2019). Zentner & Henikoff (2013) reported that the addition of the methyl group from S-adenosyl-L-methionine (SAM) into lysine or arginine residues resulted in histone methylation. Histone methyltransferases (HMTs) are the enzyme that catalyzes this process (Bannister & Kouzarides, 2011). The DNA expression is changed through histone methylation by altering the engagement and the unbreakable controlling proteins adhered with the chromatin (Hyun et al., 2017). The histone lysine methylation occurs in mono-, di-, or tri-methylated forms, however, arginine methylation as mono- or di-methylated forms (Bannister & Kouzarides, 2011). The impact of histone methylation on DNA transcription depends on the numbers and to which residues methyl groups are getting added (Zentner & Henikoff, 2013). For example, methylations of H3K4, H3K36, and H3K79 are related to assiduous transcription, whereas the methylations of H3K9, H3K27, and H4K20 are related to tranquility (Black et al., 2012). The study conducted on DNA methylation and physio-biochemical analysis of chickpea in response to cold stress (CS) has reported that CS signals are converted as physiological changes as products of gene expression and are regulated by DNA methylation patterns (Rakei et al., 2015). The major roles of antioxidant enzymes (superoxide dismutase, catalase, ascorbate peroxidase, guaiacol peroxidase, and polyphenol oxidase) along with a noticeable ratio of changes in DNA methylation/demethylation patterns were often decisive factors in the preservation of cells against cold influenced oxidative distress.

Enzymes Involved in Modifications of Histone Variants

In eukaryotic taxa, among the most conserved proteins include histone variants, linker histones, and non-histone proteins that are enciphered by highly superfluous gene families. The distinct categories of H2A and H3 histone variants, like animals also in plants based on their structure and function, have been identified (Henikoff and Smith, 2015). The physical properties of histone variants significantly affect their dynamic relations with DNA (Ingouff and Berger, 2010; Deal and Henikoff, 2011). The DNA damaged regions are discovered through phosphorylation of the H2AX variant and also assist during the recruitment of DNA repair proteins. A type of histone variant, H2A.Z exists mostly near the transcriptional start site of genes, most probably regulating transcription (Zilberman et al., 2008). The expression of this variant requires the initiative of an SWR1 chromatin-remodeling complex. Separation of DNA from H2A.Z comprising nucleosomes during heat stress is observed, which is followed by alterations in gene expression (Kumar and Wigge, 2010). During cell division, a particular histone H3 variant, i.e., CenH3 specifies the nucleosomes of centromeric regions and plays an important role in kinetochore assemblage, microtubule association, and chromosome segregation. H3.3, another histone variant having only a few different amino acids from its canonical H3 subunit, is predominantly found in regulatory regions. When specific linker histone proteins were downregulated as compensation, there is an upregulation of other histone variants. This results in a clear-cut phenotypic defective change with pleiotropy DNA hypomethylation (Jerzmanowski, 2007).

TABLE 2 | Chickpea-specific histone proteins (Chidambaranathan et al., 2016).

Gene Id	Protein ID	Groups	Protein size (aa)	MW(kDa)	pI
LOC101514392	XXP_004487028	III	150	15.95	10.75
LOC101514067	XXP_004487027.1	III	149	15.93	10.96
LOC101514719	XXP_004487029.1	III	148	15.75	10.73
LOC101492287	XXP_004493627.1	II	143	15.08	10.47
LOC101500089	XXP_004494603.1	III	146	15.42	10.35
LOC101514555	XXP_004494649.1	II	139	14.62	10.36
LOC101489870	XXP_004487625.1	I	135	14.06	10.05
LOC101490207	XXP_004487626.1	I	134	14.05	10.05
LOC101507294	XXP_004498649	IV	134	14.27	10.39
LOC101489506	XXP_004495152	IV	134	14.31	10.39
LOC101497893	XXP_004508547	IV	131	14.05	10.28

Chickpea-Specific Histone-Modifying Enzymes

Several chickpea-specific histone-modifying enzymes as given further have been reported. The histone H2B is a core component of nucleosome that wraps up long compact DNA into chromatin, limiting DNA availability to the cellular mechanisms and using DNA as a template. Therefore, histones are an integral part of transcription regulation, chromosome stability, DNA repair, and DNA replication. A complex set of post-translational modifications of histones, known as histone code, and nucleosome remodeling help in regulating DNA accessibility. Histone H2B performs the molecular functions of DNA binding and protein heteromerization (https://www.uniprot.org/uniprot/Q9M3H6). It was reported that under heat stress conditions expression of six chromatin remodeling complex genes (*SWR1*) in diverse tissues of chickpea was based on nucleosome response through histone H2A.Z variants. A group of seven genes that are homologous to chromatin remodeling complexes (*SWR1*) of *Arabidopsis* was also identified in the chickpea genome. Three genes of chickpea homologs of photoperiod independent early flowering 1 (*PIE*), Actin associated protein (*ARP6*), two serrated leaves, and early flowering (*SEF*) for histone 2A variant-Z (H2A.Zs-a thermal sensor in plants) the three genes were analyzed for their appearances under heat distress and five diverse tissues. A significant role in chromatin remodeling complexes under heat stress conditions might be played by *CarPIE1* gene. The entire three histone CarH2A.Z variants acted as potential candidate genes for the characterization of their specific function (Chidambaranathan et al., 2016). These chickpea-specific histone proteins are summarized in **Table 2**.

Furthermore, various forms of histone modifications and their sites along with effects on transcriptional activity of genes are summarized in **Table 3**.

TABLE 3 | Histone alterations, their sites and impacts on the activities of transcription.

Alterations, sites, abbreviations	Impact on transcription
Acetylation of histone	
Histone 3 Lysine 4 acetylation (H3K4ac)	Activating/Permissive
Histone 3 lysine 9 acetylation (H3K9ac)	-do-
Histone 3 lysine 14 acetylation (H3K14ac)	-do
Histone 3 lysine 18 acetylation (H3K18ac)	-do
Histone 3 lysine 27 acetylation (H3K18ac)	-do
Histone 4 lysine 16 acetylation (H4K16ac)	-do
Histone 3 pan acetylation (H3ac)	-do
Histone 4 pan acetylation (H4ac)	-do
Methylation of histone	
Histone 3 lysine 4 methylation (H3K4me1)	Activating/permissive
Histone 3 lysine 4 dimethylation (H3K4me2)	-do-
Histone 3 lysine 4 trimethylation (H3K4me3)	-do
Histone 3 lysine 9 dimethylation (H3K9me2)	Repressive
Histone 3 lysine 9 trimethylation (H3K9me3)	-do
Histone 3 lysine 27 trimethylation (H3K27me3)	-do
Histone 3 lysine 36 trimethylation (H3K36me3)	Activating/permissive
Histone 3 lysine 79 methylation (H3K79me1)	Activating/permissive
Histone phosphorylation	
Histone 2A ubiquitination (H3S10ph)	Activating/permissive
Histone ubiquitination	
Histone 2A ubiquitination (H2Aub)	Repressive
Histone 2B ubiquitination (H2Bub)	Activating/permissive

Role of Histone Modifications on Vernalization

The plant may continue to grow vegetatively through cell division during the cold period. When new seeds are produced, after the vernalization of the parent plant, the seeds are "reset." The new plants they produce from the seeds will themselves have to go through their cold season before flowering. The key gene intricate in vernalization is referred to as FLOWERING LOCUS C or *FLC*. *FLC* encodes a protein known as a transcriptional repressor. It binds to other genes and stops them from getting switched on. These three genes *FT, SOC1*, and *FD* specifically regulate flowering in *Arabidopsis thaliana*, and show that the epigenetic status of *FLC* alters after a prolonged duration of cold. Experiments with mutated versions of epigenetic enzymes have shown that the changes in histone modifications at the *FLC* gene are critically important in controlling the flowering response. For example, there is a gene called *SDG27* that adds methyl groups to the lysine amino acid at position four on histone H3, so it is an epigenetic writer that is associated with a vigorous gene expression. The *SDG27* gene can be mutated experimentally

TABLE 4 | Epigenetic studies related to biotic and abiotic stresses in chickpea.

S. No.	Biotic/Abiotic stress tolerance through an epigenetic mechanism	References
1	Salt tolerance mechanism in chickpea	Arefian et al. (2019)
2	Mechanism of drought stress	Khan et al. (2019)
3	The study on DNA methylation pattern Development and differentiation of seed size	Rajkumar et al. (2020)
4	Mechanism of salt tolerance in chickpea	Khandal et al. (2017)
5	Physio-biochemical analysis of chickpea in response to cold stress	Rakei et al. (2015)
6	Chickpea drought, water, and osmotic stress	Kaur et al. (2002) and Elkoca et al. (2007)
7	DNA methylation patterns in cultivated chickpea to understand the regulation of gene expression in different organs	Bhatia et al. (2018)
8	Drought and salinity resistance by an epigenetic mechanism in chickpea	Sen et al. (2017)
9	DNA methylation and epigenetics mechanism on physio-biochemical analysis of chickpea in response to cold stress	Rakei et al. (2015)
10	The epigenetic mechanism to heat stress in chickpea	Chidambaranathan et al. (2016)
11	Role of epigenetics in drought yield index	Sharma et al. (2019)
12	Chickpea, drought water and osmotic stress	Kilian et al. (2007)

TABLE 5 | Intergenerational stress memory resistance development in crop plants through epigenetic modifications.

Crop species	Stress resistance	Treatment/Pathway	References
Chickpea (*Cicer arietinum*)	Drought	Water and osmotic stress	Elkoca et al. (2007) and Kaur et al. (2008)
Mung bean (*Vigna radiate*)	Drought/salinity	Halopriming of seeds with NaCl and PEG	Jisha and Puthur (2014)
Alfalfa (*Medicago sativa*)	Drought	Seed osmotic treatment with PEG	Mouradi et al. (2016)
Cowpea (*Vigna unguiculata*)	Drought	Water, osmotic, and hormonal seed stress	Eskandari and Kazemi (2011) and Boucelha and Djebbar (2015)
Arabidopsis (*Arabidopsis thaliana*)	Drought/salinity/biotic stress	β-amino-butyric acid, hyperosmotic priming of seedlings	Slaughter et al. (2012) and Sani et al. (2013)
Soybean (*Glycine max*)	Drought/salt	Indole acetic acid and NaCl stress on seedlings induced long non-coding RNAs and DNA methylation	Umezawa et al. (2000) and Chen et al. (2019)
Mung bean (*Vigna radiate*)	Drought/heavy metals	Indole-3-butyric acid	Li et al. (2018b)

so that it no longer encodes an active protein. Plants with this mutation have less of this active histone modification at the *FLC* gene promoter. They produce less *FLC* protein, and so are not so good at repressing the gene that triggers flowering. The *SDG27* mutants flower earlier than the normal plants. Cold weather induces protein in plant cells called VIN3 that works as chromatin and can bind to the *FLC* promoter. When *VIN3* binds to the *FLC* promoter, it alters the local structure of the chromatin instead of how tightly chromatin is wrapped up, making it often available to other proteins. Often, opening up chromatin leads to an increase in gene expression. However, in this case, *VIN3* attracts yet another *FD* (FLOWERING DETERMINATE) enzyme that can add methyl groups at position 27 on lysine residue amino acid of histone H3 protein. This modification represses gene expression and is one of the most important methods that plant cell uses to switch off the FLC gene. Following cold weather, the cells in *Arabidopsis thaliana* produce a long RNA, which does not code for a protein called *COLDAIR*. The *COLDAIR* non-coding RNA is present, particularly in the FLC gene. When localized, it binds to the enzyme complex that creates the important repressive mark at position 27 on histone H3. *COLDAIR,* therefore, acts as a targeting mechanism for the enzyme complex. From these data, we can see that flowering plants use some of the same epigenetic machinery as many animal cells. These include histone protein alternations, and the utilization of long non-coding RNAs to target these changes. Earlier, it has been inferred that destabilization of the cells is a consequence of global DNA hypomethylation through DNMT1-depletion led, which ultimately leads to the production of aneuploids (Barra et al., 2012). In the case of aneuploids (45, XO; 46, XX; and 47,XXX) expression of DNA methyltransferase 1 and DNA methylation enzymatic gene showed a positive association with inactive X chromosomes (Rajpathak and Deobagkar, 2017).

Role of Histone in Epigenetic Regulation in Antiviral Innate Immunity

To deduce and establish the critical role of histones in epigenetic regulation during the process of viral innate immunity, there is a great need for better grasping of these complicated interactions through the epigenetic lens, which may have therapeutic opportunities in the clinic. A grasping of the parts played by the key epigenetic controllers—chromatin remodeling and histone alterations—in atonements of chromatin candidness in the process of host defense against virus, how the RNA alteration m6A (N6-methyladenosine) influences basic features of hostvirus interplaying and conclusions with subsequent orchestrations for better understanding about epigenetic regulations in host and viruses' contaminations are required (Xiao et al., 2021).

Non-Histone Proteins and Their Roles

Similar to other eukaryotes, non-histone chromosomal proteins are also found in plants, which may assist epigenetic gene regulations, including HMG proteins. The HMGB family of proteins is the foremost assayed and varied subgroup of proteins in plants, members of whom differentiate in the level of expression, localization, style, and inter-playing with DNA along with other proteins. The partial sub-functionalization of individual family members results from mutation and abnormal expression showing their role in developmental stages and response to various stress stimuli (Pedersen and Grasser, 2010). The structure-specific recognition protein (*SSRP1*) indirectly contributed to the demethylation of DNA the recognized genes in the female gametophyte's central cell (Ikeda et al., 2011). Yuan et al. (2011) have reported that the structure, assembly, and rejection of cohesion seem to be highly protected. There are only limited family members in plants that might have specific functionalities. It was also found that the defective meristem silencing 3, which was involved in *de novo* (*DMS3/IDN1*) is required for transcription of DNA-dependent RNA polymerase V (RNA Pol V) (Haag and Pikaard, 2011). This plays a very important role in the establishment of *RdDM* (RNA-controlled DNA methylation) proteins (Varshney et al., 2019a). The various mutant screens related to epigenetic regulators are the sites of *REPLICATION FACTOR C1* and *REPLICATION PROTEIN A2* (*RPA2*) (Elmayan et al., 2005; Kapoor et al., 2005). The role of this protein related to stem cell and meristem repair is also revealed when chromatin mutants having increased phenotypes and mutants having topoisomerase homolog *MGOUN* (MGO) functional loss were merged (Graf et al., 2010). This signified that many other non-histone proteins involved with DNA will also act as direct or indirect epigenetic regulators.

Nucleosome-Organizing Proteins

Short-term or long-term modifications in the nucleosomes' positioning and their connection with DNA were always required for replication, transcription, recombination, and maintenance. As a result, vigorous mechanisms on the chromatin amend DNA or protein modifications, incorporate alterations in nucleosome's possession and, constitution, along with the attainability of the DNA to variegated proteins (Pikaard and Mittelsten, 2014).

Chromatin-Remodeling Complexes

Becker and Workman (2013) have reported that chromatin remodeling can be used for the relocation or dissociation of nucleosomes. It was first reported in yeast and was named after the respective processes that ATPases such as the *SWI/SNF* complexes have influenced the mutants. Many similar complexes were also reported in plants (Jerzmanowski, 2007). The functional information for very few putative chromatin remodelers has been got through genetic screens, the first recognized being DECREASE IN DNA METHYLATION1 (*DDM1*). The function of *DDM1* included genome-wide decreased activity of methylation of DNA and H3K9me2, activating repetitive elements for transcription, and downregulation of many such genes. Therefore, the mutants for *ddm1* exhibit severe defects in developmental and morphological growth that may reach an extreme in future generations. The amalgamation of epimutations and insertional mutations induced through reactivated transposons is the main reason behind the gradually diminished fitness of ddm1 mutants. In ddm1 mutants, epigenetic information is permanently deleted and can be restored through backcrosses with wild-type plants as epigenetic patterns at several loci, mainly because of the outcome of *de novo* methylation (Teixeira et al., 2009). *DDM1* also shows *in vitro* ATP protease nucleosome moving initiative (Brzeski and Jerzmanowski, 2003). Mutants deficient in *DDM1* and linker histone H1 originate when cytosine methylation deficiency occurs in ddm1 mutants (Zemach et al., 2013) signaling that the requirement of *DDM1* is essential for the repair of methylation mechanics to discover DNA in nucleosomes consisting of core and linker histones. The defective RNA-mediated DNA Methylation 1 (*DRD1*) and *CLASSY 1* (*CLSY1*), extremities of the *SWI2/SNF2* family found in *Arabidopsis*, are unique to the plant kingdom including chickpea, and play a peculiar part in RNA-induced DNA methylation. Four additional *SWI2/SNF2* derived proteins namely; BRAHMA (BRM), MINUSCULE (1, 2), and SPLAYED (SPD) are intricate in the RNA-led DNA methylation (Sang et al., 2012). Other than ATPases, quintessence parts of *SWI/SNF* remakes are also reported in plants, inclusive of one SNF5 homolog (BSH), two *SWP73* homologs, and many *SWI3* family members (*AtSWI3 A-D*) (Jerzmanowski, 2007). However, their direct contributions to plants are still unexplored. However, it has been unraveled that *SWI3* interplays with RNA binding proteins lead to RNA-directed DNA methylation (Zhu et al., 2012).

The role of histone protein for drought and yield index (DYI) in chickpeas was studied, and it was disclosed that the development of functional molecular tags derived from the cis-regulatory sequence components of genes is crucial for their deployment and identification of several conserved non-coding SNPs (CNSNP). Among those, the two made-up natural haplotypes and alleles are derived from a histone H3 protein-coding gene and its transcriptional regulator NAC transcription factor (TF) anchoring the major QTLs and trans-acting eQTL controlling drought yield index (DYI) in chickpea (Sharma et al., 2019).

RNA-Mediated Interference

RNA interference (RNAi) is a technique in which tiny RNA molecules are combined with other molecules to target homologous DNA regions. They bring together the agents that alter chromatin, resulting in heterochromatin formation and gene suppression. Pre-transcriptional gene silencing can stop transcription from happening. As a result, DNA methylation at genomic locations corresponding to complex siRNA or miRNA is catalyzed by an enzyme complex. RNA interference in chickpea and other legume crops has been found to have a large number of drought-responsive miRNAs. In response to salt stress, 259 miRNAs were shown to be differentially expressed in the root tip of chickpea during drought and salinity stress,

which were also seen in other legumes such as soybean root apex (Khandal et al., 2017). TIR1 (TRANSPORT INHIBITOR RESPONSE 1), an auxin receptor, and AUXIN RESPONSIVE FACTOR 10 (ARF10) and ARF16 are targets of miR393 and miR160 (Chen et al., 2011). Overexpression of miR160 causes unregulated cell division and a loss of gravity sensing at the root tip during primary root development (Mallory et al., 2005). MiR164 inhibited auxin signaling for lateral root initiation by targeting the transcription factor NAC1. ARF6 and ARF8, which are positive regulators of adventitious root growth, were targeted by miR167 (Gutierrez et al., 2009; Guo et al., 2005). Comparative miRNA expression profiling of *Medicago truncatula* (Medicago) in the root tip and elongation zone, as well as root-forming callus and non-root forming callus, revealed 107 miRNAs from 44 families expressed in these tissues and predicted conservation of some of the miRNA/target relationships seen in other species (Eyles et al., 2013). Overexpression of MiR396 in Medicago roots inhibits cell-cycle gene expression and limits root development (Bazin et al., 2013). miRNA expression analysis in normal soybean roots, as well as comparisons between phosphate-starved and phosphate-sufficient soybean roots, revealed some new miRNA/target interactions (Xu et al., 2013). In Arabidopsis, an increase of miR393, miR397b, and miR402 expression occurs under dehydration and salt stress, according to a study. miRNA has a vital function in controlling root growth under abiotic stresses (Ding et al., 2009; Lu et al., 2014). Drought stress increases the expression of miR398a/b and miR408 in the Medicago root (Trindade et al., 2010) and miR169g in rice roots (Zhao et al., 2007). In Arabidopsis, mi RNA165/166 regulates root development by targeting transcripts of leucine-zipper family proteins (Singh et al., 2014). In Medicago, overexpression of miR160 altered root development and nodule number (Bustos-Sanmamed et al., 2013). In another work, epigenetic modulation of drought stress in chickpea was investigated. They notably researched MicroRNAs (miRNAs), non-coding RNAs that have been identified as significant controllers of gene performances like BHLH23 operating at post-transcriptional stages, which are implicated in tolerance to water constraints, as well as extra abiotic distress. BHLH23 transcription factor, which encodes for low copper levels, was found to be downregulated, and another drought stress-responsive gene, APETALA2/Ethylene Response Factors (ERF/AP2), was found to have a lower expression profile in miR408 over-expressed chickpea plants when compared to vector control plants after stress treatment (Hajyzadeh et al., 2015).

Methylation of cytosine and modifications of histone plays an important role in the gene regulatory mechanisms of genes responsible for epigenetic changes in plants. These modulations serve as gene regulators during transcriptional activities. Post-transcriptional modifications in context to epigenetic regulation occur through targeted degradation of mRNA. Finally, translational repression occurs and post-transcriptional gene silencing (PTGS) of mRNAs acts as a defense molecule against several pathogens. These are viruses, bacteria, fungi, molds, and transgene (Ruiz and Voinnet, 2009; Vazquez et al., 2010). The small RNAs play an active role in transcriptional and post-transcriptional gene silencing in plants (Chapman and Carrington, 2007). The actions of these miRNAs or siRNAs in plants have a resemblance to eukaryotic biogenesis (Shabalina and Koonin, 2008). However, multiple pathways have been involved in the duplication and sub-functionalization of genes guiding miRNA or siRNA-mediated processes in plants as follows (Herr, 2005; Baulcombe, 2006).

i. Biogenesis process for the miRNAs that are complementary to the targeted sequences;
ii. A ropeway in which a miRNA activates the origin of secondary trans-acting siRNAs with no complementarities to the starting miRNAs;
iii. A route for siRNA-intervened abasement of infringed viral RNAs or transgene RNAs along with
iv. A route for siRNA-intervened methylation of DNA, transcriptional mute of transposons/viruses, and other genes.

Variegation of the core machinery for siRNA biogenesis along with function pinpoints the evolutionary process of the various small RNA suppressing mechanisms in plants (Vazquez et al., 2010). Plants, like fission yeast (*S. pombe*) and nematodes (*C. elegans*), also make use of RNA-dependent RNA polymerases in dsRNA production. *Arabidopsis* genome enciphers six unique RdRPs. The plant Dicer produces diversified sizes of small RNAs, viz, miRNAs of 21 nt (*DCL1*), or siRNAs of different sizes 23–24 nt (*DCL3*), 22 nt (*DCL2*), or 21 nt (DCL4). These diverse siRNAs differ in size but overlap in functions, due to their relatedness to variegated AGO protein that contains 10 extremities in *Arabidopsis* (Vaucheret, 2008).

Plant-Specific RNA-Directed DNA Methylation

Various proteins, such as AGO, Dicers, and RdRPs, were employed in a kind of permutations for accomplishing *de novo* methylation that occurs during the process of RdDM. It is initiated by the methyltransferases (DNA) of DRM grade. The exhaustive recruitment processes of *DRM2* in DNA that takes place are still not conspicuous. The process is undertaken in green algae before plants prominently produce RNA Pols IV, and V, the complex configurations of RNA Pol II (Luo and Hall 2007; Tucker et al., 2010). RNA Pols II, IV, and V each have 12 basic parts in *Arabidopsis*, nearly half part of that is often for the above three explained polymerases and distorted by the interchangeable genes (Ream et al., 2009). The genes that originated through the reoccurrence of RNA Pol II subunit genes, bolstered with sub engaged for certain subunits enciphered the subunits that are specific to RNA Pols IV or V (Ream et al., 2009; Law et al., 2011).

Paramutation is an interaction between alleles of a gene in such a way that an allele is heritably affected by another allele. This phenomenon is explained nicely through the booster-1 (b-1) site in maize (Brink, 1956). A para mutable (B-l) allele (active allele) after getting affiliated with a para-mutagenic (B') allele (inactive allele) becomes a paramount (B-l*) allele. Uniform DNA sequences for the two alleles at the b-1 locus are found but vary in

the system of methylation (DNA). Para mutant allele itself displays as para mutagenic and is unchanged through one or more following generations. However, most alleles are neither para mutable nor para-mutagenic.

The above discussed various mechanisms and processes governing epigenetics can be adopted in chickpeas and presented through a schematic flow diagram depicted in **Figure 2**.

EPIGENETIC AND EPIGENOMIC STUDIES IN CHICKPEA

As presented in the **Table 4**, the epigenetic studies in biotic and abiotic stress response, DNA methylation is a crucial component in gene assertion control. The DNA methylation status in seven resistant and susceptible cultivars of chickpea for *Fusarium oxysporum* f. sp. was determined using the methylation-sensitive amplified polymorphism (MSAP) assay and 27, 468 DNA fragments were obtained, each of which represented a recognition site cleaved by one or both isoschizomers amplified using selected primers (Mohammadi et al., 2015). They showed DNA methylation patterns in leaves, stems, and roots from both controlled and inoculated plants, and found extensive cytosine methylation modifications in pathogen-treated/infected plants, but none in controls. Heterologous expression of WRKY40 promoter and its transcriptional regulation via epigenetic alteration controls the fusarium stress resistance. This expression aids in the prevention of bacterial infections spreading due to resistance (Chakraborty et al., 2018). The important function of the WRKY40 transcription factor in the susceptibility of chickpea to *Fusarium oxysporum* f. sp. ciceri race 1 (Foc1) and resistance to this strain (WR315) has been demonstrated. In a controlled and Fusarium-affected environment, the histone changes in two chickpea genotypes were evaluated using immunoblotting and real-time PCR techniques. In the process of resistance interaction with Foc1, location-specific Histone three lysine nine acetylation, a positive signal of transcription, becomes reinforced at the WRKY40 promoter. In Foc1-infected susceptible plants, the H3K9 Ac is reduced at the WRKY40 promoter. The salt tolerance mechanism in chickpea was studied using an epigenetic approach in FLIP 97-43C (salt-tolerant) and FLIP 97–196C (salt-susceptible), which aids in the discovery of proteins that regulate photosynthesis, distress responsiveness, and protein assimilation (Arefian et al., 2019).

Epigenetics Studies in Other Legumes

The cytosine residues in the DNA of pea root tips subjected to water deficit were investigated to see if there was a link between environmental stress and DNA methylation. Two complementary approaches were used to assess DNA methylation: (i) immunolabeling with a monoclonal antibody against 5-methylcytosine, and (ii) MSAP (Methylation-Sensitive Amplified Polymorphism) to see if methylation and demethylation in response to water deficit could be linked to specific DNA sequences (Labra et al., 2002).

Plant microRNAs were investigated in beans (Dela et al., 2019). They are generally transcribed in transcripts with a single microRNA precursor, which is processed by DICERLIKE 1 and associated proteins to produce a short RNA, which is then incorporated into an AGO-containing protein complex to direct silencing of an mRNA with a complementary target sequence. Certain microRNA loci have several precursor stem-loop structures, encoding multiple microRNAs in a single transcript that is one-of-a-kind example in which the evolutionarily conserved miR398a is encoded in the same transcript as the legume-specific miR2119. Other legumes showed the same dicistronic configuration as the common bean. The role of small RNAs in reaction to water stress was investigated in *Phaseolus vulgaris*, and it was discovered that mature miR398 and miR2119 are repressed in response to water deficit, but that they are functional since they target the mRNAs for CSD1 and ADH1, respectively. The down-regulation of miRNA with the consequences of upregulation of CSD1 and ADH1 genes in common beans and possibly in other legumes respond to water deprivation (Naya et al., 2014). In the case of cowpea, the homology search was used to predict miRNAs and their targets. Real-time quantitative PCR was used to confirm the identified cowpea miRNAs in the leaves and roots of drought-stricken cowpea plants. Target gene prediction reveals that a group of miRNA target genes is implicated in metabolic pathways associated with physiological changes caused by drought stress. We looked at the expression levels of some key genes involved in physiological responses to drought stress and discovered that differences in their expression levels corresponded to the various drought responses of drought-sensitive and drought-resistant cowpeas (Shui et al., 2013).

The legume miR1514a activates phasiRNA by modulating a NAC transcription factor transcript. MicroRNAs have been identified as post-transcriptional regulators implicated in stress responses in recent investigations. In Phaseolus vulgaris (common bean), miR1514a is a legume microRNA that is activated in response to drought stress and has varying levels of accumulation in roots during water deficit in two cultivars with different drought-resistance phenotypes. The role of miR1514 in the regulation of a NAC transcription factor gene via phasiRNA synthesis during response to drought has been reported in case of soybean (Sosa et al., 2017).

INTEGRATING EPIGENOMICS WITH OMICS APPROACHES FOR BIOTIC AND ABIOTIC TOLERANCE IN CHICKPEA

Many biotic and abiotic stress tolerance gene(s) in plants have been discovered through recent advances in next-generation sequencing (NGS) (Garg et al., 2016). Integration of omics-generated data from several platforms, such as transcriptomics, which is coupled with proteomics, and finally, metabolomics, is essential to close the genome-to-phenome gap in agricultural plants. These platforms and their data enable to identify the certain phenotypes based on genetic contribution (Choi, 2019). The use of the omics strategy to gather genomic information to

influence various biological processes, as well as the discovery of differentially expressed genes in various environmental situations and positional cloning. This strategy can also be utilized in the targeted region with an mRNA or protein shift to uncover the role of connected genes associated with the trait of interest (Su et al., 2019). A comparison of salt stress generated FLIP 97-43C (salt-tolerant) and FLIP 97–196C (salt-susceptible) was undertaken to understand the salt tolerance mechanism in chickpea, which resulted in the identification of proteins regulating photosynthesis, distress responsiveness, and protein absorption (Arefian et al., 2019). The researchers discovered 134 proteins that were expressed differently in the extracellular matrix and during the dehydration response. During the comparative proteomics investigation of JG-62, these proteins were discovered in a variety of biological roles (Bhushan et al., 2007). Through a targeted metabolomics approach, Khan et al. (2019) identified key upregulated metabolites such as allot in, L-proline, L-arginine, and L-histidine, as well as downregulated metabolites such as alanine, choline, gamma-aminobutyric acid, and phenylalanine, that were differentially expressed under drought stress conditions. From sugars to organic acids, a total of 48 distinct metabolites were discovered. Under salt stress, 28 biogenic amino acids were expressed in chickpea cultivars with varying salt tolerance (Dias et al., 2015). The specified compounds were quantitatively analyzed using modern metabolomics techniques and GC and LC were integrated into a triple quadrupole mass spectrometer (GC-QqQ-MS and LC-QqQ-MS). As a result, the reports for drought tolerance mechanisms in chickpea genotypes, omics techniques, and crop production management were shown to be the best and most cost-effective.

The cultivated chickpea has a narrow genetic base (Varshney et al., 2013) and phenotypic plasticity (Berger, 2006). It is difficult to locate the stress-responsive and undeniably tolerant gene(s), especially when plants accept cross-talk to react to many concurrent distresses (Tuteja, 2007). The physiological and genomic screening revealed that there was a wide range of genetic differences among and within the tolerant and sensitive genotypes for salinity tolerance. For example, cold was included in tolerant-1 and inhibited in tolerant-2 during gene profiling using microarray aquaporin genes for drought, salinity, and homology-based induction for salinity, heat, and environmental stress (Mantri et al., 2007; Kotula et al., 2015; Kaur et al., 2008). Several haplotypes and significant numbers of alleles associated with agronomic parameters in chickpeas have been uncovered using genomic resources (Varshney et al., 2019a). Fine mapping of 'QTL-hotspot' for drought tolerance-related features for the region of 7.74 Mb–300 kb and chickpea bin mapping were done using genotyping-by-sequencing and skim-sequencing, respectively (Varshney et al., 2014; Jaganathan et al., 2015). A huge range of resources, including genetic, genomic, and transcriptome resources, have been created over the last decade as a result of developments in various NGS technologies, transforming the chickpea crop from an orphan to a genomic-rich resource (Varshney et al., 2009; Kudapa et al., 2014; Agarwal et al., 2016; Mashaki et al., 2018). In chickpea breeding projects, next-generation sequencing, high-throughput genotyping technologies, and cost-effective omics methods are critical. Translational genomics in crop breeding has been made possible by the availability of molecular markers, sequencing platforms, genotyping assays for low-to-high density, quality check panels, draught genome assemblies, and sequence-based genetic variants (Roorkiwal et al., 2014; Thudi et al., 2016; Varshney et al., 2019a; Rasheed et al., 2017; Varshney et al., 2021).

Integration and Impacts of Next-Generation Sequencing Technologies for Improving Chickpea Epigenetics

The advances in next generation sequencing (NGS) technologies have a lead impact on epigenomic research. The arrival of NGS technologies has introduced powerful sequencing methods–like, ChIP-Seq--to interrogate whole-genome histone modifications, improving on the conventional microarray-based method (ChIP-chip). More importantly, studies of DNA methylation and histone modification using NGS technologies have yielded new discoveries in plant biology too. The recent developments of third-generation sequencing technologies have shown promising results of directly sequencing methylated nucleotides and having the ability to differentiate between 5-methylcytosine and 5-hydroxymethylcytosine. The importance of 5-hydroxymethylcytosine remains largely unknown, but it has been found in various tissues. 5-hydroxymethylcytosine was particularly enriched at promoters and in intragenic regions (gene bodies) but was largely absent from non-gene regions in DNA from human brain frontal lobe tissue. The presence of 5-hydroxymethylcytosine in gene bodies was more positively correlated with gene expression levels. The importance of studying 5-methylcytosine and 5-hydroxymethylcytosine separately for their biological roles will become clearer when more efficient methods to distinguish them are available (Ku et al., 2011).

In contrast to histone modification profiling, a wide variety of approaches have been developed to profile DNA methylation utilizing next-generation sequencing platforms. Approaches to profile DNA methylation genome-wide can be broadly divided into those that rely on methylation-dependent enzymatic restriction, methyl-DNA enrichment, and direct bisulfite conversion (Fouse et al., 2010; Harris et al., 2010). Individual methods can also be combined to increase the resolution or efficiency of a single method. For example, a combination of MeDIP-seq and MRE-seq to profile both the methylated and unmethylated fractions of the genome (Maunakea et al., 2010).

Advances in next-generation sequencing (NGS) technology have considerably curtailed sequencing costs resulting in the evolution of genotyping methods from individual marker-to whole-genome sequencing-based genotyping. This has resulted in the development of large-scale genomic resources, including genome sequence assemblies, re-sequencing of a few thousand lines, high-resolution genetic maps, and a range of low-to high-density genotyping platforms. These genomic resources were used to find alleles and haplotypes linked to chickpea agronomic traits (Varshney et al., 2019b). Genetic diversity, population structure, domestication patterns, linkage

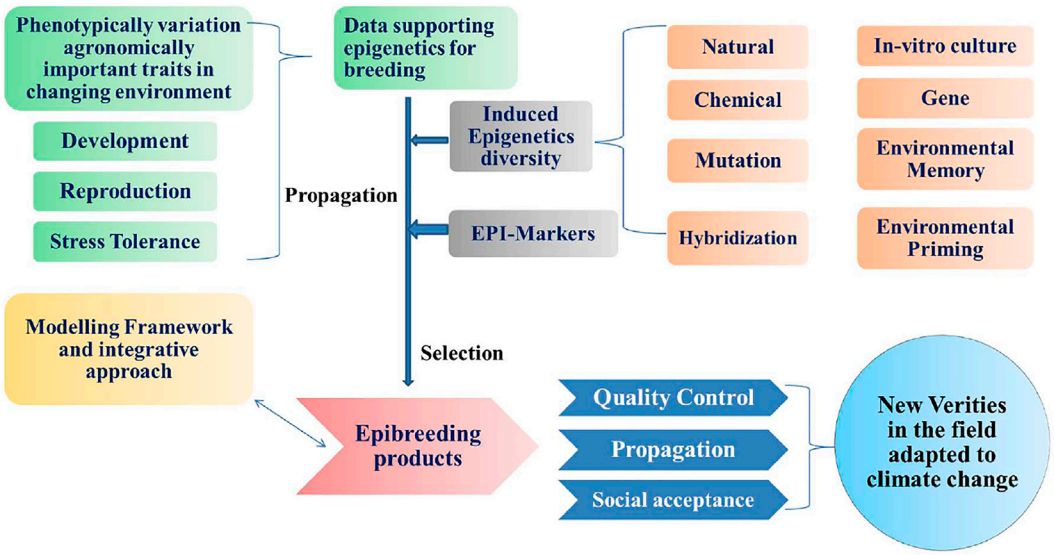


FIGURE 3 | Schematic integrated epigenetic data and tools will lead to epi-bred crops and new varieties in the field adapted to climate change (Modified from Kakoulidou et al., 2021).

disequilibrium, and the untapped genetic potential for chickpea improvement have all been studied using whole-genome re-sequencing (WGRS) (Varshney et al., 2019a). Varshney et al. (2021) conducted a study on molecular diversity in chickpeas to describe genomic diversity across cultivated and wild progenitors. They found chromosomal segments and genes that show signatures of selection during domestication, migration, and improvement. The chromosomal locations of deleterious mutations responsible for limited genetic diversity and decreased fitness were identified in elite germplasm along with the superior haplotypes for improvement-related traits. They found targets for purging deleterious alleles through genomics-assisted breeding and/or gene editing. We can use this sequence information to find the DNA methylation regions that are responsible for biotic and abiotic tolerance in the chickpea in future breeding approaches.

Advanced Technologies Assisted Epigenomics as Key Tools for Climate Resilient Chickpea

Epigenetic mechanisms have proven to have a role in enhancing plants' resilience to environmental stresses, targeting varied traits, thus, giving a significant tool in breeding for climate-resilient crops. Epigenetic variation was applied for crop improvement to increase soybean yield (Raju et al., 2018). RNAi silencing of the plant-specific gene MutS HOMOLOG1 (*MSH1*) paved the way for the growth of epi-lines with variability for the arrangement of yield-associated traits in glasshouse and field trials. New epigenetic diversity indicted by MSH1 oppression was transmitted for at least three progenies. Similarly, the identification of epigenetic variations and regulatory mechanisms in chickpea plants, which impact important agronomic traits, can be exploited for epigenetic breeding for climate-resilient crops. The following schematic presentation as depicted in **Figure 3** for the development of epigenetic data and tools will lead to breed of newer ep-breeds and varieties in the field and adapted to climatic changes.

Crop plants often have challenges of biotic and abiotic stresses, and they adopt sophisticated ways to acclimate and cope with these through the expression of specific genes. Changes in chromatin, histone, and DNA mostly serve the purpose of combating challenges and ensuring the survival of plants in stressful environments. Epigenetic changes, due to environmental stress, enable plants to remember a past stress event in order to deal with such challenges in the future. This heritable memory, called "plant stress memory", enables plants to respond against stresses in a better and more efficient way, not only for the current plant in prevailing situations but also for future generations (Chao et al., 2021). Stress memory can also be described as a mechanism to enhance the resilience of crop plants (Walter et al., 2011), and the accumulation and changes in proteins (structural and regulatory) as transcription, translation, and transduction, which play an important role in the growth, development, and memory mechanisms of plants for stress resistance (Bruce et al., 2007; Janmohammadi et al., 2015, Marcos et al., 2018b). As the epigenetic modifications are environmentally accelerated, the phenotypic changes are mostly a reflection of a specific environmental interaction, and the changes adopted by the plant for a specific period may become permanent and heritable for future generations (Chinnusamy and Zhu, 2009; Verhoeven et al., 2010). Stress memory in plants is enhanced by up and downregulated sRNAs (miRNAs, siRNAs) they mainly downregulate negative regulators, upregulate positive regulators and regulate plant hormones, reactive oxygen species (ROS), and transcriptional factors in response to abiotic stress (Banerjee et al., 2017; Zhang et al., 2019).

Plant stress memory is achieved through the coordination of physiological, translational, transcriptional, and epigenetic activities in response to stress (Kinoshita and Seki, 2014; Hu et al., 2016). These regulatory processes can occur at any stage of plant development and are primarily controlled by epigenetic changes to phenotypically remodel for environmental stress (Rehman et al., 2015; Gallusci et al., 2017). Genetic diversity has been reduced as a result of intense breeding, and now epigenetic variation has arisen as a viable option for crop genetic improvement (Gallusci et al., 2017). There have been many developments for the quantification of epigenetic variations and their impact on the growth and development of plants, leading to improved yield and quality, and ultimately, this has opened another avenue for breeders to breed desirable agronomic characters successfully (Cortijo et al., 2014)]. In epigenetic modifications, DNA methylation plays an important role in gene regulation, expression, and stabilization (Lang et al., 2017). Various enzymes (DNA methyltransferase), targeted under different plant regulatory pathway systems, take part in the process to catalyze DNA methylation for a better and quicker response against biotic and abiotic stresses (Zhang et al., 2018). Epigenetic modification has the ability to memorize the event over a long time as a plant molecular memory and the ability to respond rapidly with heritable phenotypic characteristics as an inheritance system against environmental fluxes. Some extreme abiotic stress treatments can lead to plant genome reorganization (Klumpp et al., 2004; Molinier et al., 2006) but few reports are indicating that short-term stress causes a large number of genomic mutations (Lämke and Bäurle, 2017; Cruzan et al., 2018). More evidence supports the speculation that plant stress memory is mainly regulated by epigenetic pathways (Tang et al., 2014), which means changing the expression pattern of the entire genome to form a rebalanced genome expression system, without changing the genome sequence (Habu et al., 2001; Madlung and Comai, 2004).

Epigenetic variants can be produced through chemical treatment (5-azacytidine), epigenomic editing (TALENs), zinc finger nucleases, and the CRISPR/Cas system to counter biotic and abiotic stresses. There is an immense amount of care needed because targeted genes may be involved in complex and multiple pathways, which may cause complex and unexpected pleiotropic effects (Garg et al., 2015; Hung et al., 2018). All of these methods have tremendous scope for the creation of epigenetic variations (Kapazoglou et al., 2018) For successful breeding through epigenetic memory, it is necessary that variations should be inherited. DNA methylation changes and histone modifications are often reset during meiosis, meaning stable inheritance of the epigenetic mark is a problem in successful breeding goals achievement (Danchin et al., 2019).

Several reports revealed a correlation between the regulation of gene expression and changes in chromatin modifications in plants during stress exposure (Pandey et al., 2016). Epigenetic processes are crucial adaptive mechanisms that change the expression of genes in a heritable way without accompanying changes in DNA sequences (Chinnusamy and Zhu, 2009). Thus, heritable, but simultaneously reversible alterations in the transcriptional potential of cells are possible (Chen et al., 2010b). In a eukaryotic cell, the structure and function of chromatin depend upon several regulatory epigenetics. In plants such as *Arabidopsis thaliana*, *Oryza sativa*, and *Zea mays*, it was recently reported that hyper- or hypomethylation of DNA induced by abiotic stimuli can modulate the expression of stress-responsive genes (Wang et al., 2009; Van Dijk et al., 2010) mechanisms, including DNA methylation and histone modifications (Sahu et al., 2013). Epigenetics has a major role in symbiotic nitrogen fixation in chickpea, reported and experimentally proved by epigenetic regulations in the development of symbiotic root nodules of legume plants–EPISYM project. They have discovered that epigenetic regulations, involving plant DNA (de)methylation and small interfering RNA (siRNA) populations are essential to produce nitrogen-fixing nodules. They hypothesized that epigenetic regulations play an important role in gene expression reprogramming associated with nodule differentiation.

Quantitative Epigenetic Models for Complex Traits

The accurate genetic assays of epigenetic variability and mapping of epigenetic quantitative trait loci (epiQTL) facilitated by the development of epi-RILs in Arabidopsis have provided a close association amongst epialleles and phenotypic characteristics. The epigenetics research in plants has taken a leap by employing epigenome-wide association study (EWAS) and epi-genotyping by sequencing (epi-GBS). Thus, chickpea improvement programs eventually can utilize the huge avenues provided by quantitative epigenetics to assay the contribution of epigenetic variability in trait control. Further, molecular breeding of important crop plants can be potentially facilitated by epigenome-editing tools, such as clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9 (CRISPR/Cas9), for locus-specific DNA methylation (Vijay et al., 2020).

Several statistical methods exist to detect epigenetic variations and their impact on the phenotype or epiQTLs. The significance and accuracy of epiQTLs identification are affected by several factors like recombination, transgressive segregation, instability of epialleles, and parent-off-origin effect. These factors may create confounding effects during epiQTLs analysis and result in false positives or false negatives. To deal with these interrupting factors, Johannes and Colomé-Tatché (2011) advocated as most suitable population generated from the crosses between epigenetic isogenic lines. Tal et al. (2010) deduced covariances amongst kinships owing to epigenetic transmission and environmental effect and modeled the number of events for epigenetic reset amongst generations and environmental inductions, and estimated the heritable epigenetic variance along with the rate of transmission. Furthermore, the necessity of multiple replication-wise testing due to the occurrence of several false positives is the critical bottleneck of quantitative genetics. During genome-wide recognition of epigenetic variation to encounter the bottleneck of false positivity a statistical model was developed (Jaffe et al., 2017). The missing heritability contributed by epigenetic variability may be studied by employing these models but leaving aside epigenetic-

induced phenotypic variability (Roux et al., 2011). Furthermore, another improved model was proposed to predict the proportion of genetic variation and estimate phenotypic variation explained by epigenetic variation and their effects on phenotypic values along with the interaction of genetic effects (additive and dominant) and epigenetics (Wang et al., 2012).

Quantitative Aspects of Epigenetics

Epigenetic markers are randomly present with high frequencies in the genome and are stably inherited through generations. The identification of epiQTLs is facilitated by these characteristics that permit the utilization of epigenetic markers. Unlike QTLs where polymorphism for the DNA nucleotide sequence occurs, epiQTLs are epigenomic loci that differ in cytosine methylation patterns that control phenotypic variability. Cortijo et al (2014) recognized major epiQTLs explaining 60–90% heritability by employing ddm1-derived *Arabidopsis* epi-RILs for quantitative traits root length and flowering time. These epiQTLs were observed to be useful for artificial selection and were reproducible. Furthermore, on the basis of inheritance and recombination events using mutagenic accumulation lines epigenotype map (E-map) was constructed and 99.9% of epialleles were observed to be stable (Hofmeister et al., 2017). In another study employing methylation-sensitive amplified fragment length polymorphism (MS-AFLP) and retro transposon epimarkers epiQTLs for seven agronomic characters were recognized in Brassica (Long et al., 2011). During varied developmental, environmental, and transgenerational states highly stable epigenetic marks were observed. In Sorghum, employing MSAP genotyping approach and 122 methylation polymorphic loci E-map harboring methylation hotspots, was constructed. In soybean, localization of methyl QTL (QTLs associated with DNA methylation) was facilitated by employing differentially co-segregated methylated regions (DMRs) in RILs. Thus, the crop where genetic variability is the negligible stable inheritance of epialleles through the generations makes it a potential controller of phenotypic variability. However, till date very limited number of EWAS have been accomplished in plants, but employing somatic clones (diverse for mantled abnormality and oil yield), a locus MANTLED was localized where hypomethylation in LINE retro transposon pushes the alternate splicing and premature termination epigenetic modification associated with a mantled abnormality in oil palm (Ong-Abdullah et al., 2015).

METHODS TO MODIFY THE PLANT EPIGENOME

Besides the genes inclusive of its genome, the genetic constitution of any organism also contains its epigenome including methyl classes to specific sequences within the DNA that work as epigenetic marks to minimize transcriptions, and thus the expressions of the linked genes. Several methods applied to change the plant's epigenome contained mutagenesis, carcinogenesis, plant tissue culture, CRISPR/Cas9 genome editing, and RNAi, which are explained below.

Role of Mutagenic Agents on the DNA Sequence and Epigenetics in Plants

The mutagenesis and carcinogenesis affecting DNA sequence and chromatin structure through Ethyl methane sulfonate (EMS) mutagenesis and associations with epigenetic changes have been explained (Yan et al., 2021) in rice. Whole-genome and re-sequenced data congregated from 52 rice EMS mutants facilitated mutation for altering DNA sequences and the probable linkages along with chromatin composition. Single nucleotide polymorphic sites (SNPs) along with genomic facets related to EMS anchored mutagenesis prejudices were unraveled. EMS, equated with natural SNPs available in the Rice 3K project, displayed a liking to G/C sites with flanking motifs higher in GC amounts. Efficacies of EMS mutagenesis and constituents of local dinucleotides along with trinucleotides were having associations. The prejudiced allocation of EMS indicted SNPs were affiliated in a positive direction with transposable element quantities, CpG numbers, and suppressive epigenetic markers but linked in the negative direction with active epigenetic markers and gene (s) displaying the euchromatin marker DNase I hypersensitive sites. Another example through which mutations created epigenesis was presented with *Arabidopsis thaliana* mutants originated straightway by changes in DNA methylation affecting transcription of the gene. The late-flowering mutant flowering Wageningen (FWA) created by ectopic demonstration of the FWA gene enciphers a homeodomain-containing transcription facet. In wild type, the escalating region of FWA is methylated DNA and FWA is not produced in vegetative tissues. When this methylated DNA is ousted from the ddm1 mutant, the FWA is noticed in vegetative tissues and causes late flowering. This late-flowering phenotypic form is also noticed in the mutant suggesting that silencing of FWA mainly depends on CG methylation (Soppe et al., 2000). In another recent study done at Indian Agricultural Research Institute, New Delhi, we treated Pusa 372- a high-yielding, and widely grown chickpea cultivar having moderate resistance/tolerance to major diseases with 0.3% EMS for 6 h at room temperature and found mutants with phenotypic variations for the increased number of pods (unpublished).

Role of Tissue Culture on the DNA Sequence and Epigenetics in Plants

Tissue culture techniques are soul for any alteration at the genome level in crops. These techniques are also influenced its epigenome. The high-resolution maps of DNA methylation made in rice lines have reported that the regenerated plants have less methylation than control plants. The alterations were relatively over-represented around the promoter sequences of genes and affect gene expression. Critically, the plants' offshoots also inherit the changes in methylation level (Hume et al., 2013). These aftermaths partly narrate the processes of somaclonal diversities that push forward epigenetic changes in the plants.

Role of RNAi Techniques in Genome Epigenesis in Plants

Gene silencing through RNA inference (RNAi) approach was widely used to better understand the gene function in plants. Repression of translation was achieved through post-transcriptional modifications using RNAi. Interestingly, many of the factors that mediate post-transcriptional silencing via RNAi also contribute to transcriptional gene suppression (Fire et al., 1998). In plants, RNA viruses were observed to guide DNA methylation of homologous genes along with introducing multiple transgene copies resulting in silencing (Napoli et al., 1990; Van der Krol et al., 1990).

Role of CRISPR/Cas9 in Genome Editing in Plants

The ribonucleoproteins (RNPs), containing Cas9 enzyme along with single-guide RNA (sgRNA) are successfully delivered using transformation methods or nanoparticle-based delivery approaches. The prime enzyme 4-coumarate ligase (4CL) involved in phenylpropanoid metabolism and responsible for the lignin biosynthesis process governs the congregation of lignin in distress stages. The 4CL along with the gene Reveille 7 (RVE7) linked with drought tolerance has been used for protoplast targeted mutagenesis in chickpeas. For the first time, chickpea protoplast was used as a transfection platform for CRISPR/Cas9-based genome editing in chickpeas (Badhan et al., 2021). The outcomes showed efficient editing got for the RVE7 gene *in vivo* compared with the 4CL gene.

Understanding genomic activities need site-specific modification at the loci *via* targeting systems (Papikian et al., 2019). Limited approaches for the desired manipulation of the epigenome present in plants were observed and adopted by the Cas9-system to design desired gene activation and DNA methylation in Arabidopsis.

CONNOTATIONS OF EPIGENETICS AND EPIGENOMICS IN CHICKPEA IMPROVEMENT

Epigenetics and epigenomics display certain connotations as useful immense potentials along with numerous threats and challenges as stated by Springer and Schmitz (2017). Some of the potential connotations for utilization are summarized as given below.

Resistance to Biotic and Abiotic Stresses

As we mentioned above, the role of epigenetics in disease resistance (Sen et al., 2017), cold tolerance (Rakei et al., 2015), drought, salinity tolerance (Khandal et al., 2017), and manipulating the epigenome may provide a promising breeding strategy to enhance yield, disease resistance, or adaptation for changing environmental conditions in chickpea, as shown below.

i. DNA methylation patterns in cultivated chickpeas to understand the regulation of gene expression in different organs (Bhatia et al., 2018)
ii. Drought and salinity resistance by an epigenetic mechanism in chickpeas (Sen et al., 2017)
iv. DNA methylation and epigenetics mechanism on physio-biochemical analysis of chickpea in response to cold stress (Rakei et al., 2015)
v. The epigenetic mechanism to heat stress in chickpeas (Chidambaranathan et al., 2016)
vi. Role of epigenetics in drought yield index (Sharma et al., 2019)

Avoiding the Transgene Silencing in GM Crops

Transgene methylation and transcriptional gene mutations are directly correlated to each other (Matzke et al., 1989; Park et al., 1996) mainly because of the association between methylation of the coding sequence and post-transcriptional gene oppression (Ingelbrecht et al., 1994). Although, the latest proof shows that a merging process gleaned from RNA interference is primary to both processes (Matzke and Matzke, 2004). The intricate and meticulous designing of the transgene constructs and intense dissection of transformants at the molecular level are the prerequisites of an efficient technique to avoid transgene silencing (Dewilde et al., 2000). Two dominant classes of transgene silencing, the first one results in position effects (Matzke et al., 2000) and the second one is silencing phenomena or homology-oriented gene silencing, HDGS (Meyer and Saedler, 1996). Some examples reported in plants are tobacco, transgenic tobacco (*N. tabacum*) that are constructed ectopically over-express *AtMYB90v* (*Arabidopsis thaliana MYB 90*) promoter gene in association with regulating anthocyanin production in *Arabidopsis thaliana*. Transgenic tobacco overexpressing *AtMYB90* involved in anthocyanin biosynthesis showed siRNA-mediated silencing because of systemic acquired silencing (Velten et al., 2012).

Evolution

Although epigenetics in multicellular organisms is the major mechanism for diversifications, with epigenetic motifs "reset" when organisms procreate, there were certain reflections of trans-generational epigenetic transmission, e.g., the phenomenon of para-mutation in maize (Brink, 1956). Epigenetic characters are multigenerational and eventually diminished over many generations. Then, there is a clear-cut maximum probability for explaining another aspect of evolution and adaptation. There are some speculations that the differential mutation rates associated with epigenetic features were taken as an advantage by the organisms that control the mutation rates of particular genes. Epigenetic changes have also been reflected to originate in reaction to environmental exposure; for example, epigenetic alterations are prevalent in inter-specific hybrids and polyploids. DNA methylation patterns after hybridization and/or polyploidization can be primarily changed by these re-patterning processes, as exemplified by studies in *Brassica, Arabidopsis, Triticum*, and *Oryza*. In these species, methylation-influenced AFLP assays provided widespread alterations in genomic methylation, including modifications in genes (Liu and Wendel, 2003).

A study unraveled DNA methylation systems in cultivated chickpea to explain the control of gene expression in variegated organs primarily by the methylating systems in leaf tissue of wild and cultivated chickpea. The results show a positive association of promoter hyper-methylation with increased transcript paucity through recognition of DMR of the genes governing flower development meant in cultivated chickpeas (Bhatia et al., 2018).

Genetic Variability

Epigenetic patterns in plants, once instituted, can be transmitted through the inheritance of epialleles across many generations (Kakutani, 2002). Such transmittable epigenetic alleles can be assumed as a novel source of polymorphism and may reproduce new phenotypes. Evaluating the significance of methylated epialleles in crop breeding requires the genetic variability in the selected population for the degrees that methylating modes influence superior phenotypes and the extent to which methylation is statically transmitted. DNA methylation was first reported in regeneration studies of crown gall tumor events in which phenotypic variability and methylation of T-DNA were linked (John and Amasino, 1989). The most interesting evidence suggests that a substantial proportion of somaclonal variation might be because of diverse, pre-existing epigenetic states' result in the regeneration of individual somatic cells (Neuhuber et al., 2005).

Epigenetic initiation of DNA elements through transposable elements with *Arabidopsis* suggests epigenetic modifications may also be intricated in cytogenetic instability through changes of heterochromatin, and as a basis of phenotypic diversity through the modulation of gene functionalities (Kaeppler et al., 2000).

Epigenetics plays an important role in somaclonal variation, and chromatin modulation plays an important role in gene expression regulation and genome activities (Azizi et al., 2020). Some epigenetic modifications that induced intergenerational distress memory resistance in crop plants in addition to as presented in **Table 5** are as below.

a) Chickpea-drought water and osmotic stress (Elkoca et al., 2007; Kilian et al., 2007)
b) Canola–Salt/drought, seed priming with NaCl, increased energy efficient utilization, and PGPR for halo-tolerant plant (Farhoudi et al., 2007)
c) Sugarcane–Drought/salinity, NaCl, and PEG-primed seeds (Marcos et al., 2018a)

Heterosis

Heterosis is the superiority of the F_1 hybrid phenotype over its parents. The phenomenon has been exploited extensively in agricultural breeding for decades and, despite its commercial impact; it has also improved crop performance tremendously. However, knowledge of the molecular basis underlying heterosis remains incomplete. Most studies have focused on finding genetic explanations, resulting in the classical dominance and overdominance models of heterosis (East, 1908; Shull, 1908; Bruce, 1910; Jones, 1917). Identification of better hybrids through the utilization of hybrid vigor in chickpea by assessing seven F1 hybrids inclusive of nine cultivars was executed (Ghfaffar et al., 2015). Paramount heterosis along with heterobeltiosis for plant height and subsidiary branches were observed in the cross K0014–10 × K0066-10, however, cross K0019–10 × K0031-10 performed the highest heterosis along with heterobeltiosis for principal branches and seed number plant^{-1}, the cross K0014–10 × K0052-10 reflected the highest heterosis and heterobeltiosis with 33.18 and 30.84% for 100 seed weight and 97.37 and 76.47% for seed yield plant^{-1}, discretely. Broad sense heritability for various characters observed varied from 63.14 to 77.18%. Remarkable heterosis, heritability, and genetic advance were recorded for pod number plant^{-1} that could be employed for identifying best segregate from crosses K0031–10 × K0052–10, K0019-10 × K0026-10, and K0019–10 × K0031-10. Best hybrids from the observation could be employed for the betterment of multiple traits by identifying single plants for varied characteristics.

Hybridization and Epigenetic/Epigenomic as Predictive Markers for Hybrid Performance

Molecular profiling of superior hybrids reflected that their epigenomes are substantially remodeled to their parental lines, leading to epigenetic states that deviate from the expected mid-parent values. Extensive remodeling has been observed at the level of DNA methylation in *Arabidopsis* (Zhang et al., 2016), rice (He et al., 2010), pigeon pea (Zhang et al., 2006), broccoli (Li et al., 2018a), and rapeseed (Shen et al., 2017). It occurs either at regions where parents are differentially methylated (DMRs).

POTENTIAL CHALLENGES TO EPIGENETICS AND EPIGENOMICS

Irrespective of the immense potential of epigenetics for opening new avenues for utilization in crop improvement programs, there are certain challenges and threats as stated below.

A. Determination of Epigenetic State: The first challenge is to clearly define the basis of an epigenetic state. Whereas a DNA sequence is simply defined by the order of the four bases (A, C, G, and T), the exhaustive list of components that define given chromatin or epigenetic state is yet to be established. These components include methylation of cytosines and adenines, mono-, di- or tri-methylation, acetylation, phosphorylation, ubiquitination, etc of histones at various positions (e.g., H2AK119, H3K4, H3K9, H3K27, H3K36, etc) and long or short ncRNAs produced in cis or trans.
B. Determination of Chromatin Stability: The second important challenge is to define the stability of given chromatin. Three major levels of stability can be distinguished, as given below:

Transient Chromatin States: These chromatin states are specific to a different cell or established in immediate response to biotic or abiotic stress, and do not persist after the stimulus is removed.

Metastable Epigenetic States: These epigenetic states are started by specific stress or environmental inductions and can persist across multiple cell divisions after induces.

Inherited Epigenetic States: These epigenetic states are transmitted across multiple generations and are typically correlated with TEs or other repeat sequences. It is still unclear what role the environment plays in initiating or erasing these states.

Gene and environmental interactions as epigenetics are influenced by the environment and sometimes it leads to non-stable variations.

Epigenome sequencing methods are not well established.

FUTURE PERSPECTIVE: EXPLORING EPIGENOMICS AS A NEW KIND OF PLANT BREEDING

Epigenetics has immense potential for opening new avenues for crop improvement programs stated as followings:

a) Variation: Considerable natural variability in the DNA methylation process exists within many plant species.
b) Stable Inheritance: Variability in DNA methylation can originate through processes and clonal propagation can propel epigenetic alleles.
c) Epigenome Engineering: New epigenome editing tools provide broader opportunities to create new epiallelic variants by altering the methylation of DNA or other modifications at the chromosome level. These tools can be used for crop improvement through epigenome engineering.
d) Emerging New Technologies: The development and application of methods for widespread epigenome profiling and engineering may generate new avenues for using the full potential of epigenetics in crop improvement.
e) Sources for Biotic and Abiotic Resistance: Epigenetics has become an important research focus at a time when rapid environmental changes are occurring. They enhance fitness extremely rapidly without depending on the slower process of natural selection through changing DNA-encoded genetic variants in plant populations.
f) Time and Cost-Effective: Epigenetics introduces as a time- and cost-effective tool in plants as a source of resistance against new future abiotic and biotic stresses.
g) Public and Producer's Acceptance: Successful implementation of all crop enhancement approaches at the DNA level requires support from the public and government and epigenome editing does not change the genome sequence might ease the challenges of public acceptance for epigenetically modified products.
h) Equilibrium among important agronomic traits: Plants use epigenetic variation to reprogram their transcriptome in a precise and timely manner to maintain equilibrium amongst important agronomic traits.

CONCLUSION

Climate change is altering the predominance of varied environmental situations, and improved distress tolerance has become a primary breeding goal in chickpeas. *In vivo* situations, crops are usually concomitantly opposed by diverse biotic and abiotic distresses. Hence, grasping possible processes responsible for the occurrence of stresses has become a necessity for stable crop productivity. The epigenetic mechanisms play an important role in a classical plant breeding program, mainly by genetic heritability, hybrid vigor, plant-environment interactions, abiotic and abiotic stress tolerance, and yield stability performance of crop plants. A better and deep insight into epigenetic mechanisms might facilitate plant breeders in creating novel and more super crop varieties that can include natural phenotypic diversity. It is a very interesting fact that the environmental shielding effects of epigenetics are directly associated with those genes that play a very important role in the regulation of plant growth and yield in chickpeas and other crops. Furthermore, understanding the molecular bottom of trans-generational epigenetic transmission puts forward the development of epialleles identified for specific environmental status through combined and multidisciplinary efforts of researchers and targeted epigenetic modifications in genes of interest. Thus, epigenomics, either as exploitation of existing epigenomic variability or alteration of the epigenome, can complement conventional plant breeding to ensure global food security and sustainable agriculture.

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

RK conceptualized and supervised the manuscript writing. BC, RM, and RS collected the related literature and contributed to the original writing. RK, RF, NV, SG, HY, and MM extended their help in the inference, review, and editing of the manuscript. All authors went through the final manuscript draft and approved it.

FUNDING

The study was supported by ICAR-Indian Agricultural Research Institute, New Delhi and DST-Science and Engineering Research Board through sanction order No.CRG2019006273. DBT-National Institute of Plant Genome Research, New Delhi, India; CSIR-National Botanical Research Institute, Lucknow, India; Griffith University, Australia; University of Southern Queensland, Queensland, Australia and Cornell University, Ithaca, New York, for providing necessary support for the study.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the ICAR-Indian Agricultural Research Institute, New Delhi, India.

REFERENCES

Agarwal, G., Garg, V., Kudapa, H., Doddamani, D., Pazhamala, L. T., Khan, A. W., et al. (2016). Genome-wide Dissection of AP2/ERF and HSP90 Gene Families in Five Legumes and Expression Profiles in Chickpea and Pigeonpea. *Plant Biotechnol. J.* 14, 1563–1577. doi:10.1111/pbi.12520

Amaral, J., Ribeyre, Z., Vigneaud, J., Sow, M. D., Fichot, R., Messier, C., et al. (2020). Advances and Promises of Epigenetics for Forest Trees. *Forests* 11, 976. doi:10.3390/f11090976

Amin, U. S. M., Biswas, S., Elias, S. M., Razzaque, S., Haque, T., Malo, R., et al. (2016). Enhanced Salt Tolerance Conferred by the Complete 2.3 Kb cDNA of the Rice Vacuolar Na+/H+ Antiporter Gene Compared to 1.9 Kb Coding Region with 5 ' UTR in Transgenic Lines of Rice. *Front. Plant Sci.* 7, 14. doi:10.3389/fpls.2016.00014

Arefian, M., Vessal, S., Malekzadeh-Shafaroudi, S., Siddique, K. H. M., and Bagheri, A. (2019). Comparative Proteomics and Gene Expression Analyses Revealed Responsive Proteins and Mechanisms for Salt Tolerance in Chickpea Genotypes. *BMC Plant Biol.* 19, 300. doi:10.1186/s12870-019-1793-z

Arumuganathan, K., and Earle, E. D. (1991). Nuclear DNA Content of Some Important Plant Species. *Plant Mol. Biol.* 9, 208–218. doi:10.1007/bf02672069

Azizi, P., Hanafi, M. M., Sahebi, M., Harikrishna, J. A., Taheri, S., Yassoralipour, A., et al. (2020). Epigenetic Changes and Their Relationship to Soma Clonal Variation: a Need to Monitor the Micro Propagation of Plantation Crops. *Funct. Plant Biol.* 47, 508–523. doi:10.1071/FP19077

Badhan, S., Ball, A. S., and Mantri, N. (2021). First Report of CRISPR/Cas9 Mediated DNA-free Editing of 4CL and RVE7 Genes in Chickpea Protoplasts. *Int. J. Mol. Sci.* 22, 396. doi:10.3390/ijms22010396

Banerjee, S., Sirohi, A., Ansari, A. A., and Gill, S. S. (2017). Role of Small RNAs in Abiotic Stress Responses in Plants. *Plant gene.* 11, 180–189. doi:10.1016/j.plgene.2017.04.005

Bannister, A., and Kouzarides, T. (2011). Regulation of Chromatin by Histone Modifications. *Cell. Res.* 21, 381–395. doi:10.1038/cr.2011.22

Barlow, D. P., and Bartolomei, M. S. (2014). Genomic Imprinting in Mammals. *Cold Spring Harb. Perspect. Biol.* 6, a018382. doi:10.1101/cshperspect.a018382

Barra, V., Schillaci, T., and Lentini, L. (2012). Bypass of Cell Cycle Arrest Induced by Transient DNMT1 Post-transcriptional Silencing Triggers Aneuploidy in Human Cells. *Cell. Div.* 7, 2. doi:10.1186/1747-1028-7-2

Baulcombe, D. C. (2006). Short Silencing RNA: The Dark Matter of Genetics? *Cold Spring Harb. Symp. Quant. Biol.* 71, 13–20. doi:10.1101/sqb.2006.71.052

Bazin, J., Khan, G. A., Combier, J. P., Bustos-Sanmamed, P., Debernardi, J. M., Rodriguez, R., et al. (2013). miR396 Affects Mycorrhization and Root Meristem Activity in the Legume *Medicago Truncatula*. *Plant J.* 74 (6), 920–934. doi:10.1111/tpj.12178

Becker, P. B., and Workman, J. L. (2013). Nucleosome Remodelling and Epigenetics. *Cold Spring Harb. Perspect. Biol.* 5, 017905. doi:10.1101/cshperspect.a017905

Benedito, V., Torez-Jerez, I., Murray, J., Andriankaja, A., Allen, S., Kakar, K., et al. (2008). A Gene Expression Atlas of the Model Legume *Medicago Truncatula*. *Plant J.* 55, 504–513. doi:10.1111/j.1365-313X.2008.03519.x

Berger, F. (2006). Endosperm: the Crossroad of Seed Development. *Curr. Opin. Plant Biol.* 6, 42–50. doi:10.1016/s1369526602000043

Berger, S. L., Kouzarides, T., Shiekhattar, R., and Shilatifard, A. (2009). An Operational Definition of Epigenetics. *Genes. & Dev.* 23, 81–783. doi:10.1101/gad.1787609

Berr, A., Shafiq, S., and Shen, W. H. (2011). Histone Modifications in Transcriptional Activation during Plant Development. *Biochim. Biophys. Acta.* 1809, 567–576. doi:10.1016/j.bbagrm.2011.07.001

Bhatia, H., Khemka, N., Jain, M., and Garg, R. (2018). Genome-wide Bisulphite-Sequencing Reveals Organ-specific Methylation Patterns in Chickpea. *Sci. Rep.* 8, 9704. doi:10.1038/s41598-018-27979-w

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. Prot.* 6, 1868–1884. doi:10.1074/mcp.M700015-MCP200

Bird, A. (2007). Perceptions of Epigenetics. *Nature* 447, 396–398. doi:10.1038/nature05913

Black, J. C., Vanrechem, C., and Whetstine, J. R. (2012). Histone Lysine Methylation Dynamics: Establishment, Regulation, and Biological Impact. *Mol. Cell.* 48, 491–507. doi:10.1016/j.molcel.2012.11.006

Blewitt, M., and Whitelaw, E. (2013). The Use of Mouse Models to Study Epigenetics. *Cold Spring Harb. Perspect. Biol.* 5, 017939. doi:10.1101/cshperspect.a017939

Boucelha, L., and Djebbar, R. (2015). Influence of Different Pre-germination Treatments of Vigna Unguiculata (L.) Walp. Seeds on Germination Performance and Water Stress Tolerance. *Biotechnol. Agron. Société Environ.* 19, 160–172.

Bratzel, F., Yang, C., Angelova, A., López-Torrejón, G., Koch, M., del Pozo, J. C., et al. (2012). Regulation of the New *Arabidopsis* Imprinted Gene AtBMI1C Requires the Interplay of Different Epigenetic Mechanisms. *Mol. Plant* 5, 260–269. doi:10.1093/mp/ssr078

Brink, R. A. (1956). A Genetic Change Associated with the R Locus in Maize Which Is Directed and Potentially Reversible. *Genetics* 41, 872–889. doi:10.1093/genetics/41.6.872

Bruce, A. B. (1910). The Mendelian Theory of Heredity and the Augmentation of Vigor. *Science* 32, 627–628. doi:10.1126/science.32.827.627-a10.1126/science.32.827.627.b

Bruce, J. A., Matthes, M., C., Napier, J. A., and Pickett, J. A. (2007). Stressful "Memories" of Plants: Evidence and Possible Mechanisms. *Plant Sci.* 173, 603–608. doi:10.1016/j.plantsci.2007.09.002

Brzeski, J., and Jerzmanowski, A. (2003). Deficiency in DNA Methylation 1 (DDM1) Defines a Novel Family of Chromatin-Remodeling Factors. *J. Biol. Chem.* 278, 823–828. doi:10.1074/jbc.M2092602010.1074/jbc.m209260200

Burggren, W. (2016). Epigenetic Inheritance and its Role in Evolutionary Biology: Re-evaluation and New Perspectives. *Biology* 5 (2), 24. doi:10.3390/biology5020024

Burggren, W. (2020). Phenotypic Switching Resulting from Developmental Plasticity: Fixed or Reversible? *Front. Physiol.* 10. doi:10.3389/fphys.2019.01634

Bustos-Sanmamed, P., Mao, G., Deng, Y., Elouet, M., Khan, G. A., Bazin, J. R. M., et al. (2013). Overexpression of miR160 Affects Root Growth and Nitrogen-Fixing Nodule Number in Medicago Truncatula. *Funct. Plant Biol.* 40, 1208–1220. doi:10.1071/FP13123

Chakraborty, J., Ghosh, P., Sen, S., and Das, S. (2018). Epigenetic and Transcriptional Control of Chickpea WRKY40 Promoter Activity under *Fusarium* Stress and its Heterologous Expression in *Arabidopsis* Leads to Enhanced Resistance against Bacterial Pathogen. *Plant Sci.* 276, 250–267. doi:10.1016/j.plantsci.2018.07.014

Chandler, V. L., Eggleston, W. B., and Dorweiler, J. E. (2010). Paramutation in Maize. *Plant Mol. Biol.* 43, 121–145. doi:10.1023/a:1006499808317

Chapman, E. J., and Carrington, J. C. (2007). Specialization and Evolution of Endogenous Small RNA Pathways. *Nat. Rev. Genet.* 884–896. doi:10.1038/nrg2179

Chao, S., Kazim, A., Kan, Y., Sajid, F., and Richard, D. (2021). Plants Exploration of Epigenetics for Improvement of Drought and Other Stress Resistance in Crops. *Rev. Plants* 10 (6), 1226. doi:10.3390/plants10061226

Chen, H. M., Chen, L. T., Patel, K., Li, Y. H., Baulcombe, D. C., and Wu, S. H. (2010b). 22-Nucleotide RNAs Trigger Secondary siRNA Biogenesis in Plants. *Proc. Natl. Acad. Sci. U. S. A.* 107, 15269–15274. doi:10.1073/pnas.10017310.1073/pnas.1001738107

Chen, M., Lv, S., and Meng, Y. (2010a). Epigenetic Performers in Plants. *Dev. Growth Differ.* 52, 555–566. doi:10.1111/j.1440-169x.2010.01192.x

Chen, R., Li, M., Zhang, H., Duan, L., Sun, X., Jiang, Q., et al. (2019). Continuous Salt Stress-Induced Long Non-coding RNAs and DNA Methylation Patterns in Soybean Roots. *BMC Genom* 20, 730. doi:10.1186/s12864-019-6101-7

Chen, Z. H., Bao, M. L., Sun, Y. Z., Yang, Y. J., Xu, X. H., Wang, J. H., et al. (2011). Regulation of Auxin Response by miR393-Targeted Transport Inhibitor Response Protein 1 Is Involved in Normal Development in Arabidopsis. *Plant Mol. Biol.* 77, 619–629. doi:10.1007/s11103-011-9838-1

Chidambaranathan, P., Jagannadham, P. T. K., Satheesh, V., Jain, P. K., and Srinivasan, R. (2016). Expression Analysis of Six Chromatin and Remodeling Complex Genes (SWR1) in Chickpea in Different Tissues and during Heat Stress. *Indian J. Genet.* 76, 47–56. doi:10.5958/0975-6906.2016.00007.9

Chinnusamy, V., and Zhu, J. K. (2009). Epigenetic Regulation of Stress Responses in Plants. *Curr.Opin. Plant Biol.* 12, 133–139. doi:10.1016/j.pbi.2008.12.006

Choi, H. K. (2019). Translational Genomics and Multi-Omics Integrated Approaches as a Useful Strategy for Crop Breeding. *Genes. Genom* 41 (2), 133–146. doi:10.1007/s13258-018-0751-8

Collins, B. E., Greer, C. B., and Coleman, B. C. (2019). Histone H3 Lysine K4 Methylation and its Role in Learning and Memory. *Epigenetics Chromatin* 12, 7. doi:10.1186/s13072-018-0251-8

Cortijo, S., Wardenaar, R., Colome-Tatche, M., Gilly, A., Etcheverry, M., Labadie, K., et al. (2014). Mapping the Epigenetic Basis of Complex Traits. *Science* 343, 1145–1148. doi:10.1126/science.1248127

Cruzan, M. B., Streisfeld, M. A., and Schwoch, J. A. (2018). Phenotypic Effects of Somatic Mutations Accumulating during Vegetative Growth. *bioRxiv*. doi:10.1007/s10682-022-10188-3

Cubas, P., Vincent, C., and Coen, E. (1999). An Epigenetic Mutation Responsible for Natural Variation in Floral Symmetry. *Nature* 401, 157–161. doi:10.1038/43657

Danchin, E., Pocheville, A., Rey, O., Pujol, B., and Blanchet, S. (2019). Epigenetically Facilitated Mutational Assimilation: Epigenetics as a Hub within the Inclusive Evolutionary Synthesis. *Biol. Rev.* 94, 259–282. doi:10.1111/brv.12453

Deal, R. B., and Henikoff, S. (2011). Histone Variants and Modifications in Plant Gene Regulation. *Curr.Opin. Plant Biol.* 14, 116–122. doi:10.1016/j.pbi.2010.11.005

Dela, R. C., Covarrubias, A. A., and Reyes, J. L. (2019). A Dicistronic Precursor Encoding miR398 and the Legume-specific miR2119 Coregulates CSD1 and ADH1 mRNAs in Response to Water Deficit. *Plant, Cell. & Environ.* 42, 133–144. doi:10.1111/pce.13209

Dewilde, C., Vanhoudt, H., and Debuck, S. (2000). Plants as Bioreactors for Protein Production: Avoiding the Problem of Transgene Silencing. *Plant. Mol. Biol.* 43, 347–359. doi:10.1023/a:1006464304199

Dias, D. A., Hill, C. B., Jayasinghe, N. S., Atieno, J., Sutton, T., and Roessner, U. (2015). Quantitative Profiling of Polar Primary Metabolites of Two Chickpea Cultivars with Contrasting Responses to Salinity. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* 1000, 1–13. doi:10.1016/j.jchromb.2015.07.002

Ding, D., Zhang, L., Wang, H., Liu, Z., Zhang, Z., and Zheng, Y. (2009). Differential Expression of miRNAs in Response to Salt Stress in Maize Roots. *Ann. Bot.* 3 (1), 29–38. doi:10.1093/aob/mcn205

Dowen, R. H., Mattia, P., Robert, J. S., Ryan, L., Jill, M. D., Joseph, R. N., et al. (2012). Widespread Dynamic DNA Methylation in Response to Biotic Stress. *PNAS* 109, 12858–12859. doi:10.1073/pnas.1209329109

East, E. M. (1908). "Inbreeding in Corn," in *Report of the Connecticut Agricultural Experiment Station for the Year 1907* (New Haven: Connecticut Agricultural Experiment Station), 419–428.

Elgin, S. C. R., and Reuter, G. (2013). Position Effect Variegation, Heterochromatin Formation, and Gene Silencing in Drosophila. *Cold Spring Herb. Perspect. Biol.* 5, 017780. doi:10.1101/cshperspect.a017780

Elkoca, E., Haliloglu, K., Esitken, A., and Ercisli, S. (2007). Hydro- and Osmopriming Improve Chickpea Germination. *Acta. Agric. Scand. Sect. b. Soil Plant Sci.* 57, 193–200. doi:10.1080/09064710600914087

Elmayan, T., Proux, F., and Vaucheret, H. (2005). *Arabidopsis* RPA2: A Genetic Link Among Transcriptional Gene Silencing, DNA Repair, and DNA Replication. *Curr. Biol.* 15, 1919–1925. doi:10.1016/j.cub.2005.09.044

Eskandari, H., and Kazemi, K. (2011). Effect of Seed Priming on Germination Properties and Seedling Establishment of Cowpea (Vigna Sinensis). *Not. Sci. Biol.* 3, 113–116. doi:10.15835/nsb346338

Eyles, R. P., Williams, P. H., Ohms, S. J., Weiller, G. F., Ogilvie, H. A., Djordjevic, M. A., et al. (2013). microRNA Profiling of Root Tissues and Root Forming Explant Cultures in *Medicago Truncatula*. *Planta* 238, 91–105. doi:10.1007/s00425-013-1871-7

Farhoudi, R., Sharifzadeh, F., Poustini, K., Makkizadeh, M. T., and Kochakpor, M. (2007). The Effects of NaCl Priming on Salt Tolerance in Canola (*Brassica Napus*) Seedlings Grown under Saline Conditions. *Seed Sci. Technol.* 35, 754–759. doi:10.15258/sst.2007.35.3.23

Feng, S., Cokus, S. J., Zhang, X., Chen, P. Y., Bostick, M., Goll, M. G., et al. (2010). Conservation and Divergence of Methylation Patterning Inplants and Animals. *Proc. Natl. Acad. Sci.* 107, 8689–8694. doi:10.1073/pnas.1002720107

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998). Potent and Specific Genetic Interference by Double-Stranded RNA in *Caenorhabditis elegans*. *Nature* 391806811. doi:10.1038/35888

Flowers, T. J., Galal, H. K., and Bromham, L. (2010). Evolution of Halophytes: Multiple Origins of Salt Tolerance in Land Plants. *Funct. Plant Biol.* 37 (7), 604–612. doi:10.1071/FP09269

Fouse, S., D., Nagarajan, R., P., and Costello, J. F. (2010). Genome-scale DNA Methylation Analysis. *Epigenomics* 2, 105–117. doi:10.2217/epi.09.35

Gallusci, P., Dai, Z., Génard, M., Gauffretau, A., Leblanc-Fournier, N., Richard-Molard, C., et al. (2017). Epigenetics for Plant Improvement. *Curr. Knowl. Model. Ave. Trends Plant Sci* 22, 610–623. doi:10.1016/j.tplants.2017.04.009

Garg, R., Narayana, C. V., Shankar, R., and Jain, M. (2015). Divergent DNA Methylation Patterns Associated with Gene Expression in Rice Cultivars with Contrasting Drought and Salinity Stress Response. *Sci. Rep.* 5, 14922. doi:10.1038/srep14922

Garg, R., Shankar, R., Thakkar, B., Kudapa, H., Krishnamurthy, 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

Garg, R., Tiwari, R., Tiwari, S., and Goyal, S. (2014). Genomic Survey, Gene Expression Analysis and Structural Modeling Suggest Diverse Roles of DNA Methyltransferases in Legumes. *PLoS One* 9, e88947. doi:10.1371/journal.pone.0088947

Ghfaffar, A., Hussain, N., Asiam, M., Hussain, K., and Israd, M. (2015). Identification of Superior Hybrid through Exploitation of Hybrid through Exploitation of Hybrid Vigour in Chickpea (*Cicer Arietinum* L.). *Int. J. Adv. Res. Biol.Sci.* 2 (2), 180–184.

González, P. H., and Álvarez, M. P. (2013). Epigenetics and its Implications for Psychology. *Psicothema* 25 (1), 3–12. doi:10.7334/psicothema.2012.327

González-Benito, M., E., Ibáñez, M. Á., Pirredda, M., Mira, S., and Martín, C. (2020). Application of the MSAP Technique to Evaluate Epigenetic Changes in Plant Conservation. *Int. J. Mol. Sci.* 21, 7459. doi:10.3390/ijms21207459

Graf, P., Dolzblasz, A., Wuerschum, T., Lenhard, M., Pfreundt, U., and Laux, T. (2010). MGOUN1 Encodes an Arabidopsis Type IB DNA Topoisomerase Required in Stem Cell Regulation and to Maintain Developmentally Regulated Gene Silencing. *Plant Cell.* 22, 716–728. doi:10.1105/tpc.109.068296

Guo, H. S., Xie, Q., Feia, J. F., and Chuad, N. H. (2005). MicroRNA Directs mRNA Cleavage of the Transcription Factor NAC_1 to Downregulate Auxin Signals for Arabidopsis Lateral Root Development. *Plant Cell.* 17 (5), 1376–1386. doi:10.1105/tpc.105.030841

Gutierrez, L., Bussell, J. D., Pacurar, D. I., Schwambach, J., Pacurar, M., and Bellini, C. (2009). Phenotypic Plasticity of Adventitious Rooting in Arabidopsis Is Controlled by Complex Regulation of AUXIN RESPONSE FACTOR Transcripts and microRNA Abundance. *Plant Cell.* 21, 3119–3313. doi:10.1105/tpc.108.064758

Haag, J. R., and Pikaard, C. S. (2011). Multi Subunit RNA Polymerases IV and V: Purveyors of Non-coding RNA for Plant Gene Silencing. *Nat. Rev. Mol. Cell. Biol.* 12, 483–492. doi:10.1038/nrm3152

Habu, Y., Kakutani, T., and Paszkowski, J. (2001). Epigenetic Developmental Mechanisms in Plants: Molecules and Targets of Plant Epigenetic Regulation. *Curr. Opin. Genet. Dev.* 11, 215–220. doi:10.1016/s0959-437x(00)00182-9

Hajyzadeh, M., Turktas, M., Khawar, K., and Unver, T. (2015). miR408 over Expression Causes Increased Drought Tolerance in Chickpea. *Gene* 555, 186–193. doi:10.1016/j.gene.2014.11.002

Hand, M. L., and Koltunow, A. M. (2014). The Genetic Control of Apomixis: Asexual Seed Formation. *Genetics* 197, 441–450. doi:10.1534/genetics.114.163105

Harris, R. A., Ting, W., and Cristian, C. (2010). Sequence-based Profiling of DNA Methylation: Comparisons of Methods and Catalogue of Allelic Epigenetic Modifications. *NBiotechnol* 28, 1097–1105. doi:10.1200/jco.2009.27.7111

He, G., Zhu, X., Elling, A. A., Chen, L., Wang, X., Guo, L., et al. (2010). Global Epigenetic and Transcriptional Trends Among Two Rice Subspecies and Their Reciprocal Hybrids. *Plant Cell.* 22, 17–33. doi:10.1105/tpc.109.072041

Heitz, E. (1929). Heterochromatin, Chromocentren, Chromomeren.(*Vorlaufige Mitteilung.*). *erichte Dtsch. Bot. Schen. Gesell. Schaft.* 47 (4), 274–284. doi:10.1111/j.1438-8677.1929.tb01609.x

Henikoff, S., and Smith, M. M. (2015). Histone Variants and Epigenetics. *Cold Spring Harb. Perspect. Biol.* 125 (3), 346–364. doi:10.1101/cshperspect.a019364

Herr, A. J. (2005). Pathways through the Small RNA World of Plants. *FEBS Lett.* 579, 5879–5888. doi:10.1016/j.febslet.2005.08.040

Hofmeister, B. T., Lee, K., Rohr, N. A., Hall, D.,W., and Schmitz, R., J. (2017). Stable Inheritance of DNA Methylation Allows Creation of Epigenotype Maps and the Study of Epiallele Inheritance Patterns in the Absence of Genetic Variation. *Genome Biol.* 18, 155. doi:10.1186/s13059-017-128810.1186/s13059-017-1288-x

Hong, Y., Zhang, H., Huang, L., Li, D., and Song, F. (2016). Overexpression of a Stress-Responsive NAC Transcription Factor Gene ONAC022 Improves Drought and Salt Tolerance in Rice. *Front. Plant Sci.* 7, 4. doi:10.3389/fpls.2016.00004

Hu, T., Jin, Y., Li, H., Amombo, E., and Fu, J. (2016). Stress Memory Induced Transcriptional and Metabolic Changes of Perennial Ryegrass (Lolium Perenne) in Response to Salt Stress. *Physiol. Plant.* 156, 54–69. doi:10.1111/ppl.12342

Hume, S., Bo, D., Stacey, A. S., Suhua, F., Maria, B., Matteo, P., et al. (2013). Plants Regenerated from Tissue Culture Contain Stable Epigenome Changes in Rice. *ELife* 2, e00354. doi:10.7554/eLife.00354

Hung, Y. H., Liu, F., Zhang, X. Q., Xiao, W., and Hsieh, T., F. (2018). Sexual and Non-sexual Reproduction: Inheritance and Stability of Epigenetic Variations and Consequences for Breeding Application. *Adv. Bot. Res.* 88, 117–163. doi:10.1016/bs.abr.2018.09.002

Hyun, K., Jeon, J., and Park, K. (2017). Writing, Erasing and Reading Histone Lysine Methylations. *Exp. Mol. Med.* 49, e324. doi:10.1038/emm.2017.11

Ikeda, Y., Kinoshita, Y., Susaki, D., Ikeda, Y., Iwano, M., Takayama, S., et al. (2011). HMG Domain Containing SSRP1 Is Required for DNA Demethylation and Genomic Imprinting in *Arabidopsis*. *Dev. Cell.* 21, 589–596. doi:10.1016/j.devcel.2011.08.013

Ingelbrecht, I., Vanhoudt, H., Vanmontagu, M., and Depicker, A. (1994). Post-Transcriptional Silencing of Reporter Transgenes in Tobacco Correlates with DNA Methylation. *Proc. Natl. Acad. Sci. USA.* 91, 10502–10506. doi:10.1073/pnas.91.22.10502

Ingouff, M., and Berger, F. (2010). Histone3 Variants in Plants. *Chromosoma* 119, 27–33. doi:10.1007/s00412-009-0237-1

Jaffe, A. E., Feinberg, A. P., Irizarry, R. A., and Leek, J. T. (2017). Significance Analysis and Statistical Dissection of Variably Methylated Regions. *Biostatistics* 13, 166–178. doi:10.1093/biostatistics/kxr013

Jaganathan, D., Thudi, M., Kale, S., Azam, S., Roorkiwal, M., Gaur, P. M., et al. (2015). Genotyping-by-sequencing Based Intra-specific Genetic Map Refines a "QTL-Hotspot" Region for Drought Tolerance in Chickpea. *Mol. Genet. Genomics* 290, 559–571. doi:10.1007/s00438-014-0932-3

Janmohammadi, M., Zolla, L., and Rinalducci, S. (2015). Low Temperature Tolerance in Plants: Changes at the Protein Level. *Phyto Chem.* 117, 76–89. doi:10.1016/j.phytochem.2015.06.003

Jerzmanowski, A. (2007). SWI/SNF Chromatin Remodeling and Linker Histones in Plants. *Biochim. Biophys. Acta* 1769, 330–345. doi:10.1016/j.bbaexp.2006.12.003

Jha, U. C., Chaturvedi, S. K., Bohra, A., Basu, P. S., Khan, M. S., and Barh, D. (2014). Abiotic Stresses, Constraints and Improvement Strategies in Chickpea. *Plant Breed.* 133 (2), 163–178. doi:10.1111/pbr.12150

Jisha, K., C., and Puthur, J.,T. (2014). Halopriming of Seeds Imparts Tolerance to NaCl and PEG Induced Stress in Vigna Radiata (L.) Wilczek Varieties. *Physiol. Mol. Biol. Plants.* 303, 312. doi:10.1007/s12298-014-0234-6

Johannes, F., and Colomé-Tatché, M. (2011). Quantitative Epigenetics through Epigenomic Perturbation of Isogenic Lines. *Genetics* 188, 215–227. doi:10.1534/genetics.111.127118

John, M. C., and Amasino, R. M. (1989). Extensive Changes in DNA Met Hylation Patterns Accompany Activation of a Silent T-DNA Ipt Gene in *Agrobacterium Tumefaciens*-transformed Plant Cells. *Mol. Cell. Biol.* 9, 4298–4303. doi:10.1128/mcb.910.4298-4303.198910.1128/mcb.9.10.4298

Jones, D. F. (1917). Dominance of Linked Factors as a Means of Accounting for Heterosis. *Proc. Natl. Acad. Sci. U. S. A.* 3, 310–312. doi:10.1073/pnas.3.4.310

Jones, P. A., and Baylin, S. B. (2007). The Epigenomics of Cancer. *Cell.* 128, 683–692. doi:10.1016/j.cell.2007.01.029

Joseph, C., and David, L. (2006). Epigenetic Gene Regulation in the Bacterial World. *Am. Soc. Microbiol. Rev.* 70, 830–856. doi:10.1128/MMBR.00016-06

Jurkowska, R. Z., and Jeltsch, A. (2016). Enzymology of Mammalian DNA Methyltransferases. *Adv. Exp. Med. Biol.* 945, 87–122. doi:10.1007/978-3-319-43624-1-510.1007/978-3-319-43624-1_5

Jurkowski, T. P., and Jeltsch, A. (2011). On the Evolutionary Origin of Eukaryotic DNA Methyltransferases and Dnmt 2. *PLoS One* 6, e28104. doi:10.1371/journal.pone.0028104

Kaeppler, S. M., Kaeppler, H. F., and Rhee, Y. (2000). Epigenetic Aspects of Soma Clonal Variation in Plants. *Plant Mol. Biol.* 43, 179–188. doi:10.1023/A:1006423110134

Kakoulidou, I., Avramidou, E. V., Baránek, M., Brunel-Muguet, S., Farrona, S., Johannes, F., et al. (2021). Epigenetics for Crop Improvement in Times of Global Change. *Biol. (Basel)* 210 (8), 766. doi:10.3390/biology10080766

Kakutani, T. (2002). Epi-alleles in Plants: Inheritance of Epigenetic Information over Generations. *Plant Cell. Physiol.* 43, 1106. doi:10.1093/pcp/pcf131

Kapazoglou, A., Ganopoulos, I., Tani, E., and Tsaftaris, A. (2018). "Chapter Nine—Epigenetics, Epigenomics and Crop Improvement," in *Advances in Botanical Research*. Editor M. Kuntz (New York, NY, USA: Academic Press), 86, 287–324. doi:10.1016/bs.abr.2017.11.007

Kapoor, A., Agarwal, M., Andreucci, A., Zheng, X., Gong, Z., Hasegawa, P. M., et al. (2005). Mutations in a Conserved Replication Protein Suppress Transcriptional Gene Silencing in a DNA-methylation-independent Manner in *Arabidopsis*. *Curr. Biol.* 15, 1912–1918. doi:10.1016/j.cub.2005.09.013

Kaur, H., Shukla, R. K., Yadav, G., Chattopadhyay, D., and Majee, M. (2008). Two Divergent Genes Encoding L-Myo-Inositol 1-phosphate Synthase1 (CaMIPS1) and 2 (CaMIPS2) Are Differentially Expressed in Chickpea. *Plant Cell. Environ.* 31 (11), 1701–1716. doi:10.1111/j.1365-3040.2008.0187710.1111/j.1365-3040.2008.01877.x

Kaur, N. (2000). Effect of Osmo- and Hydropriming of Chickpea Seeds on Seedling Growth and Carbohydrate Metabolism under Water Deficit Stress. *Plant Growth Regul.* 37, 17–22. doi:10.1023/A:1020310008830

Kaur, S., Gupta, A., K., and Kaur, N. (2002). Effect of Osmo- and Hydropriming of Chickpea Seeds on Seedling Growth and Carbohydrate Metabolism under Water Deficit Stress. *Plant Growth Regul.* 37, 17–22. doi:10.1023/a:1020310008830

Kelkar, A., Thakur, V., Ramaswamy, R., and Deobagkar, D. (2009). Characterisation of Inactivation Domains and Evolutionary Strata in Human X Chromosome through Markov Segmentation. *PloS One* 4, e7885. doi:10.1371/journal.pone.0007885

Khan, N., Bano, A., Rahman, M. A., Guo, J., Kang, Z., and Babar, M. A. (2019). Comparative Physiological and Metabolic Analysis Reveals a Complex Mechanism Involved in Drought Tolerance in Chickpea (*Cicer Arietinum* L.) Induced by PGPR and PGRs. *Sci. Rep.* 9, 2097. doi:10.1038/s41598-019-38702-8

Khandal, H., Parween, S., Roy, R., Meena, M. K., and Chattopadhyay, D. (2017). MicroRNA Profiling Provides Insights into Post-transcriptional Regulation of Gene Expression in Chickpea Root Apex under Salinity and Water Deficiency. *Sci. Rep.* 7, 4632. doi:10.1038/s41598-017-04906-z

Kilian, J., Whitehead, D., Horak, J., Wanke, D., Weinl, S., Batistic, O., et al. (2007). The at Gen Express Global Stress Expression Data Set: Protocols, Evaluation and Model Data Analysis of UV-B Light, Drought and Cold Stress Responses. *Plant J.* 50, 347–363. doi:10.1111/j.1365-313X.2007.03052.x

Kinoshita, T., and Seki, M. (2014). Epigenetic Memory for Stress Response and Adaptation in Plants. *Plant Cell. Physiol.* 55, 1859–1863. doi:10.1093/pcp/pcu125

Klumpp, A., Ansel, W., Fomin, A., Schnirring, S., and Pickl, C. (2004). Influence of Climatic Conditions on the Mutations in Pollen Mother Cells of Tradescantia Clone 4430 and Implications for the Trad-MCN Bioassay Protocol. *Hereditas* 141, 142–148. doi:10.1111/j.1601-5223.2004.01806.x

Kobayashi, I. A., Nobusato, N., Kobayashi, T., and Uchiyama, I. (1999). Shaping the Genome Restriction-Modification Systems as Mobile Genetic Elements. *Curr. Opin. Genet. Dev.* 9, 649–656. doi:10.1016/s0959-437x(99)00026-x

Kobayashi, I. (2001). Behavior of Restriction-Modification Systems as Selfish Mobile Elements and Their Impact on Genome Evolution. *Nucleic Acids Res.* 29, 3742–3756. doi:10.1093/nar/29.18.3742

Kotula, L., Khan, H. A., Quealy, J., Turner, N. C., Vadez, V., Siddique, K. H., et al. (2015). Salt Sensitivity in Chickpea (*Cicer Arietinum* L.): Ions in Reproductive Tissues and Yield Components in Contrasting Genotypes. *Plant Cell. Environ.* 38 (8), 1565–1577. doi:10.1111/pce.12506

Ku, C. S., Naidoo, N., Wu, M., and Soong, R. (2011). Studying the Epigenome Using Next Generation Sequencing. *J.Med. Genet.* 48 (11), 721–730. Epub 2011 Aug 8. PMID: 21825079. doi:10.1136/jmedgenet-2011-100242

Kudapa, H., Azam, S., Sharpe, A. G., Taran, B., Li, R., Deonovic, B., et al. (2014). Comprehensive Transcriptome Assembly of Chickpea (*Cicer Arietinum* L.) Using Sanger and Next Generation Sequencing Platforms: Development and Applications. *PLoS One* 9 (1), e86039. doi:10.1371/journal.pone.0086039

Kumar, S., and Singh, A. (2016). Epigeneticregulation of Abiotic Stress Tolerance in Plants. *Adv. Plants Agric. Res.* 5, e00179. doi:10.15406/apar.2016.05.00179

Kumar, S. V., and Wigge, P. A. (2010). H2A.Z-containing Nucleosomes Mediate the Thermo Sensory Response in Arabidopsis. *Cell.* 140, 136–147. doi:10.1016/j.cell.2009.11.006

Labra, M., Ghiani, A., Citterio, S., Sgorbati, S., Sala, F., Vannini, C., et al. (2002). Analysis of Cytosine Methylation Pattern in Response to Water Deficit in Pea Root Tips. *Plant Biol.* 4 (06), 694–699. doi:10.1055/s-2002-37398

Lämke, J., and Bäurle, I. (2017). Epigenetic and Chromatin-Based Mechanisms in Environmental Stress Adaptation and Stress Memory in Plants. *Genome Biol.* 18, 124. doi:10.1186/s13059-017-1263-6

Lang, Z., Wang, Y., Tang, K., Tang, D., Datsenka, T., Cheng, J., et al. (2017). Critical Roles of DNA Demethylation in the Activation of Ripening-Induced Genes and Inhibition of Ripening-Repressed Genes in Tomato Fruit. *Proc. Natl. Acad. Sci. USA.* 114, E4511. doi:10.1073/pnas.1705233114

Lauria, M., and Rossi, V. (2011). Epigenetic Control of Gene Regulation in Plants. *Biochim. Biophys. Acta* 1809, 369–378. doi:10.1016/j.bbagrm.2011.03.002

Law, J. A., Vashisht, A. A., Wohlschlegel, J. A., and Jacobsen, S. E. (2011). SHH1, a Homeodomain Protein Required for DNA Methylation, as Well as RDR2, RDM4, and Chromatin Remodeling Factors, Associate with RNA Polymerase IV. *PLoS Genet.* 7, e1002195. doi:10.1371/journal.pgen.10021910.1371/journal.pgen.1002195

Li, H., Yuan, J., Wu, M., Han, Z., Li, L., Jiang, H., et al. (2018a). Transcriptome and DNA Methylomereveal Insights into Yield Heterosis in the Curds of Broccoli (*Brassica oleracea* L Var. Italic). *BMC Plant. Biol.* 18, 168. doi:10.1186/s12870-018-1384-4

Li, S., W., Zeng, X., Y., Leng, Y., Feng, L., and Kang, X., H. (2018b). Indole-3-butyric Acid Mediates Antioxidative Defense Systems to Promote Adventitious Rooting in Mung Bean Seedlings under Cadmium and Drought Stresses. *Ecotoxicol. Environ. Saf.* 161, 332–341. doi:10.1016/j.ecoenv.2018.06.003

Liu, B., and Wendel, J. (2003). Epigenetic Phenomena and the Evolution of Plant Allopolyploids. *Mol. Phylogenet. Evol.* 29, 365–379. doi:10.1016/s1055-7903(03)00213-6

Long, Y., Xia, W., Li, R., Wang, J., Shao, M., Feng, J., et al. (2011). Epigenetic QTL Mapping in Brassica Napus. *Genetics* 189, 1093–1102. doi:10.1534/genetics.111.131615

Lu, Y. B., Yang, L. T., Qi, Y. P., Li, Y., Li, Z., Chen, Y. B., et al. (2014). Identification of Boron-Deficiency-Responsive microRNAs in Citrus Sinensis Roots by Illumina Sequencing. *BMC Plant Biol.* 14, 123. doi:10.1186/1471-2229-14-123

Luo, J., and Hall, B. D. (2007). A Multi Step Process Gave Rise to RNA Polymerase IV of Land Plants. *J. Mol. Evol.* 64, 101–112. doi:10.1007/s00239-006-0093-z

Madlung, A., and Comai, L. (2004). The Effect of Stress on Genome Regulation and Structure. *Ann. Bot.* 94, 481–495. doi:10.1093/aob/mch172

Mallory, A. C., Bartel, D. P., and Bartel, B. (2005). MicroRNA-directed Regulation of Arabidopsis AUXIN RESPONSE FACTOR17 Is Essential for Proper Development and Modulates Expression of Early Auxin Response Genes. *Plant Cell.* 17 (5), 1360–1375. doi:10.1105/tpc.105.031716

Mantri, N. L., Ford, R., Coram, T. E., and Pang, E. C. (2007). Transcriptional Profiling of Chickpea Genes Differentially Regulated in Response to High-Salinity, Cold and Drought. *BMC Genomics* 8, 303. doi:10.1186/1471-2164-8-303

Marcos, F. C. C., Silveira, N. M., Marchiori, P. E. R., Machado, E. C., Souza, G. M., Landell, M. G. A., et al. (2018a). Drought Tolerance of Sugarcane Propagules Is Improved when Origin Material Faces Water Deficit. *PLoS One* 13 (12), e0206716. doi:10.1371/journal.pone.0206716

Marcos, F. C. C., Silveira, N. M., Mokochinski, J. B., Sawaya, A. C. H. F., Marchiori, P. E. R., Machado, E. C., et al. (2018b). Drought Tolerance of Sugarcane Is Improved by Previous Exposure to Water Deficit. *J. Plant Physiol.* 223, 9–18. doi:10.1016/j.jplph.2018.02.001

Mashaki, K. M., Garg, V., Ghomi, A. A. N., Kudapa, H., Chitikineni, A., Nezhad, K. Z., et al. (2018). RNA-seq Analysis Revealed Genes Associated with Drought Stress Response in Kabuli Chickpea *(Cicer Arietinum* L.). *PLoS ONE* 13, e0199774. doi:10.1371/journal.pone.0199774

Matzke, M. A., and Matzke, A. J. M. (2004). Planting the Seeds of a New Paradigm. *PloS Biol.* 2, 582–586. doi:10.1371/journal.pbio.0020133

Matzke, M. A., Mette, M. F., and Matzke, A. J. (2000). Transgene Silencing by the Host Genome Defense: Implications for the Evolution of Epigenetic Control Mechanisms in Plants and Vertebrates. *Plant Mol. Biol.* 43, 401–415. doi:10.1023/a:1006484806925

Matzke, M. A., Primig, M., Trnovsky, J., and Matzke, A. J. M. (1989). Reversible Methylation and Inactivation of Marker Genes in Sequentially Transformed Tobacco Plants. *EMBO J.* 8, 643–649. PMCID: PMC40085. doi:10.1002/j.1460-2075.1989.tb03421.x

Maunakea, A., K., Nagarajan, R., P., and Bilenky, M. (2010). Conserved Role of Intragenic DNA Methylation in Regulating Alternative Promoters. *Nature* 466, 253–257. doi:10.1038/nature09165

McClintock, B. (1950). The Origin and Behavior of Mutable Loci in Maize. *Proc. Natl. Acad. Sci.* 36, 344–355. doi:10.1073/pnas.36.6.344

Meyer, P., and Saedler, H. (1996). Homology-dependent Gene Silencing in Plants. *Annu. Rev. Plant Physiol.* 47, 23–48. doi:10.1146/annurev.arplant.47.1.23z

Mohammadi, P., Bahramnejada, B., Badakhshana, H., and Kanouni, H. (2015). DNA Methylation Changes in *fusarium* Wilt Resistant and Sensitive Chickpea Genotypes (*Cicer Arietinum* L.). *Physiol. Mol. Plant Pathol.* 91, 72–80. doi:10.1016/j.pmpp.2015.06.001

Molinier, J., Ries, G., Zipfel, C., and Hohn, B. (2006). Transgeneration Memory of Stress in Plants. *Nature* 442, 1046–1049. doi:10.1038/nature05022

Mouradi, M., Bouizgaren, A., Farissi, M., Latrach, L., Qaddoury, A., and Ghoulam, C. (2016). Seed Osmopriming Improves Plant Growth, Nodulation, Chlorophyll Fluorescence and Nutrient Uptake in Alfalfa (Medicago Sativa L.)—Rhizobia Symbiosis under Drought Stress. *Sci. Hortic.* 213, 232–242. doi:10.1016/j.scienta.2016.11.002

Nakayama, Y., and Kobayashi, I. (1998). Restriction-modification Gene Complexes as Selfish Gene Entities: Roles of a Regulatory System in Their Establishment, Maintenance, and Apoptotic Mutual Exclusion. *Proc. Natl. Acad. Sci. U. S. A.* 95, 6442–6447. doi:10.1073/pnas.95.11.6442

Napoli, C., Lemieux, C., and Jorgensen, R. (1990). Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-suppression of Homologous Gene in Trans. *Plant Cell.* 2, 279–289. doi:10.1105/tpc.2.4.279

Naya, L., Paul, S., Valdes Lopez, O., Mendoza Soto, A. B., Soto, A. B., Nova Franco, B., et al. (2014). Regulation of Copper Homeostasis and Biotic Interactions by microRNA 398b in Common Bean. *PLoS ONE* 9, e84416. doi:10.1371/journal.pone.0084416

Neuhuber, B., Himes, B. T., Shumsky, J. S., Gallo, G., and Fischer, I. (2005). Axon Growth and Recovery of Function Supported by Human Bone Marrow Stromal Cells in the Injured Spinal Cord Exhibit Donor Variations. *Brain Res.* 1035, 73–85. doi:10.1016/j.brainres.2004.11.055

Ni, Z., Kim, E. D., Ha, M., Lackey, E., Liu, J., Zhang, Y., et al. (2009). Altered Circadian Rhythms Regulate Growth Vigour in Hybrids And Allopolyploidss. *Nature* 457, 327–331. doi:10.1038/nature07523

Ong-Abdullah, M., Ordway, J. M., Jiang, N., Ooi, S. E., Kok, S. Y., Sarpan, N., et al. (2015). Loss of Karma Transposon Methylation Underlies the Mantled Soma Clonal Variant of Oil Palm. *Nature* 525, 533–537. doi:10.1038/nature15365

Pandey, G., Sharma, N., Sahu, P. P., and Prasad, M. (2016). Chromatin-Based Epigenetic Regulation of Plant Abiotic Stress Response. *Curr. Genomics* 17 (6), 490–498. doi:10.2174/1389202917666160520103914

Papikian, A., Liu, W., Gallego-Bartolomé, J., and Jacobsen, S. E. (2019). Site-specific Manipulation of *Arabidopsis* Loci Using CRISPR-Cas9 Sun Tag Systems. *Nat. Commun.* 10, 729. doi:10.1038/s41467-019-08736-7

Park, C. L., Cohen, L. H., and Murch, R. L. (1996). Assessment and Prediction of Stress-Related Growth. *J. Pers.* 64, 71–105. doi:10.1111/j.1467-6494.1996.tb0081510.1111/j.1467-6494.1996.tb00815.x

Pedersen, D. S., and Grasser, K. D. (2010). The Role of Chromosomal HMGB Proteins in Plants. *Biochim. Biophys. Acta* 1799, 171–174. doi:10.1016/j.bbagrm.2009.11.004

Phillips, T. (2008). The Role of Methylation in Gene Expression. *Nat. Educ.* 1, 116.

Pikaard, C. S., and Mittelsten, S. O. (2014). Epigenetic Regulation in Plants. *Cold Spring Harb. Perspect. Biol.* 6, a019315. doi:10.1101/cshperspect.a019315

Pushpavalli, R., Krishnamurthy, L., Thudi, M., Gaur, P. M., Rao, M. V., Siddique, K. M. H., et al. (2015). Two Key Genomic Regions Harbour QTLs for Salinity Tolerance in ICCV 2 X JG 11 Derived Chickpea (*Cicer Arietinum* L.)

Recombinant Inbred Lines. *BMC Plant Biol.* 15, 124. doi:10.1186/s12870-015-0491-8

Rajkumar, M. S., Gupta, K., and Khemka, N. K. (2020). DNA Methylation Reprogramming during Seed Development and its Functional Relevance in Seed Size/weight Determination in Chickpea. *Commun. Biol.* 3, 340. doi:10.1038/s42003-020-1059-1

Rajpathak, S. N., and Deobagkar, D. D. (2017). Micro RNAs and DNA Methylation Are Regulatory Players in Human Cells with Altered X Chromosome to Autosome Balance. *Sci. Rep.* 7, 43235. doi:10.1038/srep43235

Raju, S. K. K., Shao, M. R., Sanchez, R., Xu, Y. Z., Sandhu, A., Graef, G., et al. (2018). An Epigenetic Breeding System in Soybean for Increased Yield and Stability. *Plant Biotechnol. J.* 16, 1836–1847. doi:10.1111/pbi.12919

Rakei, A., Maali, A. R., Zeinali, H., and Ranjbar, M. (2015). DNA Methylation and Physio-Biochemical Analysis of Chickpea in Response to Cold Stress. *Protoplasma* 253, 61–76. doi:10.1007/s00709-015-0788-3

Rasheed, A., Hao, Y., Xia, X., Khan, A., Xu, Y., Varshney, R. K., et al. (2017). Crop Breeding Chips and Genotyping Platforms: Progress, Challenges, and Perspectives. *Mol. Plant* 10, 1047–1064. doi:10.1016/j.molp.2017.06.008

Ream, T. S., Haag, J. R., Wierzbicki, A. T., Nicora, C. D., Norbeck, A. D., Zhu, J. K., et al. (2009). Sub Unit Compositions of the RNA-Silencing Enzymes Pol IV and Pol V Reveal Their Origins as Specialized Forms of RNA Polymerase II. *Mol. Cell.* 33, 192–203. doi:10.1016/j.molcel.2008.12.015

Rehman, H., Iqbal, H., Basra, S. M. A., Afzal, I., Farooq, M., Wakeel, A., et al. (2015). Seed Priming Improves Early Seedling Vigor, Growth and Productivity of Spring Maize. *J. Integr. Agric.* 14, 1745–1754. doi:10.1016/s2095-3119(14)61000-5

Richards, E. J. (2006). Inherited Epigenetic Variation Revisiting Soft Inheritance. *Nat. Rev. Genet.* 7, 395–401. doi:10.1038/nrg1834

Roorkiwal, M., vonWettberg, E. J., Upadhyaya, H. D., Warschefsky, E., Rathore, A., and Varshney, R. K. (2014). Exploring Germplasm Diversity to Understand the Domestication Process in Cicer Spp. Using Snp and Dart Markers. *PLoS One* 9 (7), e102016. doi:10.1371/journal.pone.0102016

Ruiz, F. V., and Voinnet, O. (2009). Roles of Plant Small RNAs in Biotic Stress Responses. *Annu. Rev. Plant. Biol.* 60, 485–510. doi:10.1146/annurev.arplant.043008.092111

Roux, F., Colome Tatche, M., Edelist, C., Warenaar, R., Guerche, P., Hospital, F., et al. (2011). Genome-Wide Epigenetic Perturbation Jumpstarts Patterns of Heritable Variation Found in Nature. *Genetics* 188, 1015–1017.

Sahu, P. P., Pandey, G., Sharma, N., Puranik, S., Muthamilarasan, M., and Prasad, M. (2013). Epigenetic Mechanism of Plant Stress Responses and Adaptation. *Plant Cell. Rep.* 32 (8), 1151–1159. doi:10.1007/s00299-013-1462-x

Sang, Y., Silvaortega, C. O., Wu, S., Yamaguchi, N., Wu, M. F., Pfluger, J., et al. (2012). Mutations in Two Non-canonical *Arabdopsis* SWI2/SNF2 Chromatin Remodelling ATPases Cause Embryogenesis and Stem Cell Maintenance Defects. *Plant J.* 72, 1000–1014. doi:10.1111/tpj.12009

Sani, E., Herzyk, P., Perrella, G., Colot, V., and Amtmann, A. (2013). Hyperosmotic Priming of Arabidopsis Seedlings Establishes a Long-Term Somatic Memory Accompanied by Specific Changes of the Epigenome. *Genome Biol.* 14, R59. doi:10.1186/gb-2013-14-6-r59

Sen, S., Chakraborty, J., Ghosh, P., Basu, D., and Das, S. (2017). Chickpea WRKY70 Regulates the Expression of a Homeodomain-Leucine Zipper (HD-Zip) I Transcription Factor CaHDZ12, Which Confers Abiotic Stress Tolerance in Transgenic Tobacco and Chickpea. *Plant Cell. Physiol.* 58, 1934–1952. doi:10.1093/pcp/pcx126

Shabalina, S. A., and Koonin, E. V. (2008). Origins and Evolution of Eukaryotic RNA Interference. *Trends Ecol. Evol.* 23, 578–587. doi:10.1016/j.tree.2008.06.005

Sharma, A., Basu, U., Malik, N., Daware, A., Thakro, V., Narnoliya, L., et al. (2019). Genome-wide Cis-Regulatory Signatures for Modulation of Agronomic Traits as Exemplified by Drought Yield Index (DYI) in Chickpea. *Funct. Integr. Genomics* 19, 973–992. doi:10.1007/s10142-019-00691-2

Shen, Y., Sun, S., Hua, S., Shen, E., Ye, C., Cai, D., et al. (2017). Analysis of Transcriptional and Epigenetic Changes in Hybrid Vigor of Allopolyploid Brassica Napus Uncovers Key Roles for Small RNAs. *Plant. J.* 91, 874–893. doi:10.1111/tpj.13605

Shui, X. R., Chen, Z. W., and Li, J. X. (2013). MicroRNA Prediction and its Function in Regulating Drought-Related Genes in Cowpea. *Plant Sci.* 210, 25–35. doi:10.1016/j.plantsci.2013.05.002

Shull, G. H. (1908). The Composition of a Field of Maize. *Am. Breed. Assoc. Rep.* 4, 296–301. doi:10.1093/jhered/os-4.1.296

Singh, A., Singh, S., Panigrahi, K., C., Reski, R., and Sarkar, A., K. (2014). Balanced Activity of microRNA166/165 and its Target Transcripts from the Class III Homeodomain-Leucine Zipper Family Regulates Root Growth in *Arabidopsis thaliana*. *Plant Cell. Rep.* 33 (6), 945–953. doi:10.1007/s00299-014-157310.1007/s00299-014-1573-z

Slaughter, A., Daniel, X., Flors, V., Luna, E., Hohn, B., and MauchMani, B. (2012). Descendants of Primed Arabidopsis Plants Exhibit Resistance to Biotic Stress. *Plant Physiol.* 158, 835. doi:10.1104/pp.111.191593

Soppe, W. J., Jacobsen, S. E., Alonso-Blanco, C., Jackson, J. P., Kakutani, T., Koornneef, M., et al. (2000). The Late Flowering Phenotype of Fwa Mutants Is Caused by Gain-Of-Function Epigenetic Alleles of a Homeodomain Gene. *Mol. Cell.* 6, 791–802. doi:10.1016/s1097-2765(05)00090-0

Sosa, V. G., Palomar, M., Covarrubias, A. A., and Reyes, J. L. (2017). The Legume miR1514a Modulates a NAC Transcription Factor Transcript to Trigger phasiRNA Formation in Response to Drought. *J. Exp. Bot.* 68 (8), 2013–2026. PMID: 28338719; PMCID: PMC5429018. doi:10.1093/jxb/erw380

Spillane, C., Steimer, A., and Grossniklaus, U. (2001). Apomixis in Agriculture: the Quest for Clonal Seeds. *Sex. Plant Reprod.* 14, 179–187. doi:10.1007/s00497-001-0117-1

Springer, N., and Schmitz, R. (2017). Exploiting Induced and Natural Epigenetic Variation for Crop Improvement. *Nat. Rev. Genet.* 18, 563–575. doi:10.1038/nrg.2017.45

Su, Z., Bernardo, A., Tian, B., Chen, H., Wang, S., Ma, H., et al. (2019). A Deletion Mutation in TaHRC Confers Fhb1 Resistance to Fusarium Head Blight in Wheat. *Nat. Genet.* 51, 1099–1105. doi:10.1038/s41588-019-0425-8

Tang, X. M., Tao, X., Wang, Y., Ma, D. W., Li, D., Yang, H., et al. (2014). Genomics. Analysis of DNA Methylation of Perennial Ryegrass under Drought Using the Methylation-Sensitive Amplification Polymorphism (MSAP) Technique. *Mol. Genet.* 289, 1075–1084. doi:10.1007/s00438-014-0869-6

Teixeira, F. K., Heredia, F., Sarazin, A., Roudier, F., Boccara, M., Ciaudo, C., et al. (2009). A Role for RNAi in the Selective Correction of DNA Methylation Defects. *Science* 323, 1600–1604. doi:10.1126/science.1165313

Thudi, M., Chitikineni, A., Liu, X., He, W., Roorkiwal, M., Yang, W., et al. (2016). Recent Breeding Programs Enhanced Genetic Diversity in Both Desi and Kabuli Varieties of Chickpea (*Cicer Arietinum* L.). *Sci. Rep.* 6, 38636. doi:10.1038/srep38636

Trindade, I., Capitao, C., Dalmay, T., Fevereiro, M., P., and Santos, D., M. (2010). miR398 and miR408 Are Up-Regulated in Response to Water Deficit in Medicago Truncatula. *Planta* 231 (3), 705–716. doi:10.1007/s00425-009-1078-0.38636

Tsaftaris, A. S., and Polidoros, A. N. (2000). DNA Methylation and Plant Breeding. *Plant Breed. Rev.* 18, 87–176. doi:10.1002/9780470650158.ch3

Tucker, S. L., Reece, J., Ream, T. S., and Pikaard, C. S. (2010). Evolutionary History of Plant Multi Subunit RNA Polymerases IV and V: Subunit Origins via Genome-wide and Segmental Gene Duplications, Retro Transposition, and Lineage-specific Sub Functionalization. *Cold Spring Harb. Symp. Quant. Biol.* 75, 285–297. doi:10.1101/sqb.2010.75.037

Tuteja, N. (2007). Mechanisms of High Salinity Tolerance in Plants. *Methods Enzymol.* 428, 419–438. doi:10.1016/S0076-6879(07)28024-3

Umezawa, T., Shimizu, K., Kato, M., and Ueda, T. (2000). Enhancement of Salt Tolerance in Soybean with NaCl Pretreatment. *Physilogia Plant* 110, 59–63. doi:10.1034/j.1399-3054.2000.110108.x

Van der Krol, A. R., Mur, L. A., Beld, M., Mol, J. N., and Stuitje, A. R. (1990). Flavonoid Genes in Petunia: Addition of a Limited Number of Gene Copies May Lead to a Suppression of Gene Expression. *Plant Cell.* 2, 291–299. doi:10.1105/tpc.2.4.291

Van Dijk, K., Ding, Y., Malkaram, S., Riethoven, J. J., Liu, R., Yang, J., et al. (2010). Dynamic Changes in Genome-wide Histone H3lysine 4 Methylation Patterns in Response to Dehydration Stress in *Arabidopsis thaliana*. *BMC Plant Biol.* 10, 238. doi:10.1186/1471-2229-10-238

Varshney, R.,K., Thudi, M., Nayak, S., N., Gaur, P., M., Kashiwagi, J., Krishnamurthy, L., et al. (2014). Genetic Dissection of Drought Tolerance in Chickpea (Cicer Arietinum L.). *Theor. Appl. Genet.* 127, 445–462. doi:10.1007/s00122-013-2230-6

Varshney, R. K., Hiremath, P., J., Lekha, P., Kashiwagi, J., Balaji, J., Deokar, A. A., et al. (2009). A Comprehensive Resource of Drought- and Salinity- Responsive ESTs for Gene Discovery and Marker Development in Chickpea (*Cicer Arietinum* L.). *BMC Genomics* 10, 523. doi:10.1186/1471-2164-10-523

Varshney, R. K., Mohan, S. M., Gaur, P. M., Gangarao, N. V., Pandey, M. K., Bohra, A., et al. (2013). Achievements and Prospects of Genomics-Assisted Breeding in Three Legume Crops of the Semi-arid Tropics. *Biotechnol. Adv.* 31 (8), 1120–1134. doi:10.1016/j.biotechadv.2013.01.001

Varshney, R. K., Pandey, M., Bohra, K. A., Singh, V. K., Thudi, M., et al. (2019b). Toward the Sequence-Based Breeding in Legumes in the Post-genome Sequencing Era. *Theor. Appl. Genet.* 132, 797–816. doi:10.1007/s00122-018-3252-x

Varshney, R. K., Roorkiwal, M., Sun, S., Bajaj, P., Chitikineni, A., Thudi, M., et al. (2021). A Chickpea Genetic Variation Map Based on the Sequencing of 3,366 Genomes. *Nature* 599, 622–627. Available at: www.nature.com/articles/s41586-021-04066-1. doi:10.1038/s41586-021-04066-1

Varshney, R. K., Thud, M., Roorkiwal, M., He, W., Upadhyaya, H. D., Yang, W., et al. (2019a). Resequencing of 429 Chickpea Accessions from 45 Countries Provides Insights into Genome Diversity, Domestication and Agronomic Traits. *Nat. Genet.* 51, 857–864. doi:10.1038/s41588-019-0401-3

Vaucheret, H. (2008). Plant Argonautes. *Trends Plant Sci.* 13, 350–358. doi:10.1016/j.tplants.2008.04.007

Vazquez, F., Legrand, S., and Windels, D. (2010). The Biosynthetic Pathways and Biological Scopes of Plant Small RNAs. *Trends Plant Sci.* 15, 337–345. doi:10.1016/j.tplants.2010.04.001

Velten, J., Cakir, C., Youn, E., Chen, J., and Cazzonelli, I. C. (2012). Transgene Silencing and Transgene-Derived siRNA Production in Tobacco Plants Homozygous for an Introduced AtMYB90 Construct. *PLoS One* 7, e30141. doi:10.1371/journal.pone.0030141

Verhoeven, K. J. F., Jansen, J. J., Van Dijk, P. J., and Biere, A. (2010). Stress-induced DNA Methylation Changes and Their Heritability in Asexual Dandelions. *New Phytol.* 185, 1108–1118. doi:10.1111/j.1469-8137.2009.03121.x

Vijay, G., Zinta, G., kumar, S., and Jaswal, V. (2020). Quantitative Epigenetics: A New Avenue for Crop Improvement. *Epigenomes* 4 (4), 25. doi:10.3390/epigenomes404002

Waddington, C. H. (1942). The Epigenotype. *Int. J. Epidemiol.* 41 (1), 10–13. doi:10.1093/ije/dyr184

Walter, J., Nagy, L., Hein, R., Rascher, U., Beierkuhnlein, C., Willner, E., et al. (2011). Do plants Remember Drought? Hints towards a Drought-Memory in Grasses. *Environ. Exp. Bot.* 71, 34–40. doi:10.1016/j.envexpbot.2010.10.020

Wang, X., Elling, A. A., Li, X., Li, N., Peng, Z., He, G., et al. (2009). Genome-Wide and Organ-specific Landscapes of Epigenetic Modifications and Their Relationships to mRNA and Small RNA Transcriptomes in Maize. *Plant Cell.* 21, 1053–1069. doi:10.1105/tpc.109.065714

Wang, Z., Wang, Z., Wang, J., Sui, Y., Zhang, J., Liao, D., et al. (2012). A Quantitative Genetic and Epigenetic Model of Complex Traits. *BMC Bioinforma.* 13, 274. doi:10.1186/1471-2105-13-274

Woo, H. R., Pontes, O., Pikaard, C. S., and Richards, E. J. (2007). VIM1, a Methylcytosine-Binding Protein Required for Centromeric Hetero Chromatinization. *Genes. Dev.* 21, 267–277. doi:10.1101/gad.1512007

Xiao, W., Huawei, X., Shengde, L., Lili, C., and Fuping, Y. (2021). Epigenetic Regulation in Antiviral Innate Immunity. *Eur. J. Immunol.* 51, 1641–1651. doi:10.1002/eji.202048975

Xu, F., Liu, Q., Chen, L., Kuang, J., Walk, T., Wang, J., et al. (2013). Genome-wide Identification of Soybean microRNAs and Their Targets Reveals Their Organ-Specificity and Responses to Phosphate Starvation. *BMC Genomics* 14, 66. doi:10.1186/1471-2164-14-66

Yan, W., Deng, X. W., Yang, C., and Tang, X. (2021). The Genome-wide EMS Mutagenesis Bias Correlates with Sequence Context and Chromatin Structure in Rice. *Front. Plant Sci.* 12, 370–462. doi:10.3389/fpls.2021.579675

Yuan, L., Yang, X., and Makaroff, C. A. (2011). Plant Cohesins, Common Themes and Unique Roles. *Curr. Prot. Pept. Sci.* 12 (2), 93–104. doi:10.2174/138920311795684904

Zemach, A., Kim, M. Y., Hsieh, P. H., Coleman, D. D., Eshed, W. L., Thao, K., et al. (2013). The Arabidopsis Nucleosome Remodeler DDM1 Allows DNA Methyltransferases to Access H1-Containing Heterochromatin. *Cell.* 153, 193–205. doi:10.1016/j.cell.2013.02.033

Zentner, G. E., and Henikoff, S. (2013). Regulation of Nucleosome Dynamics by Histone Modifications. *Nat. Struct. Mol. Biol.* 20, 259–266. doi:10.1038/nsmb.2470

Zhang, H., Lang, Z., and Zhu, J. K. (2018). Dynamics and Function of DNA Methylation in Plants. *Nat. Rev. Mol. Cell. Biol.* 19, 489–506. doi:10.1038/s41580-018-0016-z

Zhang, J., and Zhou, J. M. (2010). Plant Immunity Triggered by Microbial Molecular Signatures. *Mol. Plant* 3, 783–793. doi:10.1093/mp/ssq035

Zhang, M., An, P., Li, H., Wang, X., Zhou, J., Dong, P., et al. (2019). The miRNA-Mediated Post-Transcriptional Regulation of Maize in Response to High Temperature. *Int. J. Mol. Sci.* 20, 1754. doi:10.3390/ijms20071754

Zhang, Q., Wang, D., Lang, Z., He, L., Yang, L., Zeng, L., et al. (2016). Methylation Interactions in Arabidopsis Hybrids Require RNA-Directed DNA Methylation and Are Influenced by Genetic Variation. *Proc. Natl. Acad. Sci. USA.* 113, E4248–E4256. doi:10.1073/pnas.1607851113

Zhang, X., Yazaki, J., Sundaresan, A., Cokus, S., Chan, S. W., Chen, H., et al. (2006). Genome-wide High Resolution Mapping and Functional Analysis of DNA Methylation in Arabidopsis. *Cell.* 126, 1189–1201. doi:10.1016/j.cell.2006.08.003

Zhao, B., Liang, R., Ge, L., Li, W., Xiao, H., Lin, H., et al. (2007). Identification of Drought-Induced microRNAs in Rice. *Biochem. Biophys. Res. Commun.* 354 (2), 585–590. doi:10.1016/j.bbrc.2007.01.022

Zhu, N., Cheng, S., Liu, X., Du, H., Dai, M., Zhou, D. X., et al. (2015). The R2R3-type MYB Gene OsMYB91 Has a Function in Coordinating Plant Growth and Salt Stress Tolerance in Rice. *Plant Sci.* 236, 146–156. doi:10.1016/j.plantsci.2015.03.023

Zhu, Y., Rowley, M. J., Bohmdorfer, G., and Wierzbicki, A. T. (2012). A SWI/SNF Chromatin-Remodeling Complex Acts in Noncoding RNA-Mediated Transcriptional Silencing. *Mol. Cell.* 49, 298–309. doi:10.1016/j.molcel.2012.11.011

Zilberman, D., Coleman, D. D., Ballinger, T., and Henikoff, S. (2008). HistoneH2A.Z and DNA Methylation Are Mutually Antagonistic Chromatin Marks. *Nature* 456, 125–129. doi:10.1038/nature07324

Zilberman, D., Gehring, M., Tran, R. K., Ballinger, T., and Henikoff, S. (2007). Genome-wide Analysis of *Arabidopsis thaliana* DNA Methylation Uncovers an Interdependence between Methylation and Transcription. *Nat. Genet.* 39, 61–69. doi:10.1038/ng1929

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.

Copyright © 2022 Chandana, Mahto, Singh, Ford, Vaghefi, Gupta, Yadav, Manohar and Kumar. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Genetics

TYPE Review
PUBLISHED 25 August 2022
DOI 10.3389/fgene.2022.909007

Check for updates

OPEN ACCESS

EDITED BY
Abhishek Bohra,
Indian Institute of Pulses Research (ICAR), India

REVIEWED BY
Muhammad Waseem,
Hainan University, China
Md Mahmudul Hassan,
Patuakhali Science and Technology University, Bangladesh

*CORRESPONDENCE
Rajiv Kumar,
rajiv@ihbt.res.in

SPECIALTY SECTION
This article was submitted to Plant Genomics, a section of the journal Frontiers in Genetics

RECEIVED 31 March 2022
ACCEPTED 21 July 2022
PUBLISHED 25 August 2022

CITATION
Satyakam, Zinta G, Singh RK and Kumar R (2022), Cold adaptation strategies in plants—An emerging role of epigenetics and antifreeze proteins to engineer cold resilient plants. *Front. Genet.* 13:909007. doi: 10.3389/fgene.2022.909007

COPYRIGHT
© 2022 Satyakam, Zinta, Singh and Kumar. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Cold adaptation strategies in plants—An emerging role of epigenetics and antifreeze proteins to engineer cold resilient plants

Satyakam[1,2], Gaurav Zinta[1,2], Rajesh Kumar Singh[1,2] and Rajiv Kumar[1,2]*

[1]Biotechnology Division, CSIR-Institute of Himalayan Bioresource Technology, Palampur, Himachal Pradesh, India, [2]Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, India

Cold stress adversely affects plant growth, development, and yield. Also, the spatial and geographical distribution of plant species is influenced by low temperatures. Cold stress includes chilling and/or freezing temperatures, which trigger entirely different plant responses. Freezing tolerance is acquired via the cold acclimation process, which involves prior exposure to non-lethal low temperatures followed by profound alterations in cell membrane rigidity, transcriptome, compatible solutes, pigments and cold-responsive proteins such as antifreeze proteins. Moreover, epigenetic mechanisms such as DNA methylation, histone modifications, chromatin dynamics and small non-coding RNAs play a crucial role in cold stress adaptation. Here, we provide a recent update on cold-induced signaling and regulatory mechanisms. Emphasis is given to the role of epigenetic mechanisms and antifreeze proteins in imparting cold stress tolerance in plants. Lastly, we discuss genetic manipulation strategies to improve cold tolerance and develop cold-resistant plants.

KEYWORDS

cold acclimation, freezing stress, DNA methylation, genetic engineering, antifreeze proteins

Introduction

Plants are sessile organisms constantly challenged by environmental stresses such as temperature extremes, UV radiation, salinity, drought, flooding, mineral toxicity, and pathogen attack. Among different environmental stresses, cold severely alters membrane fluidity, water and ionic balance, generates reactive oxygen species (ROS) that impair DNA, RNA, and protein stability, hamper photosynthetic efficiency and slow down biochemical reactions. These cellular and physiological changes reduce growth, development, and productivity and limit the geographical distribution of plants (Steponkus and Lynch, 1989; Hüner et al., 2013; Barrero-Sicilia et al., 2017; Bailey-

Serres et al., 2019). Therefore, understanding the plant responses and adaptation processes is important for developing of cold resilient plants, which is critical for global food security. The last decades have witnessed tremendous efforts to understand cold adaptation mechanisms in plants (Theocharis et al., 2012; Jeon & Kim, 2013; Zhao et al., 2015; Shi et al., 2018; Guo et al., 2018; Baier et al., 2019; Ding et at., 2019).

Plants encounter two types of low-temperature regimes in their natural habitat. The temperature range between 0 and 15°C causes cold/chilling stress, while temperatures below 0°C cause freezing stress, and distinct adaptive mechanisms help plants to deal with these two cold stress types. Plants use avoidance and tolerance strategies to mitigate cold stress. Avoiding mechanism involves preventing the formation of ice crystals inside the cell and is primarily associated with structural aspects. However, cold tolerance involves acquiring tolerance to low non-freezing temperature through a process known as cold acclimation, which includes prior exposure to nonlethal low temperature (Guy 1990; Thomashow, 1999). Cold acclimation is mainly characterized by the regulation of gene expression and metabolic changes that lead to various morphological, biochemical, and physiological alterations in plants (Liu et al., 2018; Liu et al., 2019; Wang B. et al., 2020; Fang et al., 2021; Guo et al., 2021).

Cold acclimation involves plasma membrane rigidification that affects ion concentration and metabolite transport from apoplast, endomembranes, and organelles, which initiates downstream cold signaling. These signaling cascades ultimately regulate the expression of cold-responsive (COR) genes. COR genes are induced by C-repeat Binding Factors (CBFs), which are under the control of the Inducer of CBF Expression (ICE). CBFs are genes encoding transcriptional activators having important roles in plant cold adaptation. Further, the ICE-CBF-COR regulatory module is a central pathway affecting cold response in plants (Chinnusamy et al., 2003; Jin et al., 2018). Cold inducible genes regulate the synthesis of compatible solutes (soluble sugars and proline), pigments (xanthophylls and carotenoids), and cold-responsive proteins like antifreeze proteins (AFPs), late embryogenesis abundant (LEA) proteins, heat shock proteins (HSPs), cold shock proteins (CSPs), and dehydrins, which eventually impart cold tolerance (Griffith et al., 1992a; Rinehart et al., 2007; Latowski et al., 2011; Liu et al., 2016; Yu et al., 2017; Meena et al., 2019). AFPs provide cold tolerance by arresting the growth of intracellular ice formation by binding to miniature ice-crystals formed due to freezing stress. Cold stress driven regulation of gene expression often depends on chromatin properties and small RNAs. In recent years, it has been documented that dynamics of histone modifications, DNA methylation, and biogenesis of miRNAs termed as “epigenetic regulators” were largely involved in the regulation of transcriptional and post-transcriptional gene expression in response to abiotic stress, including cold (Park et al., 2018; Hereme et al., 2021). Most of the epigenetic modifications are stable in the genome and forwarded to the next generation as epigenetic stress memory that could act more effectively towards subsequent cold stress. Our understanding of key genes that impart cold/freezing tolerance is crucial for developing cold resilient plants. The key genes primarily include signaling components like protein kinases, ion transporter, biochemical/ metabolic enzymes, and transcription factors are potential targets of crop improvement. With the help of genetic engineering techniques, the generation of overexpression/silencing line of key regulatory genes involved in cold adaptation could be an important strategy for developing cold resilient plants (Wang Q. et al., 2020; Li et al., 2021; Rivero et al., 2022).

This review summarises various aspects of cold stress response in plants and discusses the underlying adaptive mechanisms. Specifically, we discuss the role of epigenetic mechanisms and antifreeze proteins in cold stress tolerance. At the end, we highlight how modern genetic engineering tools can be utilized to develop cold resilient crops, and the industrial application of antifreeze proteins is also discussed. This review should provide us to characterize the process responsible for cold tolerance in plants that will be helpful in developing stress-resilient crops.

Cold sensing and signaling pathways

The early events upon cold stress include changes in cell membrane structure and lipid composition that provide the basis of low temperature sensing (Pareek et al., 2017; Guo et al., 2018; Poovaiah and Du, 2018; Guan et al., 2021). These changes induce a downstream cold signaling cascade by changing the ion and metabolite transport and redox state of the cell (Steponkus, 1984; Carpaneto et al., 2007; Jeon & Kim, 2013; Zhang et al., 2021a). Many plasma-membrane localized receptors, such as receptors like protein kinases (RLK) and leucine-rich repeats receptor-like protein kinase (LRR-RLK) have been shown to induce cold signaling pathways (Chen et al., 2021; Ye et al., 201; Su. et al., 2022). In addition, ion leakage, which is a common symptom of cold stress in plants, also causes Ca^{2+} changes when temperature ebbs. Perception of cold stress on the plasma membrane activates Ca^{2+} permeable channel that leads to the release of Ca^{2+} inside the cell. However, the frequency, duration, and amplitude of calcium ions, combinedly known as calcium signature, depends on the strength of stress condition (Zheng et al., 2021). Ca^{2+} imaging based on aequorin and yellow Cameleon has provided evidence of transient cold-induced Ca^{2+} channel activation in Arabidopsis (Krebs et al., 2012; Sun et al., 2021). Further experiments with Arabidopsis and moss show that cyclic nucleotide-gated calcium channel (CNGC) and glutamate-like receptor homologs (GLRs) function in Ca^{2+} signaling (Finka et al., 2012; Wang J. et. al., 2021; Ghosh et al., 2022). These Ca^{2+} are sensed by many calcium binding proteins like calmodulin (CaM), CaM-like proteins

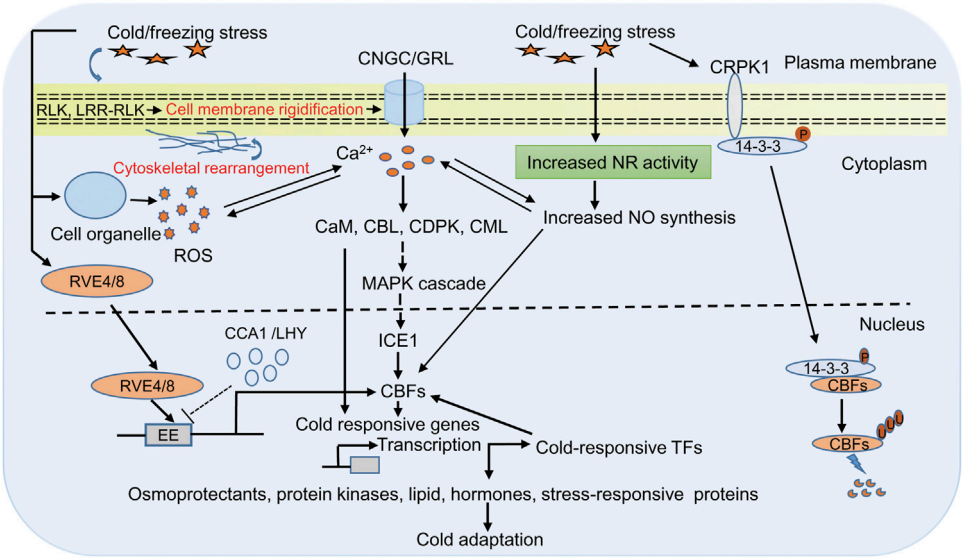


FIGURE 1
Representative diagram of cold responsive signaling pathway in plants. Plants sense cold/freezing signals through membrane receptor (RLK and LRR-RLK) and membrane rigidification. Cold sensing activates calcium channels (CNGC/GRL) that lead to increase Ca^{2+} in cytoplasm, which in turn activates of Ca^{2+} related protein likases (CaM, CML, CDPKs, and CBLs) and downstream signaling including MAPK signaling. These signaling cascades finally interacts with ICE1 and controls expression of CBFs/COR genes. COR genes encode proteins required for the biosynthesis of osmoprotectants, cryoprotectants, protein kinases, lipid, hormone, and stress-responsive proteins that are directly involved in cold tolerance. In addition, COR gene-dependent responses involve expression of diverse cold-induced transcription factors, which regulates CBFs expression in either positive or negative manner. The cold/freezing stress and increased Ca^{2+} activates the NADPH to generate more ROS. ROS and Ca^{2+} regulate each other's concentration, and this cross talk controls the expression of defense gene in the nucleus. Cold/freezing stress also triggers NO synthesis that is essential for cold acclimation response through CBF dependent manner. In another cold signaling pathway, 14-3-3 protein get phosphorylation by CRPK1 followed by translocation from the cytoplasm to the nucleus where it interacts with CBFs and trigger its degradation through the 26S proteasome pathway. In Arabidopsis, clock related MYB proteins RVE4/RVE8 plays an direct transcriptional activators of DREB1 expression in cold stress. In unstressed condition CCA1 and LHY suppressed DREB1 expression, wheras in stressed condition RVE4/RVE8 translocate from cytoplasm to the nucleus and induces the expression of CBFs/DREB1 through cis acting element EE by rapidly degrading CCA1and LHY. Abbreviations: RLK, receptors like protein kinases; LRR-RLK leucine-rich repeats receptor-like protein kinase; CNGC, cyclic nucleotide-gated calcium channel; GLRs, glutamate-like receptor homologs; Ca^{2+}, calcium ion; calcium binding proteins like CaM, calmodulin; CML, CaM-like proteins; CDPKs, Ca^{2+} dependent protein kinases; CBLs, calcineurin B-like proteins; MAPK, mitogen-activated protein kinase; CBFs, C-repeat Binding Factors; ICE, Inducer of CBF Expression; COR, cold-responsive; ROS, reactive oxygen species; NO, nitric oxide; CRPK1, cold-responsive protein kinase 1; RVE4/8, reveille4/8; lhy-cca1-Like1 (/LCL1); CCA1, circadian clock's oscillator component circadian clock-associated1; LHY, late elongated hypocotyl; EE, cis acting element; TFs, transcription factors; DREB1, dehydration responsive element binding-protein 1.

(CML), Ca^{2+} dependent protein kinases (CDPKs), and calcineurin B-like proteins (CBLs), which are essential for the regulation of CBF/COR gene expression (Huang et al., 2011; Zeng et al., 2015; Atif et al., 2019; Zhang et al., 2022). Also, calcium-binding proteins induce cold tolerance by modulating different mitogen-activated protein kinase (MAPK) through phosphorylation at their threonine and tyrosine residue that finally interacts with ICE1 and controls the expression of CBF genes (Yuan et al., 2018; Figure 1).

Besides Ca^{2+} and reactive oxygen species (ROS), nitric oxide (NO) is also a prominent component of a cold signaling pathway in plants (Figure 1). ROS is a highly reactive, short-lived secondary messenger and has a conserved signaling pathway in diverse stress conditions (Choudhury et al., 2013; You and Chan, 2015; Lim et al., 2019). Production of ROS (superoxide, hydroxyl radicals, and hydrogen peroxide) upon stress encounter is one of the early steps, and if its accumulation exceeds the threshold level, it could harm to lipid, protein, RNA, and DNA and generate oxidative stress in the cell (Mittler, 2002; Sharma et al., 2021). The ROS enters the cell through aquaporin membrane proteins, perceived by membrane receptors and modifies the cytoplasmic proteins to regulate signaling and cellular processes. For example, mucolipin 1 (MCOLN1; a member of the transient receptor potential channel gene family) is a ROS sensor in lysosomes that regulates autophagy, hydrogen peroxide sensor 1 (HPC1; LRR receptor kinase gene family) and GUARD CELL HYDROGEN PEROXIDE-RESISTANT1 (GHR1; plasma membrane LRR receptor kinase gene family) is a H_2O_2 sensor in Arabidopsis regulate Ca^{2+} driven stomatal movement (Hua et al., 2012; Zhang et al., 2016; Wu et al., 2020). Previous reports provide evidence that Ca^{2+} and ROS are interlinked and low ROS level induces Ca^{2+} influx into the cytoplasm (Mazars et al., 2010; Verhage, 2021). The increased Ca^{2+} activates the NADPH to generate more ROS, which is then

converted to H_2O_2 under the influence of superoxide dismutase. Thus, ROS and Ca^{2+} regulate each other's concentration, and this cross-talk controls the expression of defense gene in the nucleus (Mazars et al., 2010). The trade-off between ROS and Ca^{2+} in response to cold stress is a matter of debate and needs to dissect their role in cold tolerance. NO is another crucial signaling molecule, and its role in combating abiotic stress has also been studied well (Shi et al., 2012; Puyaubert and Baudouin, 2014; Fancy et al., 2017; Liu et al., 2021). NO has also been found to be working in association with hormones such as ABA, jasmonic acid, ethylene, and other molecules like Ca^{2+}, phosphatidic acid, H_2O_2, and melatonin to mitigate cold stress (Costa-Broseta et al., 2018; Feng et al., 2021). Therefore, it is prudent to assume that NO should play a crucial role in combating cold stress. The previous study has shown that level of endogenous NO increases with cold acclimation due to enhanced activity of the *NITRATE REDUCTASE 1* gene (Zhao et al., 2009). Further, trehalose also triggers NO upon cold stress (Liu et al., 2021). It also acts as an intermediate in lipid-based signaling and gene regulation during cold acclimation (Cantrel et al., 2011; Sohag et al., 2020). Triple mutants *nia1nia2noa1-2* is impaired in the nitrate reductase (NIA/NR) and Nitric Oxide-Associated (NOA1)-mediated NO production, and are thus NO deficient. Study in Arabidopsis has shed light on the importance of NO-induced cold acclimation. In this study, the author demonstrated that NO accumulation is essential for cold acclimation response through CBF-dependent and CBF-independent gene expression (Costa-Broseta et al., 2019).

C-repeat binding factors and cold-responsive signaling pathway

Once the temperature goes below the optimum, the *COR* gene springs into action to maintain homeostasis by mitigating the impact of cold stress. The first type of COR gene-dependent response involves encoding proteins required for the biosynthesis of osmoprotectants, cryoprotectants, protein kinases, lipid, hormone, and stress-responsive proteins like AFPs, HSPs, LEA, dehydrins that are directly involved in cold tolerance (Holmberge and Bülow, 1998; Thomashow, 2010; Kidokoro et al., 2022). Other sets of COR gene-dependent responses include genes such as early response to dehydration, low temperature-induced, response to abscisic acid, and cold-Induced transcription factors (Alves et al., 2011; Shi et al., 2015; Cao et al., 2021). The expression of COR genes is regulated by both CBF-dependent and CBF-independent pathways. CBFs are the transcription factor belonging to the superfamily ethylene-responsive element-binding factors and *APETALA2* (AP2/ERF), which recognize RCCGAC, a c-repeat dehydration-responsive element (CRT/DRE) (Akhtar et al., 2012; Zhu et al., 2021). CBFs have signature sequences, PKK/RPAGRxKFxETRHP and DSAWR, distinguishing them from other superfamily members (Canella et al., 2010; Fan et al., 2021). Out of four CBF genes found in the Arabidopsis genome, three *CBF* genes (*CBF1*, *CBF2*, *CBF3*) are induced by cold, whereas *CBF4* is induced by drought and salt stress. *CBF1*, *CBF2*, and *CBF3* are also known as DREB1b, DREB1c, and DREB1a, respectively. Of the three cold-induced *CBF* genes, only *CBF1* and *CBF3* positively regulate cold acclimation, whereas *CBF2* negatively regulates both *CBF 1* and *CBF3* by inducing different sets of genes (Novillo et al., 2007). A recent study suggests that CBF is also under the regulation of redox-dependent structural changes by Thioredoxin-H2(Trx-h2), a cytosolic redox protein, which enhances its function in the cold stress mitigation pathway. Trx-h2, which resides in the cytoplasm under normal conditions, upon cold stress migrates to the nucleus and binds to CBF and reduces the oxidized/inactive CBF (Lee et al., 2021). In CBF-dependent COR regulation, C-repeat binding factor/dehydration responsive element binding-protein-cold regulated (ICE-CBF3/DREB1-COR) is the major pathway regulating cold stress. CBF is under the direct control of several positive and negative regulatory elements and chromatin remodeling complexes (Ding et al., 2019). A recent study by Kidokoro et al. (2021) reports the circadian clock's oscillator component circadian clock-associated1 (CCA1) and late elongated hypocotyl (LHY) negatively regulate DREB1 expression under normal growth conditions. Upon cold stress, CCA1 and LHY were rapidly degraded, followed by translocation of MYB transcription factor reveille4/lhy-cca1-Like1 (RVE4/LCL1) and RVE8/LCL5 from the cytoplasm to the nucleus and inducing the expression of DREB1 through cis-acting element EE (Kidokoro et al., 2021).

Different hormones also known to regulate CBFs. For example, gibberellic acid (GA) metabolism and signaling are under the target of cold stress. Cold induction leads to the activation of Gibberellin 2-oxidase (GA2ox), which leads to the hydroxylation and inactivation of bioactive GA. Moreover, overexpression of CBF leads to enhanced DELLA protein accumulation through post-translational modification (Eremina et al., 2016; Devireddy et al., 2021). The previous study has shown that GA deficient mutants of Arabidopsis and rice have altered chilling and freezing tolerance (Eremina et al., 2016). GA also controls CBF expression by mediating the regulation of PHYTOCHROME-INTERACTING FACTOR 4 (PIF4). In long day (LD; 16 h photoperiod) plants, PIF4 represses the expression of CBF and is directly controlled by DELLA protein. Further, jasmonic acid is the central hub of the JAZ-BBX37-ICE1-CBF pathway, which positively regulates cold stress tolerance (An et al., 2021). CBFs also associated with ABA metabolism genes such as 9-CIS-EPOXYCAROTENOID DIOXYGENASE 2 (NCED2), NCED3, NCED5, and cytochrome P450 monooxygenase (CYP707A3 and CYP707A4) (Song et al., 2021). The recent results from various studies suggest that CBF regulation is more abstruse than it appears. Evidence suggests that

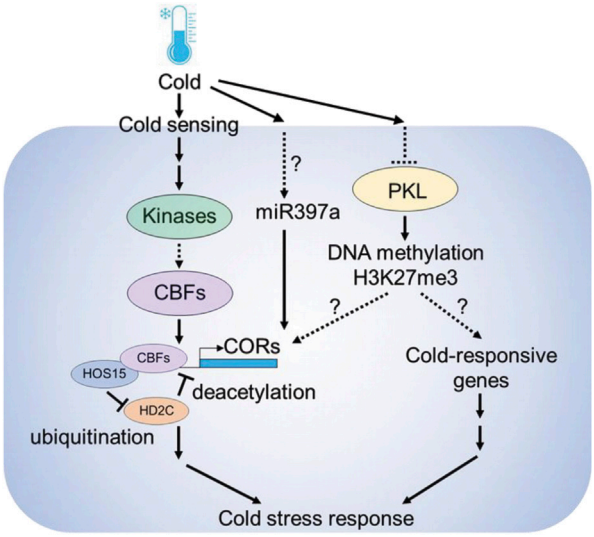


FIGURE 2
Epigenetic components involved in cold stress response in plants. Cold is sensed by upstream sensors followed by the activation of downstream gene expression. Under normal temperature conditions, HIGH EXPRESSION OF OSMOTICALLY RESPONSIVE GENES 15 (HOS15) interacts with HISTONE DEACETYLASE 2C (HD2C), and represses COLD RESPONSIVE (*COR*) gene expression by deacetylation. However, under cold stress conditions, HOS15 promotes HD2C degradation by ubiquitination, resulting in the increase of H3 acetylation on COR promoters. HOS15 also recruits CBFs to the COR promoters to activate COR gene expression. The chromatin remodeler PICKLE (PKL) modulates the chromatin status of *COR* genes through H3K27me3-dependent silencing. Also, under cold miR397a leads to the up-regulation of *COR* genes and enhanced cold tolerance.

repression of CBF3/DREB1a in ice1-1 is achieved not only by genetic regulation but also by DNA methylation-mediated gene silencing and overexpression of ICE1 has no impact on cold-induced CBF3/DREB1a (Kidokoro et al., 2021). Mieura et al. have demonstrated that overexpression of ICE1 leads to increased expression of CBF1 and CBF3 by 30% and CBF2 by 24% (Miura et al., 2007). Moreover, the same group has also shown that ICE1 mutation (S403A) leads to stabilization of ICE1 and a twofold increase in the expression of CBF3/DREB1a during cold stress. Earlier studies have pointed out that ICE is the master regulator of cold-induced genes (Chinnusamy et al., 2003; Miura et al., 2007; Miura et al., 2011), however, recent study contradicts this hypothesis.

Epigenetic regulation of cold stress

Several epigenetic components including microRNAs, DNA methylation and histone modifications are involved in cold stress responses in plants (Figure 2). MicroRNAs (miRNAs) are known to play a role in cold-stress responses in plants (Megha et al., 2018). Analyzing small RNA libraries identified cold regulated miRNAs in Arabidopsis (Sunkar and Zhu, 2004). Furthermore, the Arabidopsis overexpression lines of miR397a showed higher CBFs and COR genes expression and improved tolerance against cold stress (Dong and Pei, 2014). SICKLE (SIC) is a proline-rich protein known to participate in miRNAs biogenesis, and *sic-1* mutant shows high sensitivity to cold and salt stress suggesting that microRNAs play a central role in stress responses (Zhan et al., 2012). Other than microRNAs, long noncoding RNAs are also important. Recently a cold-induced long noncoding RNA called SVALKA was identified, which negatively regulates CBF1 expression by producing a cryptic antisense transcript (Kindgren et al., 2018)

DNA methylation is a conserved epigenetic mark playing a crucial role in plant development and stress responses via modulating chromatin packaging and gene expression (Banerjee et al., 2017). Exposure to chilling and freezing stress resulted in the alterations of cytosine methylation in the alpine plant *Chorispora bungeana* as revealed by methylation-sensitive amplified fragment-length polymorphism (Song et al., 2015).

The chromatin remodeler PICKLE (PKL) plays a crucial role in cold stress responses. Arabidopsis *pkl* mutants are sensitive to cold stress and the expression of *CBF3* and *COR* family genes such as *COR15B* and *RD29A* were downregulated in *pkl* mutants (Yang et al., 2019). Also, PKL cooperates with the members of SWR1 chromatin remodeling complex members such as PHOTOPERIOD INDEPENDENT EARLY FLOWERING 1 (PIE1) to deposit H3K27me3 at gene loci (Carter et al., 2018). In Arabidopsis, cold stress caused the decline of H3K27me3 at COR15A and GALACTINOL SYNTHASE 3 (GOLS3) genes, suggesting that PKL modulates cold stress responses *via* H3K27me3 histone modifications at *COR* genes (Kwon et al., 2009). Also, enrichment in histone acetylation marks occurs in the promoters of several *COR* genes, including *COR15A* and *COR47* under cold stress (Zhu et al., 2008; Pavangadkar et al., 2010; Park et al., 2018). The process of histone acetylation is regulated by the concurrent action of HISTONE ACETYLTRANSFERASES (HATs) and HISTONE DEACETYLASES (HDACs). The Arabidopsis lines overexpressing HISTONE DEACETYLASE 2D (HD2D) were tolerant to cold stress as revealed by lesser accumulation of malondialdehyde (MDA) levels in transgenic plants (Han et al., 2016).

HIGH EXPRESSION OF OSMOTICALLY RESPONSIVE GENE 1 (HOS1) is an E3 ubiquitin ligase that encodes a WD40-repeat protein. HOS1 is involved in the regulation of cold-responsive gene regulation by histone deacetylation (Zhu et al., 2008). HOS15 interacts with *HISTONE DEACETYLASE 2C* (HD2C) to regulate the expression of *COR* genes by binding to their promoters (Park et al., 2018). Under optimal temperature the HOS15-HD2C complex occupies the promoters of *COR* genes and induces the hypoacetylation of COR chromatin,

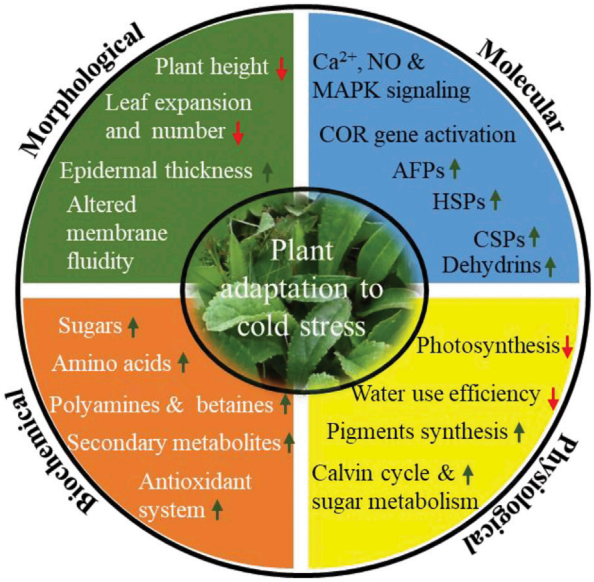


FIGURE 3
Different morphological, biochemical, physiological and molecular mitigation strategy acquired by plants upon cold stress. Change in different mitigation strategy are indicated by arrow (up) indicates increased concentration/expression, whereas arrow (down) indicates decreased concentration/expression). Abbreviations: Ca^{2+}, calcium ion; NO, nitric oxide; MAPK, mitogen-activated protein kinase; *COR* genes, cold responsive genes; AFPs, antifreeze proteins; HSPs, heat shock proteins; CSPs, cold shock proteins.

leading to the inhibition of COR gene expression. However, HOS15 functions as an E3 ubiquitin ligase during cold stress by recruiting CUL4 (CULLIN4) to degrade HD2C. This process results in the hyperacetylation of H3 on COR chromatin, which consequently enhances the ability of CBFs to bind to COR promoters (Yang et al., 2019). Recently, it has been shown that, POWERDRESS (PWR) interacts with HOS15 to modulate the cold stress response. Arabidopsis *pwr* mutants show low expression of COR genes and are sensitive to freezing stress, suggesting that PWR-HOS15-HD2C histone-modifying complex regulates the COR gene expression and freezing tolerance in plants (Lim et al., 2020). These findings suggest that epigenetic regulation is an important mechanism for plant responses to cold stress.

The fluctuating and recurring exposure to low-temperature stress can result in cold stress memory, which can significantly improve plant fitness under cold stress conditions (Markovskaya et al., 2008). For instance, sustained cold stress exposure of *Arabidopsis thaliana* seedlings resulted in freezing tolerance which was further increased in response to triggering stress after 3 days of priming (Leuendorf et al., 2020). Importantly, the *cbf* mutants did not show much difference suggesting that CBFs have a limited function in cold stress priming. In the other study, *Arabidopsis* seedlings that were cold primed at 4°C, then placed at 20°C as a lag phase, and finally subjected to 4°C (cold stress) exhibited significant freezing tolerance. The results show that this is due to raffinose accumulation after the lag phase, suggesting that raffinose metabolism may be involved in the retention of cold memory (Zuther et al., 2019). Arabidopsis *pkl-1* mutant was less primable to cold stress than wild type as mutant showed poorer survival after being primed by mild cold stress. These findings suggest that chromatin remodeler PKL plays a major role in cold-stress memory (Yang et al., 2019). Vernalization is a well studied process that requires cold, regulated epigenetically by *POLYCOMB REPRESSIVE COMPLEX 2* (*PRC2*) to mediate the suppression of a flower repressor *FLOWERING LOCUS C* (*FLC*) (Schubert et al., 2006). Overall, epigenetic components are involved in regulating cold induced stress and developmental responses in plants.

Cold adaptation in plants

Adaptation is a long term evolutionary process, whereas acclimation is a short term process contributing to overcome stress episodes. Both these processes contribute to combat the environmental odds and ensure plant survival. Cold acclimation occurs when temperature goes below normal at which plant is adapted to complete its life cycle.

Morphological adaptation

At the morphological level, plants reduce their height, leaf show reduced expansion and numbers, increase epidermal thickens, and induce rigidification of plasma membrane by changing the nature and composition of membrane lipids (Figure 3). Under low-temperature conditions, cells get dehydrated, which results in osmotic stress that impacts on membrane integrity and permeability. Parallelly cold/freezing condition causes metabolic disbalance in the cytomembrane, resulting in excessive accumulation of ROS that leads to oxidative stress. Consequently, the cell membranes get damaged. Therefore, the plasma membrane maintains its structural and functional integrity by increasing lipid unsaturation, altering lipid class/composition and lipid/protein ratio (Takahashi et al., 2013; De, 2014; Holthuis and Menon, 2014; Wang B. et al., 2020). Membrane fluidity is largely determined by the desaturation of fatty acids and fractions of phospholipids, galactolipids, sphingolipids, and sterols. It has been reported earlier that many plants, including monocotyledonous, dicotyledonous, herbaceous, and woody like oat, rye, mulberry, orchard grass, and Arabidopsis have increased phospholipid (phosphatidylcholine and phosphatidylethanolamine) and decreased sterols and cerebrosides composition of the plasma membrane in response to cold stress (Uemura and Steponkus, 1999;

Uemura et al., 2006). In contrast, Zhao et al. (2021) reported approximately 10% decrease in phosphatidylcholine in maize under low temperature suggesting regulation of membrane lipid composition predominantly takes place to counter low-temperature stress (Zhao et al., 2021). In addition, many lipid pathway enzymes get expressed to increase the proportion of unsaturated fatty acids. For example, genetic studies suggest that fatty acid desaturase (FAD) likely induces cold tolerance in *tetrahymena* thermophile to regulate membrane fluidity (Granel et al., 2019).

Biochemical adaptation

Further, to mitigate these effects at biochemical level, plants synthesize and accumulate an array of cryoprotectant *viz.* soluble sugars, specific amino acids (Proline, glycine, alanine and serine), polyamines, betaines and secondary metabolites (Ramazan et al., 2022; Saddhe et al., 2021; Ming et al., 2021; Figure 2). Sugars like sucrose, glucose, fructose, galactose, raffinose, and trehalose regulate osmotic potential, prevent ice crystal formation, scavenges ROS, and thus increases membrane stability in order to survive in cold stress. In addition, sugars also act as signaling molecule involved in plant growth and development as well as in stress conditions. A large number of articles have been reported that describe the possible link between sugars accumulation and cold tolerance in plants (Heidarvand and Amiri, 2010; Tarkowski and Van den Ende, 2015). Proline is considered a biomarker for cold tolerance and has been accumulated in cold adapted plants like *A. thaliana*, Chickpea, barley, and winter wheat (Kumar and Yadav, 2009; Primo-Capella et al., 2021). Proline maintains osmotic potential, stabilize membrane and proteins, scavenge ROS, and regulates osmotic stress-related gene expression. Further, polyamines (PAs), particularly putrescine, spermidine, and spermine accumulation and degree of abiotic stress tolerance have been discussed in many plants (Gill & Tuteja, 2010; Tiburcio et al., 2014; Alcázar et al., 2020). In cold/freezing stress, PAs increases osmolyte accumulation, control redox homeostasis, stabilize membranes, promote seed germination, improve fruit quality, protect photosynthetic apparatus and regulate gene expression (Oufir et al., 2008; Alcázar et al., 2020). Excessive accumulation of ROS species warrants a greater concentration of antioxidants to neutralize the harmful impact of ROS. It leads to the synthesis of secondary metabolites including phenols, flavonoids, coumarins, catechins, tocopherols, cinnamic acid derivatives, lignins, and polyfunctional organic acids (Ahmed et al., 2015; Robe et al., 2021; Shkryl et al., 2021).

Physiological adaptation

At the physiological level, photosynthesis is strongly affected by cold/freezing stress. Cold/freezing stress damage chloroplast membrane integrity, decreases the efficiency of photosystems I and II, intercellular CO_2, chlorophyll a, and chlorophyll b, which in turn reduces net photosynthesis (Lin et al., 2007; Gao et al., 2022; Figure 3). In order to maintain photosynthetic efficiency, plant synthesize accessory pigments like carotenoids, violaxanthin, antheraxanthin, and zeaxanthin, increases the activity of enzymes associated with the calvin cycle and sugar metabolism as well as increases the synthesis of D1 protein, which is essential to maintain the efficiency of PSII (Fang et al., 2019; Lu K. et. al., 2021; Wu et al., 2021). Cold stress also alters the morphology and anatomy of stomata. Results from stevia and various experiments with other plants has sown that stomatal size increases under cold stress (Hajihashemi et al., 2018). Altered stomatal morphology might cause a significant reduction in the intracellular CO_2 concentration and water use efficiency (https://doi.org/10.3389/fpls.2018.01430, 2018).

Molecular adaptation

At the molecular level, plants synthesized stress-responsive protein like AFPss, HSPs, CSPs, LEA and dehydrins to assist protein folding and refolding activities, preventing the denaturation of globular macromolecule and cellular protein transport and inhibition of ice crystal growth (Figure 3). AFPs are the most important proteins, which we discuss in the next section.

Antifreeze proteins and their role in cold tolerance

To survive in freezing stress, cold-hardy plants produce a specific type of protein called AFPs, which lowers the freezing temperature of the cytosol, inhibits the growth of ice-crystal in the apoplast, and attenuates freezing damage. AFPs were first discovered in arctic fish when De Vries and Wohlschlag discovered some plasma protein with the capacity to lower the freezing point of blood (DeVries and Wohlschlag, 1969). Few were proteinaceous and termed as AFPs, and others were glycosylated, named as antifreeze glycoproteins (AFGPs). These proteins are combinedly known as AF(G)Ps. Since its first discovery in marine teleosts, AF(G)Ps now have been reported from plants (Griffith et al., 1992a), molds fungi (Hoshino et al., 2003), sea-ice diatoms (Gwak et al., 2010), snow algae (Leya, 2013), and bacteria (Singh et al., 2014). The evolution of AF(G)Ps genes is thought to be done under adaptive conflict/environmental pressure by intra-gene/whole gene duplication, sequence divergence of selected genes like C-type lectin, trypsinogen, and sialic acid synthase, and from non-coding DNA sequences (Chen et al., 1997; Deng et al., 2010; Sorhannus, 2012; Zhuang et al., 2019). Although both AFPs and AFGPs have similar functions to mitigate freezing stress, they are

different structurally. AFPs have distinct primary, secondary and tertiary structures, whereas AFGPs have repeated tripeptide units Ala-Ala-Thr (Harding et al., 2003; Urbańczyk et al., 2017; Sun et al., 2021). AFPs act in a non-colligative manner, and their unique structure equips them to bind to the minute ice-crystals and prevent their growth (Chapsky and Rubinsky, 1997; Carvajal-Rondanelli et al., 2011). Thermal hysteresis (TH) and ice recrystallization inhibition (IRI) are the two distinct properties of AFPs. TH refers to the separation of melting and freezing point, whereas IRI is the ability to inhibit the growth of ice crystals. Few reports have discussed the properties of AFPs in detail (Kuiper et al., 2001; Zhang et al., 2004; Middleton et al., 2009; Gupta and Deswal, 2014; Maddah et al., 2021). Here, in this review, we will majorly focus on plant AFPs in brief.

Plant antifreeze proteins

In plants, AFPs were discovered for the first time in winter rye (Griffith et al., 1992a). It was demonstrated that the apoplastic extract from rye leaves acclimated for cold has a similar capacity to modify the ice formation as seen in fish and insects. These apoplastic extracts changed the ice-crystal morphology and had thermal hysteresis characteristics unique to the AFPs found in the animal kingdom. Since then, AFPs have been reported from more than 60 plants, including monocots, dicots, and gymnosperms (Smallwood et al., 1999; Wisniewski et al., 1999; Kuiper et al., 2001; Kawahara et al., 2009; Lauresen et al., 2011; Gupta and Deswal, 2012; Wang Q. et al., 2020; Arya et al., 2021; Yu et al., 2021). The AFPs from plants are different from the AFPs reported in animals in that it does not act as a cryoprotectant but as an ice interacting protein (Griffith et al., 2005). Plant AFPs, in general, were found to have low TH of around 0.1–0.5°C (Kuiper et al., 2001; Zhang et al., 2004) and high IRI compared to fishes, insects, and bacterial AFPs (2–13°C), however, plant AFPs with higher TH has also been reported (Gupta and Deswal, 2014). AFPs with low TH and high IRI seem to be of an evolutionary adaptation in plants as low TH allows the more controlled growth of ice crystals, and high IRI allows AFP to work even at minute concentrations.

Ice binding sites of antifreeze proteins

AFPs bind to ice using ice-binding site, but to date, there is no consensus on the actual mechanism of this interaction. Studies have shown that amino acid sequence composition (Davies and Sykes, 1997), secondary structure like beta-strand rich proteins (Lu et al., 2002) motifs such as N-acetyl group at C-2 peptide chain containing O-glycosidic linkage (Urabńczyk et al., 2017), gamma -methyl group at threonine (Chakrabarty and Jana, 2019) and ice like motif (Hudait et al., 2018) assist the binding. Though ice-binding site is mainly hydrophobic, a recent molecular study has shown that hydrophobicity and hydrophilicity contribute to the ice-binding mechanism (Hudait et al., 2018). Recent work highlights the importance of hydration shell (the sphere of water molecule around each dissolved ion) in AFP and ice interactions through the use of FTIR spectroscopy and self-modelling curve resolution (Zanetti-Polzi et al., 2019). Another study has shown that higher water density at non-ice binding surfaces also contributes to AFP hyperactivity (Biswas et al., 2021). Diverse anchored clathrate motifs have also been found in assisting the ice binding activity of AFP (Hudait et al., 2018). The clathrate motif model states that AFP forms cages around the methyl group on the ice-binding surface by organizing surrounding water molecules into an ice-like lattice and thus forming a quasi-liquid like layer between water and then merging with the already formed ice-crystals. Taking the cues from fishes and insects AFPs, Walker et al. presented the theoretical model of Lolium perenne AFP (Kuiper et al., 2001). The predicted model has extended flat beta-sheet to its opposite side, similar to that of the beta helical structure of insect AFP. It also indicated the presence of two ice-binding sites. This duplication of the ice-binding site might explain the superior IRI of plant AFP. Later on, a site-directed mutation study by the same group showed that only one acts as an actual ice-binding site out of the two putative ice-binding sites. Their experimental study confirmed the theoretical model, which predicted the ice-binding site to be planar, hydrophobic, and of high order (Middleton et al., 2009). Unlike animal AFPs, which have been divided into four distinct groups, type I-IV (Xiang et al., 2020), plant AFPs have not been classified due to immense diversity (Bredow and Walker, 2017). Studies in this regard have shown that plant AFPs can be classified on the basis of amino acid composition, secondary structure, presence of certain motifs, and its homology with other proteins. For example, a 118 residues long, heat-stable, hydrophilic, AFP isolated from ryegrass was found to have repeating motifs of seven conserved residues XXNXVG and no homology with other AFP (Sidebottom et al., 2000). Similarly, DcAFP, isolated from carrot, shares sequence similarity with POLYGALACTURONASE INHIBITOR PROTEIN family from apoplastic having LEUCINE RICH REPEAT (Dang-Quan et al., 2006). Furthermore, Plakestrin Homology, WRKY proteins, pathogenesis-related proteins such as 3-beta-glucosidase and thaumatin-like proteins have also been identified in the plant as AFPs (Table 1).

Extraction, purification, and identification of antifreeze proteins

Most of the research of AFPs has been centered on extraction, purification, and identification of AFPs, evaluation of antifreeze activity, and its implication for the development of cold resilient plants. To study AFPs, knowledge of AFP location is essential. In

TABLE 1 List of antifreez proteins reported in plants.

Plant origin	Protein name	Secondary structure composition	Domain similarity	Characteristics features	Localisation/References
Secale cereale	Glc AFP	Alpha helix-0% Extended strand-48.28% Coil- 51.72%	Endoglucanase	O-linked glycosylation	Apoplastic/ Hon et al., 1995
	Cht AFP	Beta rich	Class I Endochitinase	Hexagonal bipyramidal structure of ice crystal	Secretory pathway/Hon et al., 1995; Yeh et al., 2000
	Cht AFP	Beta rich	Class II Endochitinase		
	TLP AFP	Beta rich	Thaumatin like domain		Phloem tissue/ Hon et al., 1995
Daucus carota	dcAFP	Alpha + Beta	Leucine rich repeat	N linked glycosylation High level IRI	Secretory pathway/Ding et al., 2014
Lolium perenne	lpAFP	NA	NA	O linked glycosylation	Secretory pathway/Kuiper et al., 2001
Ricinus communis	rcAFP	Beta rich	Plant agglutinin	N and O linked glycosylation	NA/ Muthukumaran et al., 2011
	rcAFP	Beta rich	Plakesterin homology	NA	
Chlorella vulgaris	cvAFP	Alpha	NA	NA	Chloroplast
Hippophae rhamnoids	hr berry AFP	Alpha helix -41.03% Beta sheet - 14.89% Coil - 44%	TLR and LRR	Hexagonal ice shaping	Cytoplasmic/Gupta and Deswal, 2012
	hr berry AFP	Alpha helix -41.03% Beta sheet - 14.89% Coil - 44%	TLR and LRR	Hexagonal ice shaping	Cytoplasmic/Gupta and Deswal, 2012
	hr leaf AFP I	Alpha helix -28.82% Beta sheet - 20.83% Coil - 50.35%	LRR	NA	Extracellular/Gupta and Deswal, 2012
	Hr leaf II	Alpha helix -25.42% Beta sheet - 9.69% Coil - 64.89%	Cystein rich secretory protein 5	NA	Cell wall/Gupta and Deswal, 2012
Solanum dulcumara	STHP-64	Alpha helix- 18.78 Extended strand-19.80 Coil-61.42	WRKY	N and O limked glycosylation	Cytoplasmic/ Huang et al., 2002
Raphanus sativas	rsAFP	NA	NA	Hexagonal ice shaping	Apoplastic/Wisniewski et al., 2020
Triticum aestivum	taAFP	Alpha helix - 13.14% Extended strand - 25.14% Coil - 61.71%	LRR	N and O linked Glycosylation Heat stable AFP High IRI	Secretory pathway/Zhang et al., 2007
Picea abies	paAFP	NA	Chitinase	No glycosylation Bipyramidal Ice crystals	Apoplastic/Jarząbek et al., 2009
Deschampsia antarctica	daAFP	NA	LRR	O linked glycosylation	Secretory pathway/Cid et al., 2018
Populus suaveolens	psAFP	Alpha helix - 37.75% Extended strand - 19.87% Coil - 42.38%	Plakestrin homology	No glycosylation	NA/ Muthukumaran et al., 2011
Festuca pratensis	fpAFP	Alpha helix-NA Extended strand-38.26% Coil- 61.74	LRR	N and O-linked glycosylation	Chloroplast/ Cid et al., 2018; Muthukumaran et al., 2011

TABLE 2 List of *in silico* tools used for predicting and analysis of AFP.

Softwares	Prediction basis	References	Web access
AFPredictor	Surface-based pattern detection algorithm	Doxey and McConkey, (2006)	Freely available on request from the authors
AFP-Pred	Random forest approach	Kandaswamy et al. (2011)	No web server
AFP-PSSM	Support vector machine and position specific scoring matrix profiles	Zhao et al. (2012)	http: //59.73.198.144/AFP_PSSM/
AFP-PseAAC	Concept of pseudo amino acid composition	Mondal and Pai, (2014)	http://www.csbio.sjtu.edu.cn/bioinf/PseAAC/
AFP-Ensemble	Random forest classifiers and ensemble method	Yang et al. (2015)	http://afp.weka.cc/afp
TargetFreeze	Combination of weights using sequence evolutionary information and pseudo amino acid composition	He et al. (2015)	http://csbio.njust.edu.cn/bioinf/TargetFreeze
iAFP-Ense	Ensemble classifier	Xiao et al. (2016)	http://www.jci-bioinfo.cn/iAFP-Ense
CryoProtect	Amino acid composition, dipeptide composition, and physicochemical property	Pratiwi et al. (2017)	http://codes.bio/cryoprotect/
AFP-LSE	Latent space learning of K-spaced amino acid pairs	Usman et al. (2020)	https://github.com/Shujaat123/AFP-LSE
AFP-CMBPred	Extending consensus sequence into multi-blocks evolutionary information	Ali et al. (2021)	--------

woody, herbaceous, and gramineous plant species, most AFP is localized in bark, root tissue, and leaf blades, respectively. For extraction of AFPs from bark/root tissue, conventional methods like grinding and stirring, while for leaf blades, the infiltration-centrifugation method is used (Chincinska, 2021).The infiltration-centrifugation method has gained recent attention because it collects apoplastic fluid without protoplasmic contamination. Once the antifreeze proteins are extracted, they are purified using conventional techniques like ultrafiltration, column chromatography, ion exchange, and ammonium precipitation, and a classical method like ice affinity chromatogram (Tasaki and Okada, 2008; Sharma et al., 2019) followed by mass spectrometry for identification. A newer method like falling water ice affinity purification has also been reported (Adar et al., 2018). This method takes advantage of the affinity of ice-binding proteins for ice. In this purification method, the crude hydrolysate falls on a chilled vertical surface of a commercial ice machine; ice-binding proteins binds to the ice, whereas non-ice-binding proteins do not bind to ice. Many plant species have been used for isolation, purification, and identification of AFPs, such as bittersweet nightshade, winter rye, carrot, ryegrass, malting barley, oat, *Hipphoae rhamnides*, *Brassica juncea*, and *Ammopiptanthus nanus* (Griffith et al., 1992b; Gupta and Deswal, 2012; Ding et al., 2014; Ding et al., 2015; Arya et al., 2021; Yu et al., 2021). Further, nanolitre osmometry, differential scanning calorimeter, sucrose sandwich splat assay, and capillary assays are used to detect antifreeze activity in the AFPs extract /purified AFPs based on how AFP changes the growth of ice crystals. The former two measure TH, whereas the latter two are used to measure IRI. Recently, a colorimetric assay based on the change in color of the freeze-labile AuNP (Gold Nanoparticle) solution has also been widely used in studying AFPs (Park et al., 2013; Mitchell et al., 2015). Proteins with AFP properties degrade the AuNP, thus changing the color of the solution, whereas non-AFP protein doesn't change the color of the solution. Furthermore, different *in silico* tools like AFPredictor, AFP-Pred, TargetFreeze, iAFP-Ense, afpCOOL, AFP-CMBPred, and AFP-LSE are also being used for predicting and analysis of AFP based on different principles (Table 2).

Applications of antifreeze proteins

Based on AFPs properties to inhibit low temperature damage, the development of AFPs based products have found in various fields such as agriculture, industries, and medicine. In agriculture, it is used as a biofertilizer (Eskandri et al., 2020), germination promoter (Kyu et al., 2019), and to develop AFPs transgenic plants. In the past decade, several AFPs-transgenic plants have been developed to adapt to cold environments like maize (Zhang et al., 2021b), tomato (Balamurugan et al., 2018), tobacco (Huang et al., 2021), sweet potato (Lai et al., 2020). Such transgenic plants might provide humanity with food security in the near future in the wake of climate change. In the medicinal field, AFPs have found their uses in cryopreservation of organs, embryos, oocytes, and improving cryopreservation efficiency (Lee et al., 2015). The use of AFPs as a cryopreservent has an advantage compared to synthetic cryoprotectant agents like DMSO. It reduces the damage and mortality of the preserved organs, cells, and tissues. For example, AFPs from winter flounder have shown a better recovery rate of red blood cells post preservation (Stevens et al., 2022). A study by the group Tomas et al., has been demonstrated that the application of extracellular AFP enhances the protection of cell monolayers (Tomas et al., 2006). AFPs have also been used in minimally

invasive surgery to destroy bad tissues (Liu et al., 2021) and improve the vitrification process of mouse oocytes (Robles et al., 2019). The global AFP market is expected to reach 26 million dollars by 2026 (https://www.marketsandmarkets.com/Market-Reports/antifreeze-protein-market-264931272.html). The major driving force in this increase is thought to be the frozen food industry. Studies have shown that the treatment of zachunni, cucumber (Hu et al., 2022), green beans (Kashyap et al., 2020), frozen desserts (Ma et al., 2022), vegetables (Kim et al., 2019), star fruit (Provesi et al., 2019), and beef (Hu et al., 2022) with AFP solution retains the original texture and sensory perception of frozen food post-preservation. Since artificial cryoprotectives like polyvenyl alcohol, polyampholyte, graphene oxide, cause health issues, natural AFPs can use as food preservatives without damaging health. In the material industry, AFPs have been utilized to make antifrosting, anti-icing polymers, and ceramics, which increase the safety measures in appliances (Eskandari et al., 2020).

Engineering cold resilient plants and its application

The development of cold resilient plants is a demand for crops for the future. Historical methods such as conventional breeding to modern biotechnological techniques like genome editing such as clustered regularly interspaced short palindromic repeat (CRISPR) technology, epigenetic modification through altered methylation tagging, and other methods can be applied to develop chilling/freezing tolerant crops. Recent advancements in understanding the plant metabolic, transcriptomic, and signaling response to cold/freezing stress have aided and paved the path for the utilization of generated knowledge via different techniques. For example, QTL identified in several crops by genome-wide association studies with freezing tolerance can be used in breeding cold/freezing tolerant crops (Zhang et al., 2014; Huang et al., 2018; Wąsek et al., 2022). Although, there is limited success in developing cold resilient plants using traditional breeding approaches. The main reason behind this is that different responses in different plant species achieve cold tolerance, and mostly it is a combined effect of multiple factors at gene and metabolite levels. Thus, a multidisciplinary approach targeting multiple factors simultaneously can provide a better success rate in successfully implementing strategies to develop cold resilient crops using any method. Below are few strategies that can be applied to develop cold resilient crops.

Strengthening the wall (primary defense)

Cold stress is received differently, and plants respond differently to them at various stages. Cell membrane systems are the primary site of freezing injury in plants, and making them more tolerant or insensitive toward freezing/cold stress can serve as the first line of defense in resisting cold. Increased cold stress tolerance has been achieved in several plants by either modulating cell membrane composition or manipulating genes and proteins involved in perceiving cold stress. For example, the level of membrane fatty acids and their saturation states is critical in determining low- temperature sensitivity. Increased unsaturation of membrane fatty acids leads to a decrease in low-temperature sensitivity. Thus, increasing the saturation of membrane fatty acids is an effective way to generate cold resilient plants (Zhang Y. et al., 2021). Similarly, Modulation of several other genes and proteins belonging to plasma membrane have shown a significant increase in cold/freezing tolerance in transgenic plants in tobacco (MpRCI, Feng et al., 2009), rice (OsSMP1, Zheng et al., 2021), Arabidopsis (PsCor413pm2, Zhou et al., 2018), (Feng et al., 2009; Zhou et al., 2018; Zheng et al., 2021).

Primed/pre-activation of cold stress signaling

One of the best approaches to contest cold stress is to pre-activate the stress signaling response, which in turn activates the downstream pathway and prepares the plant to combat the stress. Cold stress is sensed and signalled through divergent but mostly conserved signaling cascades in different plant species with ICE-CBF-COR pathway. Pre/constitutive activation of this pathway genes has been and is one of the most effective and promising ways to develop cold resilient crops (Liu et al., 1998; Chinnusamy et al., 2003; Wang et al., 2008; Zhang H. et al., 2019). Overexpression of signaling genes helps in the preparedness of endogenous freeze tolerance machinery. Overexpression (both constitutive and inducible) of TFs primarily CBFs, has enhanced freezing tolerance in transgenic plants (Park and Chen 2006; Lv et al., 2020). Similar enhancement in freezing tolerance is also reported by overexpression of cold-induced DREB and ICE TFs (Wang et al., 2008; Zhang L. et al., 2019). In addition to TFs, overexpression of COR, and LEA genes, acting downstream of CBFs also enhances freezing tolerance in many plants (Artus et al., 1996; Hara et al., 2003; Ju et al., 2021).

Increasing the warriors

Once a plant experience cold stress, it activates its warriors to defend and protect itself from the adverse effects. A plant cold stress warrior consists of different proteins and metabolites which protects plant cell from freezing. Sugars (trehalose, fructans), compatible solutes or osmolytes, proline, glycine betaine and many other small molecules and metabolites act like a warrior in scavenging ROS or providing osmoprotection. The concentration of these osmolytes and metabolites increases

during cold stress; thus, increasing their amount in advance or under specific conditions is an efficient way to engineer plants for cold stress adaptation (Giri, 2011; Miranda et al., 2017). Sometimes a constitutive overexpression for the biosynthesis of these metabolites and solutes can interfere with the plant developmental process (Goddijn et al., 1997; Cortina and Culiáñez-Macià, 2005; Giri, 2011) while sometimes not (Jang et al., 2003: Su et al., 2006). Thus a better way is to express them under stress-inducible promoters (Su and Wu, 2004; Miranda et al., 2007; Iordachescu and Imai, 2008).

Antifreeze proteins –new avenue to engineer cold resilient plants

Antifreeze proteins found mainly in cold region plants, microbes and animals, restricts formation of ice crystals within cell, and helps in the survival of plants in freezing environments. Recently, the function of AFPs and their role in providing plant freezing tolerance has gained importance. Several antifreeze proteins have been identified from many temperate plant species. Transferring single genes encoding antifreeze proteins to freezing-sensitive plants lowered their freezing temperatures by ~1°C (Griffith and Yaish, 2004). Similarly, there are ice-binding proteins (IBP) present in microbes, animals and plants, which control the growth of ice crystals and mitigate freezing damage. Several of these proteins have been identified, and heterologous expression of these proteins in plants successfully lowered the freezing temperature (Hightower et al., 1991; Kenward et al., 1999; Khanna and Daggard, 2006). The use of plant IBP has proven to have more tolerance, and their overexpression led to decreased electrolyte leakage, indicating membrane protection and enhanced freeze survival at temperatures below −5°C (Fan et al., 2002; Zhang et al., 2010; Bredow et al., 2016). Recently, it has been demonstrated that the expression of more than one IBP isoform, as would be expressed endogenously in plants, further enhanced freeze survival in transgenic *A. thaliana* (Bredow et al., 2016).

In conclusion, cold/freezing tolerance is a multifaceted process regulated at different cellular levels, starting from the plasma membrane to intercellular processes involving genes and metabolites. Thus, the strategy to develop cold resilient plants depends on plant species, degree of resilience needed, technology applied, and selection of genes/metabolites. With advancements in the understanding of cold stress response, evolvement of new biotechnological tools for precise gene editing, and multiple gene-editing, the development of cold resilience plants is very well within reach.

Author contributions

All authors wrote, edited and designed the figures of the manuscript.

Funding

This work is financially supported by CSIR funded project (MLP-201) and Department of Science and Technology SERB-ECR (Grant, ECR/2017/000771/LS).

Acknowledgments

We thank Director CSIR-IHBT for his constant support and suggestions. Satyakam acknowledges UGC-JRF fellowship. The manuscript has CSIR-IHBT publication number: 5119.

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.

References

Adar, C., Sirotinskaya, V., Bar Dolev, M., Friehmann, T., and Braslavsky, I. (2018). Falling water ice affinity purification of ice-binding proteins. *Sci. Rep.* 8 (1), 11046–11049. doi:10.1038/s41598-018-29312-x

Ahmed, N. U., Park, J. I., Jung, H. J., Hur, Y., and Nou, I. S. (2015). Anthocyanin biosynthesis for cold and freezing stress tolerance and desirable color in Brassica rapa. *Funct. Integr. Genomics* 15 (4), 383–394. doi:10.1007/s10142-014-0427-7

Akhtar, M., Jaiswal, A., Taj, G., Jaiswal, J. P., Qureshi, M. I., and Singh, N. K. (2012). DREB1/CBF transcription factors: Their structure, function and role in abiotic stress tolerance in plants. *J. Genet.* 91 (3), 385–395. doi:10.1007/s12041-012-0201-3

Alcázar, R., Bueno, M., and Tiburcio, A. F. (2020). Polyamines: Small amines with large effects on plant abiotic stress tolerance. *Cells* 9 (11), 2373. doi:10.3390/cells9112373

Ali, F., Akbar, S., Ghulam, A., Maher, Z. A., Unar, A., and Talpur, D. B. (2021). AFP-CMBPred: Computational identification of antifreeze proteins by extending consensus sequences into multi-blocks evolutionary information. *Comput. Biol. Med.* 139, 105006. doi:10.1016/j.compbiomed.2021.105006

Alves, M. S., Fontes, E. P., and Fietto, L. G. (2011). Early responsive to dehydration 15, a new transcription factor that integrates stress signaling pathways. *Plant Signal. Behav.* 6 (12), 1993–1996. doi:10.4161/psb.6.12.18268

An, J. P., Wang, X. F., Zhang, X. W., You, C. X., and Hao, Y. J. (2021). Apple B-box protein BBX37 regulates jasmonic acid mediated cold tolerance through the JAZ-BBX37-ICE1-CBF pathway and undergoes MIEL1-mediated ubiquitination and degradation. *New Phytol.* 229 (5), 2707–2729. doi:10.1111/nph.17050

Artus, N. N., Uemura, M., Steponkus, P. L., Gilmour, S. J., Lin, C., and Thomashow, M. F. (1996). Constitutive expression of the cold-regulated *Arabidopsis thaliana* COR15a gene affects both chloroplast and protoplast freezing tolerance. *Proc. Natl. Acad. Sci. U. S. A.* 93 (23), 13404–13409. doi:10.1073/pnas.93.23.13404

Arya, M., Prakash, S., Sougrakpam, Y., and Deswal, R. (2021). *Brassica juncea* leaf cuticle proteome analysis shows myrosinase protein, antifreeze activity, and post-translationally modified secretory proteins. *Plant Physiol. biochem.* 161, 234–247. doi:10.1016/j.plaphy.2021.02.012

Atif, R. M., Shahid, L., Waqas, M., Ali, B., Rashid, M., Azeem, F., et al. (2019). Insights on calcium-dependent protein kinases (CPKs) signaling for abiotic stress tolerance in plants. *Int. J. Mol. Sci.* 20 (21), 5298. doi:10.3390/ijms20215298

Baier, M., Bittner, A., Prescher, A., and van Buer, J. (2019). Preparing plants for improved cold tolerance by priming. *Plant Cell Environ.* 42 (3), 782–800. doi:10.1111/pce.13394

Bailey-Serres, J., Parker, J. E., Ainsworth, E. A., Oldroyd, G. E., and Schroeder, J. I. (2019). Genetic strategies for improving crop yields. *Nature* 575 (7781), 109–118. doi:10.1038/s41586-019-1679-0

Balamurugan, S., Ann, J. S., Varghese, I. P., Murugan, S. B., Harish, M. C., Kumar, S. R., et al. (2018). Heterologous expression of *Lolium perenne* antifreeze protein confers chilling tolerance in tomato. *J. Integr. Agric.* 17 (5), 1128–1136. doi:10.1016/S2095-3119(17)61735-0

Banerjee, A., Wani, S. H., and Roychoudhury, A. (2017). Epigenetic control of plant cold responses. *Front. Plant Sci.* 8, 1643. doi:10.3389/fpls.2017.01643

Barrero-Sicilia, C., Silvestre, S., Haslam, R. P., and Michaelson, L. V. (2017). Lipid remodelling: Unravelling the response to cold stress in Arabidopsis and its extremophile relative Eutrema salsugineum. *Plant Sci.* 263, 194–200. doi:10.1016/j.plantsci.2017.07.017

Biswas, A. D., Barone, V., and Daidone, I. (2021). High water density at non-ice-binding surfaces contributes to the hyperactivity of antifreeze proteins. *J. Phys. Chem. Lett.* 12 (36), 8777–8783. doi:10.1021/acs.jpclett.1c01855

Bredow, M., Vanderbeld, B., and Walker, V. K. (2016). Knockdown of ice-binding proteins in *Brachypodium distachyon* demonstrates their role in freeze protection. *PloS one* 11 (12), e0167941. doi:10.1371/journal.pone.0167941

Bredow, M., and Walker, V. K. (2017). Ice-binding proteins in plants. *Front. Plant Sci.* 8, 2153. doi:10.3389/fpls.2017.02153

Canella, D., Gilmour, S. J., Kuhn, L. A., and Thomashow, M. F. (2010). DNA binding by the Arabidopsis CBF1 transcription factor requires the PKKP/RAGRxKFxETRHP signature sequence. *Biochim. Biophys. Acta* 1799 (5-6), 454–462. doi:10.1016/j.bbagrm.2009.11.017

Cantrel, C., Vazquez, T., Puyaubert, J., Rezé, N., Lesch, M., Kaiser, W. M., et al. (2011). Nitric oxide participates in cold-responsive phosphosphingolipid formation and gene expression in *Arabidopsis thaliana*. *New Phytol.* 189 (2), 415–427. doi:10.1111/j.1469-8137.2010.03500.x

Cao, K., Wei, Y., Chen, Y., Jiang, S., Chen, X., Wang, X., et al. (2021). PpCBF6 is a low-temperature-sensitive transcription factor that binds the PpVIN2 promoter in peach fruit and regulates sucrose metabolism and chilling injury. *Postharvest Biol. Technol.* 181, 111681. doi:10.1016/j.postharvbio.2021.111681

Carpaneto, A., Ivashikina, N., Levchenko, V., Krol, E., Jeworutzki, E., Zhu, J. K., et al. (2007). Cold transiently activates calcium-permeable channels in Arabidopsis mesophyll cells. *Plant Physiol.* 143 (1), 487–494. doi:10.1104/pp.106.090928

Carter, B., Bishop, B., Ho, K. K., Huang, R., Jia, W., Zhang, H., et al. (2018). The chromatin remodelers PKL and PIE1 act in an epigenetic pathway that determines H3K27me3 homeostasis in Arabidopsis. *Plant Cell* 30, 1337–1352. doi:10.1105/tpc.17.00867

Carvajal-Rondanelli, P. A., Marshall, S. H., and Guzman, F. (2011). Antifreeze glycoprotein agents: Structural requirements for activity. *J. Sci. Food Agric.* 91 (14), 2507–2510. doi:10.1002/jsfa.4473

Chakraborty, S. A. N. D. I. P. A. N., and Jana, B. (2019). Antifreeze proteins: An unusual tale of structural evolution, hydration and function. *Proc. Indian Natl. Sci. Acad.* 85, 169–187. doi:10.16943/ptinsa/2018/49553

Chapsky, L., and Rubinsky, B. (1997). Kinetics of antifreeze protein-induced ice growth inhibition. *FEBS Lett.* 412 (1), 241–244. doi:10.1016/S0014-5793(97)00787-4

Chen, L., DeVries, A. L., and Cheng, C. H. C. (1997). Convergent evolution of antifreeze glycoproteins in Antarctic notothenioid fish and Arctic cod. *Proc. Natl. Acad. Sci. U. S. A.* 94 (8), 3817–3822. doi:10.1073/pnas.94.8.3817

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 (1), 53–78. doi:10.1111/jipb.13061

Chincinska, I. A. (2021). Leaf infiltration in plant science: Old method, new possibilities. *Plant Methods* 17 (1), 83–21. doi:10.1186/s13007-021-00782-x

Chinnusamy, V., Ohta, M., Kanrar, S., Lee, B. H., Hong, X., Agarwal, M., et al. (2003). ICE1: A regulator of cold-induced transcriptome and freezing tolerance in Arabidopsis. *Genes Dev.* 17 (8), 1043–1054. doi:10.1101/gad.1077503

Choudhury, S., Panda, P., Sahoo, L., and Panda, S. K. (2013). Reactive oxygen species signaling in plants under abiotic stress. *Plant Signal. Behav.* 8 (4), e23681. doi:10.4161/psb.23681

Cid, F. P., Maruyama, F., Murase, K., Graether, S. P., Larama, G., Bravo, L. A., et al. (2018). Draft genome sequences of bacteria isolated from the *Deschampsia Antarctica* phyllosphere. *Extremophiles* 22 (3), 537–552. doi:10.1007/s00792-018-1015-x

Cortina, C., and Culiáñez-Macià, F. A. (2005). Tomato abiotic stress enhanced tolerance by trehalose biosynthesis. *Plant Sci.* 169 (1), 75–82. doi:10.1016/j.plantsci.2005.02.026

Costa-Broseta, Á., Perea-Resa, C., Castillo, M. C., Ruíz, M. F., Salinas, J., and León, J. (2018). Nitric oxide controls constitutive freezing tolerance in *Arabidopsis* by attenuating the levels of osmoprotectants, stress-related hormones and anthocyanins. *Sci. Rep.* 8 (1), 9268–9310. doi:10.1038/s41598-018-27668-8

Costa-Broseta, Á., Perea-Resa, C., Castillo, M. C., Ruíz, M. F., Salinas, J., and León, J. (2019). Nitric oxide deficiency decreases C-repeat binding factor-dependent and-independent induction of cold acclimation. *J. Exp. Bot.* 70 (12), 3283–3296. doi:10.1093/jxb/erz115

Davies, P. L., and Sykes, B. D. (1997). Antifreeze proteins. *Curr. Opin. Struct. Biol.* 7 (6), 828–834. doi:10.1016/S0959-440X(97)80154-6

Deng, C., Cheng, C. H. C., Ye, H., He, X., and Chen, L. (2010). Evolution of an antifreeze protein by neofunctionalization under escape from adaptive conflict. *Proc. Natl. Acad. Sci. U. S. A.* 107 (50), 21593–21598. doi:10.1073/pnas.1007883107

Devireddy, A. R., Tschaplinski, T. J., Tuskan, G. A., Muchero, W., and Chen, J. G. (2021). Role of reactive oxygen species and hormones in plant responses to temperature changes. *Int. J. Mol. Sci.* 22 (16), 8843. doi:10.3390/ijms22168843

DeVries, A. L., and Wohlschlag, D. E. (1969). Freezing resistance in some Antarctic fishes. *Science* 163 (3871), 1073–1075. doi:10.1126/science.163.3871.1073

Ding, X., Zhang, H., Chen, H., Wang, L., Qian, H., and Qi, X. (2015). Extraction, purification and identification of antifreeze proteins from cold acclimated malting barley (*Hordeum vulgare* L.). *Food Chem.* 175, 74–81. doi:10.1016/j.foodchem.2014.11.027

Ding, X., Zhang, H., Liu, W., Wang, L., Qian, H., and Qi, X. (2014). Extraction of carrot (*Daucus carota*) antifreeze proteins and evaluation of their effects on frozen white salted noodles. *Food bioproc. Tech.* 7 (3), 842–852. doi:10.1007/s11947-013-1101-0

Ding, Y., Shi, Y., and Yang, S. (2019). Advances and challenges in uncovering cold tolerance regulatory mechanisms in plants. *New Phytol.* 222 (4), 1690–1704. doi:10.1111/nph.15696

Dong, C.-H., and Pei, H. (2014). Over-expression of miR397 improves plant tolerance to cold stress in *Arabidopsis thaliana*. *J. Plant Biol.* 57, 209–217. doi:10.1007/s12374-013-0490-y

Doxey, A. C., and McConkey, B. J. (2006). AFPredictor: A computational screening protocol for antifreeze/ice-structuring proteins. *Protoc. Exch.*. doi:10.1038/nprot.2006.213

Eremina, M., Rozhon, W., and Poppenberger, B. (2016). Hormonal control of cold stress responses in plants. *Cell. Mol. Life Sci.* 73 (4), 797–810. doi:10.1007/s00018-015-2089-6

Eskandari, A., Leow, T. C., Rahman, M. B. A., and Oslan, S. N. (2020). Antifreeze proteins and their practical utilization in industry, medicine, and agriculture. *Biomolecules* 10 (12), 1649. doi:10.3390/biom10121649

Fan, M., Zhou, R., Liu, Q., and Sun, Y. (2021). CBF transcription factors involved in the cold response of *Camellia japonica* (Naidong). *PeerJ* 9, e12155. doi:10.7717/peerj.12155

Fan, Y., Liu, B., Wang, H., Wang, S., and Wang, J. (2002). Cloning of an antifreeze protein gene from carrot and its influence on cold tolerance in transgenic tobacco plants. *Plant Cell Rep.* 21 (4), 296–301. doi:10.1007/s00299-002-0495-3

Fancy, N. N., Bahlmann, A. K., and Loake, G. J. (2017). Nitric oxide function in plant abiotic stress. *Plant Cell Environ.* 40 (4), 462–472. doi:10.1111/pce.12707

Fang, J. C., Tsai, Y. C., Chou, W. L., Liu, H. Y., Chang, C. C., Wu, S. J., et al. (2021). A CCR4-associated factor 1, OsCAF1B, confers tolerance of low-temperature stress to rice seedlings. *Plant Mol. Biol.* 105 (1), 177–192. doi:10.1007/s11103-020-01079-8

Fang, P., Yan, M., Chi, C., Wang, M., Zhou, Y., Zhou, J., et al. (2019). Brassinosteroids act as a positive regulator of photoprotection in response to chilling stress. *Plant Physiol.* 180 (4), 2061–2076. doi:10.1104/pp.19.00088

Feng, D. R., Liu, B., Li, W. Y., He, Y. M., Qi, K. B., Wang, H. B., et al. (2009). Over-expression of a cold-induced plasma membrane protein gene (MpRCI) from plantain enhances low temperature-resistance in transgenic tobacco. *Environ. Exp. Bot.* 65 (2-3), 395–402. doi:10.1016/j.envexpbot.2008.12.009

Feng, Y., Fu, X., Han, L., Xu, C., Liu, C., Bi, H., et al. (2021). Nitric oxide functions as a downstream signal for melatonin-induced cold tolerance in cucumber seedlings. *Front. Plant Sci.* 1432, 686545. doi:10.3389/fpls.2021.686545

Finka, A., Cuendet, A. F. H., Maathuis, F. J., Saidi, Y., and Goloubinoff, P. (2012). Plasma membrane cyclic nucleotide gated calcium channels control land plant thermal sensing and acquired thermotolerance. *Plant Cell* 24 (8), 3333–3348. doi:10.1105/tpc.112.095844

Gao, X., Dong, J., Rasouli, F., Pouya, A. K., Tahir, A. T., and Kang, J. (2022). Transcriptome analysis provides new insights into plants responses under phosphate starvation in association with chilling stress. *BMC Plant Biol.* 22 (1), 26–14. doi:10.1186/s12870-021-03381-z

Ghosh, S., Bheri, M., Bisht, D., and Pandey, G. K. (2022). Calcium signaling and transport machinery: Potential for development of stress tolerance in plants. *Curr. Plant Biol.* 29. 100235. doi:10.1016/j.cpb.2022.100235

Gill, S. S., and Tuteja, N. (2010). Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. *Plant Physiol. biochem.* 48 (12), 909–930. doi:10.1016/j.plaphy.2010.08.016

Giri, J. (2011). Glycinebetaine and abiotic stress tolerance in plants. *Plant Signal. Behav.* 6 (11), 1746–1751. doi:10.4161/psb.6.11.17801

Goddijn, O. J., Verwoerd, T. C., Voogd, E., Krutwagen, R. W., De Graff, P. T. H. M., Poels, J., et al. (1997). Inhibition of trehalase activity enhances trehalose accumulation in transgenic plants. *Plant Physiol.* 113 (1), 181–190. doi:10.1104/pp.113.1.181

Granel, M. L. S., Cánepa, C., Cid, N. G., Navarro, J. C., Monroig, Ó., Verstraeten, S. V., et al. (2019). Gene identification and functional characterization of a Δ12 fatty acid desaturase in *Tetrahymena thermophila* and its influence in homeoviscous adaptation to low temperature. *Biochim. Biophys. Acta. Mol. Cell Biol. Lipids* 1864 (11), 1644–1655. doi:10.1016/j.bbalip.2019.08.003

Griffith, M., Ala, P., Yang, D. S., Hon, W. C., and Moffatt, B. A. (1992a). Antifreeze protein produced endogenously in winter rye leaves. *Plant Physiol.* 100 (2), 593–596. doi:10.1104/pp.100.2.593

Griffith, M., Lumb, C., Wiseman, S. B., Wisniewski, M., Johnson, R. W., and Marangoni, A. G. (2005). Antifreeze proteins modify the freezing process in planta. *Plant Physiol.* 138 (1), 330–340. doi:10.1104/pp.104.058628

Griffith, M., Paul, A., Daniel, S. C., Yang, W-C., Moffatt, B. A., and Hon, W. C. (1992b). Antifreeze protein produced endogenously in winter rye leaves. *Plant Physiol.* 100 (2), 593–596. doi:10.1104/pp.100.2.593

Griffith, M., and Yaish, M. W. (2004). Antifreeze proteins in overwintering plants: A tale of two activities. *Trends Plant Sci.* 9 (8), 399–405. doi:10.1016/j.tplants.2004.06.007

Guan, P., Shi, W., Riemann, M., and Nick, P. (2021). Dissecting the membrane-microtubule sensor in grapevine defence. *Hortic. Res.* 8, 260. doi:10.1038/s41438-021-00703-y

Guo, Q., Li, X., Niu, L., Jameson, P. E., and Zhou, W. (2021). Transcription-associated metabolomic adjustments in maize occur during combined drought and cold stress. *Plant Physiol.* 186 (1), 677–695. doi:10.1093/plphys/kiab050

Guo, X., Liu, D., and Chong, K. (2018). Cold signaling in plants: Insights into mechanisms and regulation. *J. Integr. Plant Biol.* 60 (9), 745–756. doi:10.1111/jipb.12706

Gupta, R., and Deswal, R. (2014). Antifreeze proteins enable plants to survive in freezing conditions. *J. Biosci.* 39 (5), 931–944. doi:10.1007/s12038-014-9468-2

Gupta, R., and Deswal, R. (2012). Low temperature stress modulated secretome analysis and purification of antifreeze protein from *Hippophae rhamnoides*, a Himalayan wonder plant. *J. Proteome Res.* 11 (5), 2684–2696. doi:10.1021/pr200944z

Guy, C. L. (1990). Cold acclimation and freezing stress tolerance: Role of protein metabolism. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 41 (1), 187–223. doi:10.1146/annurev.pp.41.060190.001155

Gwak, I. G., Kim, H. J., Kang, S. H., and Jin, E. (2010). Antifreeze protein in Antarctic marine diatom, *Chaetoceros neogracile*. *Mar. Biotechnol.* 12 (6), 630–639. doi:10.1007/s10126-009-9250-x

Hajihashemi, S., Noedoost, F., Geuns, J., Djalovic, I., and Siddique, K. H. (2018). Effect of cold stress on photosynthetic traits, carbohydrates, morphology, and anatomy in nine cultivars of Stevia rebaudiana. *Front. Plant Sci.* 2018. 1430. doi:10.3389/fpls.2018.01430

Han, Z., Yu, H., Zhao, Z., Hunter, D., Luo, X., Duan, J., et al. (2016). AtHD2D gene plays a role in plant growth, development, and response to abiotic stresses in *Arabidopsis thaliana*. *Front. Plant Sci.* 7, 310. doi:10.3389/fpls.2016.00310

Hara, M., Terashima, S., Fukaya, T., and Kuboi, T. (2003). Enhancement of cold tolerance and inhibition of lipid peroxidation by citrus dehydrin in transgenic tobacco. *Planta* 217 (2), 290–298. doi:10.1007/s00425-003-0986-7

Harding, M. M., Anderberg, P. I., and Haymet, A. D. J. (2003). 'Antifreeze'glycoproteins from polar fish. *Eur. J. Biochem.* 270 (7), 1381–1392. doi:10.1046/j.1432-1033.2003.03488.x

He, X., Han, K., Hu, J., Yan, H., Yang, J. Y., Shen, H. B., et al. (2015). TargetFreeze: Identifying antifreeze proteins via a combination of weights using sequence evolutionary information and pseudo amino acid composition. *J. Membr. Biol.* 248 (6), 1005–1014. doi:10.1007/s00232-015-9811-z

Heidarvand, L., and Maali Amiri, R. (2010). What happens in plant molecular responses to cold stress? *Acta Physiol. Plant.* 32 (3), 419–431. doi:10.1007/s11738-009-0451-8

Hereme, R., Galleguillos, C., Morales-Navarro, S., and Molina-Montenegro, M. A. (2021). What if the cold days return? Epigenetic mechanisms in plants to cold tolerance. *Planta* 254 (3), 46–11. doi:10.1007/s00425-021-03694-1

Hightower, R., Baden, C., Penzes, E., Lund, P., and Dunsmuir, P. (1991). Expression of antifreeze proteins in transgenic plants. *Plant Mol. Biol.* 17 (5), 1013–1021. doi:10.1007/BF00037141

Holmberg, N., and Bülow, L. (1998). Improving stress tolerance in plants by gene transfer. *Trends plant Sci.* 3 (2), 61–66. doi:10.1016/S1360-1385(97)01163-1

Holthuis, J., and Menon, A. K. (2014). Lipid landscapes and pipelines in membrane homeostasis. *Nature* 510 (7503), 48–57. doi:10.1038/nature13474

Hoshino, T., Kiriaki, M., Ohgiya, S., Fujiwara, M., Kondo, H., Nishimiya, Y., et al. (2003). Antifreeze proteins from snow mold fungi. *Can. J. Bot.* 81 (12), 1175–1181. doi:10.1139/b03-116

Hu, R., Zhang, M., Liu, W., Mujumdar, A. S., and Bai, B. (2022). Novel synergistic freezing methods and technologies for enhanced food product quality: A critical review. *Compr. Rev. Food Sci. Food Saf.* 21, 1979–2001. doi:10.1111/1541-4337.12919

Hua, D., Wang, C., He, J., Liao, H., Duan, Y., Zhu, Z., et al. (2012). A plasma membrane receptor kinase, GHR1, mediates abscisic acid- and hydrogen peroxide-regulated stomatal movement in Arabidopsis. *Plant Cell* 24 (6), 2546–2561. doi:10.1105/tpc.112.100107

Huang, C., Ding, S., Zhang, H., Du, H., and An, L. (2011). CIPK7 is involved in cold response by interacting with CBL1 in *Arabidopsis thaliana*. *Plant Sci.* 181 (1), 57–64. doi:10.1016/j.plantsci.2011.03.011

Huang, Q., Hu, R., Peng, C., and Chen, L. (2021). Expression of multi-domain type III antifreeze proteins from the Antarctic eelpout (*Lycodichths dearborni*) in transgenic tobacco plants improves cold resistance. *Aquac. Fish.* 6 (2), 186–191. doi:10.1016/j.aaf.2019.11.006

Huang, Z., Zhao, N., Qin, M., and Xu, A. (2018). Mapping of quantitative trait loci related to cold resistance in *Brassica napus* L. *J. Plant Physiol.* 231, 147–154. doi:10.1016/j.jplph.2018.09.012

Hudait, A., Odendahl, N., Qiu, Y., Paesani, F., and Molinero, V. (2018). Ice-nucleating and antifreeze proteins recognize ice through a diversity of anchored clathrate and ice-like motifs. *J. Am. Chem. Soc.* 140 (14), 4905–4912. doi:10.1021/jacs.8b01246

Hüner, N. P., Bode, R., Dahal, K., Busch, F. A., Possmayer, M., Szyszka, B., et al. (2013). Shedding some light on cold acclimation, cold adaptation, and phenotypic plasticity. *Botany* 91 (3), 127–136. doi:10.1139/cjb-2012-0174

Iordachescu, M., and Imai, R. (2008). Trehalose biosynthesis in response to abiotic stresses. *J. Integr. Plant Biol.* 50 (10), 1223–1229. doi:10.1111/j.1744-7909.2008.00736.x

Jang, I. C., Oh, S. J., Seo, J. S., Choi, W. B., Song, S. I., Kim, C. H., et al. (2003). Expression of a bifunctional fusion of the *Escherichia coli* genes for trehalose-6-phosphate synthase and trehalose-6-phosphate phosphatase in transgenic rice plants increases trehalose accumulation and abiotic stress tolerance without stunting growth. *Plant Physiol.* 131 (2), 516–524. doi:10.1104/pp.007237

Jarząbek, M., Pukacki, P. M., and Nuc, K. (2009). Cold-regulated proteins with potent antifreeze and cryoprotective activities in spruces (*Picea* spp.). *Cryobiology* 58 (3), 268–274. doi:10.1016/j.cryobiol.2009.01.007

Jeon, J., and Kim, J. (2013). Cold stress signaling networks in Arabidopsis. *J. Plant Biol.* 56 (2), 69–76. doi:10.1007/s12374-013-0903-y

Jin, Y. N., Zhai, S., Wang, W., Ding, X., Guo, Z., Bai, L., et al. (2018). Identification of genes from the ICE–CBF–COR pathway under cold stress in Aegilops–Triticum composite group and the evolution analysis with those from Triticeae. *Physiol. Mol. Biol. Plants* 24 (2), 211–229. doi:10.1007/s12298-017-0495-y

Ju, H., Li, D., Li, D., Yang, X., and Liu, Y. (2021). Overexpression of ZmDHN11 could enhance transgenic yeast and tobacco tolerance to osmotic stress. *Plant Cell Rep.* 40 (9), 1723–1733. doi:10.1007/s00299-021-02734-0

Kandaswamy, K. K., Chou, K. C., Martinetz, T., Möller, S., Suganthan, P. N., Sridharan, S., et al. (2011). AFP-pred: A random forest approach for predicting antifreeze proteins from sequence-derived properties. *J. Theor. Biol.* 270 (1), 56–62. doi:10.1016/j.jtbi.2010.10.037

Kashyap, P., Kumar, S., and Singh, D. (2020). Performance of antifreeze protein HrCHI4 from *Hippophae rhamnoides* in improving the structure and freshness of green beans upon cryopreservation. *Food Chem.* 320, 126599. doi:10.1016/j.foodchem.2020.126599

Kawahara, H., Fujii, A., Inoue, M., Kitao, S., Fukuoka, J., and Obata, H. (2009). Antifreeze activity of cold acclimated Japanese radish and purification of antifreeze peptide. *Cryo Lett.* 30 (2), 119–131.

Kenward, K. D., Brandle, J., McPherson, J., and Davies, P. L. (1999). Type II fish antifreeze protein accumulation in transgenic tobacco does not confer frost resistance. *Transgenic Res.* 8 (2), 105–117. doi:10.1023/A:1008886629825

Khanna, H. K., and Daggard, G. E. (2006). Targeted expression of redesigned and codon optimised synthetic gene leads to recrystallisation inhibition and reduced electrolyte leakage in spring wheat at sub-zero temperatures. *Plant Cell Rep.* 25 (12), 1336–1346. doi:10.1007/s00299-006-0191-9

Kidokoro, S., Hayashi, K., Haraguchi, H., Ishikawa, T., Soma, F., Konoura, I., et al. (2021). Posttranslational regulation of multiple clock-related transcription factors triggers cold-inducible gene expression in *Arabidopsis*. *Proc. Natl. Acad. Sci. U. S. A.* 118 (10), e2021048118. doi:10.1073/pnas.2021048118

Kidokoro, S., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2022). Transcriptional regulatory network of plant cold-stress responses. *Trends Plant Sci.*. doi:10.1016/j.tplants.2022.01.008

Kim, M., Jin, E. S., Sim, D. W., Won, H. S., Kim, E. K., Jang, S., et al. (2019). Cryoprotective effect of an antifreeze protein purified from *Tenebrio molitor* larvae on vegetables. *J. Clin. Neurol.* 94, 585–587. doi:10.3988/jcn.2019.15.4.585

Kindgren, P., Ard, R., Ivanov, M., and Marquardt, S. (2018). Transcriptional read-through of the long non-coding RNA SVALKA governs plant cold acclimation. *Nat. Commun.* 9, 4561–4611. doi:10.1038/s41467-018-07010-6

Krebs, M., Held, K., Binder, A., Hashimoto, K., Den Herder, G., Parniske, M., et al. (2012). FRET-based genetically encoded sensors allow high-resolution live cell imaging of Ca^{2+} dynamics. *Plant J.* 69 (1), 181–192. doi:10.1111/j.1365-313X.2011.04780.x

Kuiper, M. J., Davies, P. L., and Walker, V. K. (2001). A theoretical model of a plant antifreeze protein from Lolium perenne. *Biophys. J.* 81 (6), 3560–3565. doi:10.1016/S0006-3495(01)75986-3

Kumar, V., and Yadav, S. K. (2009). Proline and betaine provide protection to antioxidant and methylglyoxal detoxification systems during cold stress in *Camellia sinensis* (L.) O. Kuntze. *Acta Physiol. Plant.* 31 (2), 261–269. doi:10.1007/s11738-008-0227-6

Kwon, C. S., Lee, D., Choi, G., and Chung, W. I. (2009). Histone occupancy-dependent and-independent removal of H3K27 trimethylation at cold-responsive genes in *Arabidopsis*. *Plant J.* 60, 112–121. doi:10.1111/j.1365-313X.2009.03938.x

Kyu, S. Y., Naing, A. H., Pe, P. P. W., Park, K. I., and Kim, C. K. (2019). Tomato seeds pretreated with Antifreeze protein type I (AFP I) promotes the germination under cold stress by regulating the genes involved in germination process. *Plant Signal. Behav.* 14 (12), 1682796. doi:10.1080/15592324.2019.1682796

Lai, X., Zhang, Y., Gu, Y., and Yan, L. (2020). Transformation of insect derived antifreeze gene into sweet potato *(Ipomoea batatas)* and enhanced its freeze-tolerance. *Chin. Bull. Bot.* 55 (1), 9. doi:10.11983/CBB19133

Latowski, D., Kuczyńska, P., and Strzałka, K. (2011). Xanthophyll cycle–a mechanism protecting plants against oxidative stress. *Redox Rep.* 16 (2), 78–90. doi:10.1179/174329211X13020951739938

Lauersen, K. J., Brown, A., Middleton, A., Davies, P. L., and Walker, V. K. (2011). Expression and characterization of an antifreeze protein from the perennial rye grass, Lolium perenne. *Cryobiology* 62 (3), 194–201. doi:10.1016/j.cryobiol.2011.03.003

Lee, E. S., Park, J. H., Wi, S. D., Kang, C. H., Chi, Y. H., Chae, H. B., et al. (2021). Redox-dependent structural switch and CBF activation confer freezing tolerance in plants. *Nat. Plants* 7 (7), 914–922. doi:10.1038/s41477-021-00944-8

Lee, H. H., Lee, H. J., Kim, H. J., Lee, J. H., Ko, Y., Kim, S. M., et al. (2015). Effects of antifreeze proteins on the vitrification of mouse oocytes: Comparison of three different antifreeze proteins. *Hum. Reprod.* 30 (9), 2110–2119. doi:10.1093/humrep/dev170

Leuendorf, J. E., Frank, M., and Schmülling, T. (2020). Acclimation, priming and memory in the response of *Arabidopsis thaliana* seedlings to cold stress. *Sci. Rep.* 10, 689–711. doi:10.1038/s41598-019-56797-x

Leya, T. (2013). "Polyextremophiles," in *Snow algae: Adaptation strategies to survive on snow and ice* (Dordrecht: Springer), 401–423. doi:10.1007/978-94-007-6488-0_17

Lim, C. J., Park, J., Shen, M., Park, H. J., Cheong, M. S., Park, K. S., et al. (2020). The histone-modifying complex PWR/HOS15/HD2C epigenetically regulates cold tolerance. *Plant Physiol.* 184 (2), 1097–1111. doi:10.1104/pp.20.00439

Lim, S. D., Kim, S. H., Gilroy, S., Cushman, J. C., and Choi, W. G. (2019). Quantitative ROS bioreporters: A robust toolkit for studying biological roles of ROS in response to abiotic and biotic stresses. *Physiol. Plant.* 165 (2), 356–368. doi:10.1111/ppl.12866

Lin, K. H., Hwang, W. C., and Lo, H. F. (2007). Chilling stress and chilling tolerance of sweet potato as sensed by chlorophyll fluorescence. *Photosynthetica* 45 (4), 628–632. doi:10.1007/s11099-007-0108-z

Liu, C. T., Wang, W., Mao, B. G., and Chu, C. C. (2018). Cold stress tolerance in rice: Physiological changes, molecular mechanism, and future prospects. *Yi chuan = Hered.* 40 (3), 171–185. doi:10.16288/j.yczz.18-007

Liu, C., Yang, X., Yan, Z., Fan, Y., Feng, G., and Liu, D. (2019). Analysis of differential gene expression in cold-tolerant vs. cold-sensitive varieties of snap bean (*Phaseolus vulgaris* L.) in response to low temperature stress. *Genes Genomics* 41 (12), 1445–1455. doi:10.1007/s13258-019-00870-2

Liu, Q., Kasuga, M., Sakuma, Y., Abe, H., Miura, S., Yamaguchi-Shinozaki, K., et al. (1998). Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in Arabidopsis. *Plant Cell* 10 (8), 1391–1406. doi:10.1105/tpc.10.8.1391

Liu, T., Shi, J., Li, M., Ye, X., and Qi, H. (2021). Trehalose triggers hydrogen peroxide and nitric oxide to participate in melon seedlings oxidative stress tolerance under cold stress. *Environ. Exp. Bot.* 184, 104379. doi:10.1016/j.envexpbot.2021.104379

Lu, K., Sun, J., Li, Q., Li, X., and Jin, S. (2021a). Effect of cold stress on growth, physiological characteristics, and calvin-cycle-related gene expression of grafted watermelon seedlings of different gourd rootstocks. *Horticulturae* 7 (10), 391. doi:10.3390/horticulturae7100391

Lu, M., Wang, B., Li, Z., Fei, Y., Wei, L., and Gao, S. (2002). Differential scanning calorimetric and circular dichroistic studies on plant antifreeze proteins. *J. Therm. Anal. Calorim.* 67 (3), 689–698. doi:10.1023/a:1014369208229

Lv, K., Li, J., Zhao, K., Chen, S., Nie, J., Zhang, W., et al. (2020). Overexpression of an AP2/ERF family gene, BpERF13, in birch enhances cold tolerance through upregulating CBF genes and mitigating reactive oxygen species. *Plant Sci.* 292, 110375. doi:10.1016/j.plantsci.2019.110375

Maddah, M., Shahabi, M., and Peyvandi, K. (2021). How does dc AFP, a plant antifreeze protein, control ice inhibition through the kelvin effect? *Ind. Eng. Chem. Res.* 60 (50), 18230–18242. doi:10.1021/acs.iecr.1c02559

Markovskaya, E., Sysoeva, M., and Sherudilo, E. (2008). The effect of daily exposure to low hardening temperature on plant vital activity. *Russ. J. Dev. Biol.* 39, 261–268. doi:10.1134/S1062360408050019

Mazars, C., Thuleau, P., Lamotte, O., and Bourque, S. (2010). Cross-talk between ROS and calcium in regulation of nuclear activities. *Mol. Plant* 3 (4), 706–718. doi:10.1093/mp/ssq024

Meena, M., Divyanshu, K., Kumar, S., Swapnil, P., Zehra, A., Shukla, V., et al. (2019). Regulation of L-proline biosynthesis, signal transduction, transport, accumulation and its vital role in plants during variable environmental conditions. *Heliyon* 5 (12), e02952. doi:10.1016/j.heliyon.2019.e02952

Megha, S., Basu, U., and Kav, N. N. (2018). Regulation of low temperature stress in plants by microRNAs. *Plant Cell Environ.* 41, 1–15. doi:10.1111/pce.12956

Middleton, A. J., Brown, A. M., Davies, P. L., and Walker, V. K. (2009). Identification of the ice-binding face of a plant antifreeze protein. *FEBS Lett.* 583 (4), 815–819. doi:10.1016/j.febslet.2009.01.035

Ming, R., Zhang, Y., Wang, Y., Khan, M., Dahro, B., and Liu, J. H. (2021). The JA-responsive MYC2-BADH-like transcriptional regulatory module in *Poncirus trifoliata* contributes to cold tolerance by modulation of glycine betaine biosynthesis. *New Phytol.* 229 (5), 2730–2750. doi:10.1111/nph.17063

Miranda, J. A., Avonce, N., Suárez, R., Thevelein, J. M., Van Dijck, P., and Iturriaga, G. (2007). A bifunctional TPS-TPP enzyme from yeast confers tolerance to multiple and extreme abiotic-stress conditions in transgenic Arabidopsis. *Planta* 226 (6), 1411–1421. doi:10.1007/s00425-007-0579-y

Mitchell, D. E., Congdon, T., Rodger, A., and Gibson, M. I. (2015). Gold nanoparticle aggregation as a probe of antifreeze (glyco) protein-inspired ice

recrystallization inhibition and identification of new IRI active macromolecules. *Sci. Rep.* 5 (1), 15716–15810. doi:10.1038/srep15716

Mittler, R. (2002). Oxidative stress, antioxidants and stress tolerance. *Trends Plant Sci.* 7 (9), 405–410. doi:10.1016/S1360-1385(02)02312-9

Miura, K., Jin, J. B., Lee, J., Yoo, C. Y., Stirm, V., Miura, T., et al. (2007). SIZ1-mediated sumoylation of ICE1 controls CBF3/DREB1A expression and freezing tolerance in Arabidopsis. *Plant Cell* 19 (4), 1403–1414. doi:10.1105/tpc.106.048397

Miura, K., Ohta, M., Nakazawa, M., Ono, M., and Hasegawa, P. M. (2011). ICE1 Ser403 is necessary for protein stabilization and regulation of cold signaling and tolerance. *Plant J.* 67 (2), 269–279. doi:10.1111/j.1365-313X.2011.04589.x

Mondal, S., and Pai, P. P. (2014). Chou's pseudo amino acid composition improves sequence-based antifreeze protein prediction. *J. Theor. Biol.* 356, 30–35. doi:10.1016/j.jtbi.2014.04.006

Muthukumaran, J., Manivel, P., Kannan, M., Jeyakanthan, J., and Krishna, R. (2011). A framework for classification of antifreeze proteins in over wintering plants based on their sequence and structural features. *J. Bioinforma. Sequence Analysis* 3 (4), 70–88. doi:10.5897/JBSA11.003

Novillo, F., Medina, J., and Salinas, J. (2007). Arabidopsis CBF1 and CBF3 have a different function than CBF2 in cold acclimation and define different gene classes in the CBF regulon. *Proc. Natl. Acad. Sci. U. S. A.* 104 (52), 21002–21007. doi:10.1073/pnas.0705639105

Oufir, M., Legay, S., Nicot, N., Van Moer, K., Hoffmann, L., Renaut, J., et al. (2008). Gene expression in potato during cold exposure: Changes in carbohydrate and polyamine metabolisms. *Plant Sci.* 175 (6), 839–852. doi:10.1016/j.plantsci.2008.08.010

Pareek, A., Khurana, A., K Sharma, A., and Kumar, R. (2017). An overview of signaling regulons during cold stress tolerance in plants. *Curr. Genomics* 18 (6), 498–511. doi:10.2174/1389202918666170228141345

Park, J. I., Lee, J. H., Gwak, Y., Kim, H. J., Jin, E., and Kim, Y. P. (2013). Frozen assembly of gold nanoparticles for rapid analysis of antifreeze protein activity. *Biosens. Bioelectron.* 41, 752–757. doi:10.1016/j.bios.2012.09.052

Park, J., Lim, C. J., Shen, M., Park, H. J., Cha, J. Y., Iniesto, E., et al. (2018). Epigenetic switch from repressive to permissive chromatin in response to cold stress. *Proc. Natl. Acad. Sci. U. S. A.* 115 (23), E5400–E5409. doi:10.1073/pnas.1721241115

Pavangadkar, K., Thomashow, M. F., and Triezenberg, S. J. (2010). Histone dynamics and roles of histone acetyltransferases during cold-induced gene regulation in Arabidopsis. *Plant Mol. Biol.* 74, 183–200. doi:10.1007/s11103-010-9665-9

Poovaiah, B. W., and Du, L. (2018). Calcium signaling: Decoding mechanism of calcium signatures. *New Phytol.* 217 (4), 1394–1396. doi:10.1111/nph.15003

Pratiwi, R., Malik, A. A., Schaduangrat, N., Prachayasittikul, V., Wikberg, J. E., Nantasenamat, C., et al. (2017). CryoProtect: A web server for classifying antifreeze proteins from nonantifreeze proteins. *J. Chem.*, 2017. 1–15. doi:10.1155/2017/9861752

Primo-Capella, A., Martínez-Cuenca, M. R., Gil-Muñoz, F., and Forner-Giner, M. A. (2021). Physiological characterization and proline route genes quantification under long-term cold stress in *Carrizo citrange*. *Sci. Hortic.* 276, 109744. doi:10.1016/j.scienta.2020.109744

Provesi, J. G., Neto, P. A. V., Arisi, A. C. M., and Amante, E. R. (2019). Extraction of antifreeze proteins from cold acclimated leaves of *Drimys angustifolia* and their application to star fruit (*Averrhoa carambola*) freezing. *Food Chem.* 289, 65–73. doi:10.1016/j.foodchem.2019.03.055

Puyaubert, J., and Baudouin, E. (2014). New clues for a cold case: Nitric oxide response to low temperature. *Plant Cell Environ.* 37 (12), 2623–2630. doi:10.1111/pce.12329

Ramazan, S., Nazir, I., Yousuf, W., John, R., and Allakhverdiev, S. (2022). Environmental stress tolerance in maize (*Zea mays*): Role of polyamine metabolism. *Funct. Plant Biol.*. doi:10.1071/FP21324

Rinehart, J. P., Li, A., Yocum, G. D., Robich, R. M., Hayward, S. A., and Denlinger, D. L. (2007). Up-regulation of heat shock proteins is essential for cold survival during insect diapause. *Proc. Natl. Acad. Sci. U. S. A.* 104 (27), 11130–11137. doi:10.1073/pnas.0703538104

Rivero, R. M., Mittler, R., Blumwald, E., and Zandalinas, S. I. (2022). Developing climate-resilient crops: Improving plant tolerance to stress combination. *Plant J.* 109 (2), 373–389. doi:10.1111/tpj.15483

Robe, K., Izquierdo, E., Vignols, F., Rouached, H., and Dubos, C. (2021). The coumarins: Secondary metabolites playing a primary role in plant nutrition and health. *Trends Plant Sci.* 26 (3), 248–259. doi:10.1016/j.tplants.2020.10.008

Robles, V., G Valcarce, D., and F Riesco, M. (2019). The use of antifreeze proteins in the cryopreservation of gametes and embryos. *Biomolecules* 9 (5), 181. doi:10.3390/biom9050181

Saddhe, A. A., Manuka, R., and Penna, S. (2021). Plant sugars: Homeostasis and transport under abiotic stress in plants. *Physiol. Plant.* 171 (4), 739–755. doi:10.1111/ppl.13283

Schubert, D., Primavesi, L., Bishopp, A., Roberts, G., Doonan, J., Jenuwein, T., et al. (2006). Silencing by plant Polycomb-group genes requires dispersed trimethylation of histone H3 at lysine 27. *EMBO J.* 25, 4638–4649. doi:10.1038/sj.emboj.7601311

Sharma, B., Gupta, R., Sahoo, D., and Deswal, R. (2019). Purification of dual-functioning chitinases with hydrolytic and antifreeze activities from *Hippophae rhamnoides* seedlings. *J. Proteins Proteom.* 10 (1), 69–81. doi:10.1007/s42485-019-00007-9

Sharma, P., Jha, A. B., Dubey, R. S., and Pessarakli, M. (2021). Reactive oxygen species generation, hazards, and defense mechanisms in plants under environmental (abiotic and biotic) stress conditions. *Handb. Plant Crop Physiology*, 617–658. doi:10.1201/9781003093640-37

Shi, H., Chen, Y., Qian, Y., and Chan, Z. (2015). Low temperature-induced 30 (LTI30) positively regulates drought stress resistance in *Arabidopsis*: Effect on abscisic acid sensitivity and hydrogen peroxide accumulation. *Front. Plant Sci.* 6, 893. doi:10.3389/fpls.2015.00893

Shi, H. T., Li, R. J., Cai, W., Liu, W., Fu, Z. W., and Lu, Y. T. (2012). *In vivo* role of nitric oxide in plant response to abiotic and biotic stress. *Plant Signal. Behav.* 7 (3), 437–439. doi:10.4161/psb.19219

Shi, Y., Ding, Y., and Yang, S. (2018). Molecular regulation of CBF signaling in cold acclimation. *Trends Plant Sci.* 23 (7), 623–637. doi:10.1016/j.tplants.2018.04.002

Shkryl, Y., Yugay, Y., Avramenko, T., Grigorchuk, V., Gorpenchenko, T., Grischenko, O., et al. (2021). CRISPR/Cas9-Mediated knockout of *HOS1* reveals its role in the regulation of secondary metabolism in *Arabidopsis thaliana*. *Plants* 10 (1), 104. doi:10.3390/plants10010104

Sidebottom, C., Buckley, S., Pudney, P., Twigg, S., Jarman, C., Holt, C., et al. (2000). Heat-stable antifreeze protein from grass. *Nature* 406 (6793), 256. doi:10.1038/35018639

Singh, P., Hanada, Y., Singh, S. M., and Tsuda, S. (2014). Antifreeze protein activity in Arctic cryoconite bacteria. *FEMS Microbiol. Lett.* 351 (1), 14–22. doi:10.1111/1574-6968.12345

Smallwood, M., Worrall, D., Byass, L., Elias, L., Ashford, D., Doucet, C. J., et al. (1999). Isolation and characterization of a novel antifreeze protein from carrot (*Daucus carota*). *Biochem. J.* 340 (2), 385–391. doi:10.1042/bj3400385

Sohag, A. A. M., Tahjib-Ul-Arif, M., Polash, M. A. S., Belal Chowdhury, M., Afrin, S., Burritt, D. J., et al. (2020). Exogenous glutathione-mediated drought stress tolerance in rice (*Oryza sativa* L.) is associated with lower oxidative damage and favorable ionic homeostasis. *Iran. J. Sci. Technol. Trans. Sci.* 44 (4), 955–971. doi:10.1007/s40995-020-00917-0

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

Song, Y., Zhang, X., Li, M., Yang, H., Fu, D., Lv, J., et al. (2021). The direct targets of CBFs: In cold stress response and beyond. *J. Integr. Plant Biol.* 63 (11), 1874–1887. doi:10.1111/jipb.13161

Sorhannus, U. (2012). Evolution of type II antifreeze protein genes in teleost fish: A complex scenario involving lateral gene transfers and episodic directional selection. *Evol. Bioinform. Online* 8, 535–544. EBO-S9976. doi:10.4137/EBO.S9976

Steponkus, P. L., and Lynch, D. V. (1989). Freeze/thaw-induced destabilization of the plasma membrane and the effects of cold acclimation. *J. Bioenerg. Biomembr.* 21 (1), 21–41. doi:10.1007/BF00762210

Steponkus, P. L. (1984). Role of the plasma membrane in freezing injury and cold acclimation. *Annu. Rev. Plant Physiol.* 35 (1), 543–584. doi:10.1146/annurev.pp.35.060184.002551

Stevens, C. A., Gibson, M. I., and Klok, H. A. (2022). Natural and synthetic macromolecules that interact with ice. *Biomacromolecules* 23 (2), 465–466. doi:10.1021/acs.biomac.2c00093

Su, C. C., Zhai, H. Q., Wang, C. M., Sun, L. H., and Wan, J. M. (2006). SSR mapping of Brown planthopper resistance gene Bph9 in Kaharamana, an Indica rice (*Oryza sativa* L.). *Yi chuan xue bao = Acta Genet. Sin.* 33 (3), 262–268. doi:10.1016/S0379-4172(06)60049-8

Su, J., and Wu, R. (2004). Stress-inducible synthesis of proline in transgenic rice confers faster growth under stress conditions than that with constitutive synthesis. *Plant Sci.* 166 (4), 941–948. doi:10.1016/j.plantsci.2003.12.004

Su, Y., Peng, X., and Shen, S. (2022). Identification of leucine-rich repeat receptor-like protein kinase (LRR-RLK) genes in paper mulberry and their potential roles in response to cold stress. *Comput. Biol. Chem.* 97. 107622. doi:10.1016/j.compbiolchem.2022.107622

Sun, Y., Giubertoni, G., Bakker, H. J., Liu, J., Wagner, M., Ng, D. Y., et al. (2021). Disaccharide residues are required for native antifreeze glycoprotein activity. *Biomacromolecules* 22 (6), 2595–2603. doi:10.1021/acs.biomac.1c00313

Sunkar, R., and Zhu, J.-K. (2004). Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. *Plant Cell* 16, 2001–2019. doi:10.1105/tpc.104.022830

Takahashi, D., Li, B., Nakayama, T., Kawamura, Y., and Uemura, M. (2013). Plant plasma membrane proteomics for improving cold tolerance. *Front. Plant Sci.* 4, 90. doi:10.3389/fpls.2013.00090

Tarkowski, Ł. P., and Van den Ende, W. (2015). Cold tolerance triggered by soluble sugars: A multifaceted countermeasure. *Front. Plant Sci.* 6, 203. doi:10.3389/fpls.2015.00203

Tasaki, Y., and Okada, T. (2008). Ice chromatography, A new physicochemical approach to water-ice surface. *Proc. Whistler 2008 Int. Snow Sci. Workshop Sept.* 21-27, 20081027.

Theocharis, A., Clément, C., and Barka, E. A. (2012). Physiological and molecular changes in plants grown at low temperatures. *Planta* 235 (6), 1091–1105. doi:10.1007/s00425-012-1641-y

Thomashow, M. F. (2010). Molecular basis of plant cold acclimation: Insights gained from studying the CBF cold response pathway. *Plant Physiol.* 154 (2), 571–577. doi:10.1104/pp.110.161794

Thomashow, M. F. (1999). Plant cold acclimation: Freezing tolerance genes and regulatory mechanisms. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 50 (1), 571–599. doi:10.1146/annurev.arplant.50.1.571

Tiburcio, A. F., Altabella, T., Bitrián, M., and Alcázar, R. (2014). The roles of polyamines during the lifespan of plants: From development to stress. *Planta* 240 (1), 1–18. doi:10.1007/s00425-014-2055-9

Tomás, R. M. F., Bailey, T. L., Hasan, M., and Gibson, M. I. (2019). Extracellular antifreeze protein significantly enhances the cryopreservation of cell monolayers. *Biomacromolecules* 20 (10), 3864–3872. doi:10.1021/acs.biomac.9b00951

Uemura, M., and Steponkus, P. L. (1999). Cold acclimation in plants: Relationship between the lipid composition and the cryostability of the plasma membrane. *J. Plant Res.* 112 (2), 245–254. doi:10.1007/PL00013882

Uemura, M., Tominaga, Y., Nakagawara, C., Shigematsu, S., Minami, A., and Kawamura, Y. (2006). Responses of the plasma membrane to low temperatures. *Physiol. Plant.* 126 (1), 81–89. doi:10.1111/j.1399-3054.2005.00594.x

Urbańczyk, M., Góra, J., Latajka, R., and Sewald, N. (2017). Antifreeze glycopeptides: From structure and activity studies to current approaches in chemical synthesis. *Amino Acids* 49 (2), 209–222. doi:10.1007/s00726-016-2368-z

Usman, M., Khan, S., and Lee, J. A. (2020). Afp-lse: Antifreeze proteins prediction using latent space encoding of composition of k-spaced amino acid pairs. *Sci. Rep.* 10 (1), 7197–7213. doi:10.1038/s41598-020-63259-2

Verhage, L., and Mittler, R. (2021). Integration of electric, calcium, reactive oxygen species and hydraulic signals during rapid systemic signaling in plants. *Plant J.* 107 (1), 7–20. doi:10.1111/tpj.15360

Wang, B., Chai, H., Zhong, Y., Shen, Y., Yang, W., Chen, J., et al. (2020a). The DEAD-box RNA helicase SHI2 functions in repression of salt-inducible genes and regulation of cold-inducible gene splicing. *J. Exp. Bot.* 71 (4), 1598–1613. doi:10.1093/jxb/erz523

Wang, J., Ren, Y., Liu, X., Luo, S., Zhang, X., Liu, X., et al. (2021). Transcriptional activation and phosphorylation of OsCNGC9 confer enhanced chilling tolerance in rice. *Mol. Plant* 14 (2), 315–329. doi:10.1016/j.molp.2020.11.022

Wang, Q., Guan, Y., Wu, Y., Chen, H., Chen, F., and Chu, C. (2008). Overexpression of a rice OsDREB1F gene increases salt, drought, and low temperature tolerance in both *Arabidopsis* and rice. *Plant Mol. Biol.* 67 (6), 589–602. doi:10.1007/s11103-008-9340-6

Wang, X., Feng, C., Tian, L., Hou, C., Tian, W., Hu, B., et al. (2021b). A transceptor–channel complex couples nitrate sensing to calcium signaling in *Arabidopsis*. *Mol. Plant* 14 (5), 774–786. doi:10.1016/j.molp.2021.02.005

Wąsek, I., Dyda, M., Gołębiowska, G., Tyrka, M., Rapacz, M., Szechyńska-Hebda, M., et al. (2022). Quantitative trait loci and candidate genes associated with freezing tolerance of winter triticale (× *Triticosecale Wittmack*). *J. Appl. Genet.* 63 (1), 15–33. doi:10.1007/s13353-021-00660-1

Wisniewski, M., Webb, R., Balsamo, R., Close, T. J., Yu, X. M., and Griffith, M. (1999). Purification, immunolocalization, cryoprotective, and antifreeze activity of PCA60: A dehydrin from peach (*Prunus persica*). *Physiol. Plant.* 105 (4), 600–608. doi:10.1034/j.1399-3054.1999.105402.x

Wu, F., Chi, Y., Jiang, Z., Xu, Y., Xie, L., Huang, F., et al. (2020). Hydrogen peroxide sensor HPCA1 is an LRR receptor kinase in Arabidopsis. *Nature* 578 (7796), 577–581. doi:10.1038/s41586-020-2032-3

Wu, P., Xiao, C., Cui, J., Hao, B., Zhang, W., Yang, Z., et al. (2021). Nitric oxide and its interaction with hydrogen peroxide enhance plant tolerance to low temperatures by improving the efficiency of the calvin cycle and the ascorbate–glutathione cycle in cucumber seedlings. *J. Plant Growth Regul.* 40 (6), 2390–2408. doi:10.1007/s00344-020-10242-w

Xiang, H., Yang, X., Ke, L., and Hu, Y. (2020). The properties, biotechnologies, and applications of antifreeze proteins. *Int. J. Biol. Macromol.* 153, 661–675. doi:10.1016/j.ijbiomac.2020.03.040

Xiao, X., Hui, M., and Liu, Z. (2016). iAFP-Ense: an ensemble classifier for identifying antifreeze protein by incorporating grey model and PSSM into PseAAC. *J. Membr. Biol.* 249 (6), 845–854. doi:10.1007/s00232-016-9935-9

Yang, R., Hong, Y., Ren, Z., Tang, K., Zhang, H., Zhu, J.-K., et al. (2019). A role for PICKLE in the regulation of cold and salt stress tolerance in. *Arabidopsis. Front. plant Sci.*. 900. doi:10.3389/fpls.2019.00900

Yang, R., Zhang, C., Gao, R., and Zhang, L. (2015). An effective antifreeze protein predictor with ensemble classifiers and comprehensive sequence descriptors. *Int. J. Mol. Sci.* 16 (9), 21191–21214. doi:10.3390/ijms160921191

You, J., and Chan, Z. (2015). ROS regulation during abiotic stress responses in crop plants. *Front. Plant Sci.* 6, 1092. doi:10.3389/fpls.2015.01092

Yu, H., Zheng, H., Liu, Y., Yang, Q., Li, W., Zhang, Y., et al. (2021). Antifreeze protein from *Ammopiptanthus nanus* functions in temperature-stress through domain A. *Sci. Rep.* 11 (1), 8458–8510. doi:10.1038/s41598-021-88021-0

Yu, T. F., Xu, Z. S., Guo, J. K., Wang, Y. X., Abernathy, B., Fu, J. D., et al. (2017). Improved drought tolerance in wheat plants overexpressing a synthetic bacterial cold shock protein gene SeCspA. *Sci. Rep.* 7 (1), 44050–44114. doi:10.1038/srep44050

Yuan, P., Yang, T., and Poovaiah, B. W. (2018). Calcium signaling mediated plant response to cold stress. *Int. J. Mol. Sci.* 19 (12), 3896. doi:10.3390/ijms19123896

Zanetti-Polzi, L., Biswas, A. D., Del Galdo, S., Barone, V., and Daidone, I. (2019). Hydration shell of antifreeze proteins: Unveiling the role of non-ice-binding surfaces. *J. Phys. Chem. B* 123 (30), 6474–6480. doi:10.1021/acs.jpcb.9b06375

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

Zhan, X., Wang, B., Li, H., Liu, R., Kalia, R. K., Zhu, J.-K., et al. (2012). Arabidopsis proline-rich protein important for development and abiotic stress tolerance is involved in microRNA biogenesis. *Proc. Natl. Acad. Sci. U. S. A.* 109, 18198–18203. doi:10.1073/pnas.1216199109

Zhang, B. Q., Song, X. P., Zhang, X. Q., Huang, Y. X., Liang, Y. J., Zhou, S., et al. (2022). Differential gene expression analysis of SoCBL family calcineurin B-like proteins: Potential involvement in sugarcane cold stress. *Genes* 13 (2), 246. doi:10.3390/genes13020246

Zhang, C., Fei, S. Z., Arora, R., and Hannapel, D. J. (2010). Ice recrystallization inhibition proteins of perennial ryegrass enhance freezing tolerance. *Planta* 232 (1), 155–164. doi:10.1007/s00425-010-1163-4

Zhang, D. Q., Liu, B., Feng, D. R., He, Y. M., and Wang, J. F. (2004). Expression, purification, and antifreeze activity of carrot antifreeze protein and its mutants. *Protein Expr. Purif.* 35 (2), 257–263. doi:10.1016/j.pep.2004.01.019

Zhang, H., Dong, J., Zhao, X., Zhang, Y., Ren, J., Xing, L., et al. (2019a). Research progress in membrane lipid metabolism and molecular mechanism in peanut cold tolerance. *Front. Plant Sci.* 10, 838. doi:10.3389/fpls.2019.00838

Zhang, H., Zhang, Z., Xiong, Y., Shi, J., Chen, C., Pan, Y., et al. (2021a). Stearic acid desaturase gene negatively regulates the thermotolerance of *Pinellia ternata* by modifying the saturated levels of fatty acids. *Industrial Crops Prod.* 166, 113490. doi:10.1016/j.indcrop.2021.113490

Zhang, H., Zhu, J., Gong, Z., and Zhu, J. K. (2021b). Abiotic stress responses in plants. *Nat. Rev. Genet.* 23, 104–119. doi:10.1038/s41576-021-00413-0

Zhang, L., Zhao, T., Sun, X., Wang, Y., Du, C., Zhu, Z., et al. (2019b). Overexpression of VaWRKY12, a transcription factor from *Vitis amurensis* with increased nuclear localization under low temperature, enhances cold tolerance of plants. *Plant Mol. Biol.* 100 (1), 95–110. doi:10.1007/s11103-019-00846-6

Zhang, Q., Chen, Q., Wang, S., Hong, Y., and Wang, Z. (2014). Rice and cold stress: Methods for its evaluation and summary of cold tolerance-related quantitative trait loci. *Rice* 7 (1), 24–12. doi:10.1186/s12284-014-0024-3

Zhang, X., Cheng, X., Yu, L., Yang, J., Calvo, R., Patnaik, S., et al. (2016). MCOLN1 is a ROS sensor in lysosomes that regulates autophagy. *Nat. Commun.* 7 (1), 12109–12112. doi:10.1038/ncomms12109

Zhang, Y., Cao, Y., Zheng, H., Feng, W., Qu, J., Fu, F., et al. (2021c). Ectopic expression of antifreeze protein gene from *Ammopiptanthus nanus* confers chilling tolerance in maize. *Crop J.* 9 (4), 924–933. doi:10.1016/j.cj.2020.08.011

Zhao, C., Lang, Z., and Zhu, J. K. (2015). Cold responsive gene transcription becomes more complex. *Trends Plant Sci.* 20 (8), 466–468. doi:10.1016/j.tplants.2015.06.001

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 (2), 755–767. doi:10.1104/pp.109.140996

Zhao, X., Ma, Z., and Yin, M. (2012). Using support vector machine and evolutionary profiles to predict antifreeze protein sequences. *Int. J. Mol. Sci.* 13, 2196–2207.

Zhao, X., Wei, Y., Zhang, J., Yang, L., Liu, X., Zhang, H., et al. (2021). Membrane lipids' metabolism and transcriptional regulation in maize roots under cold stress. *Front. Plant Sci.* 12, 639132. doi:10.3389/fpls.2021.639132

Zheng, S., Su, M., Wang, L., Zhang, T., Wang, J., Xie, H., et al. (2021). Small signaling molecules in plant response to cold stress. *J. Plant Physiol.* 266, 153534. doi:10.1016/j.jplph.2021.153534

Zhou, A., Liu, E., Li, H., Li, Y., Feng, S., Gong, S., et al. (2018). PsCor413pm2, a plasma membrane-localized, cold-regulated protein from *Phlox subulata*, confers low temperature tolerance in. *Int. J. Mol. Sci.* 19 (9), 2579. doi:10.3390/ijms19092579

Zhu, J., Jeong, J. C., Zhu, Y., Sokolchik, I., Miyazaki, S., Zhu, J.-K., et al. (2008). Involvement of Arabidopsis HOS15 in histone deacetylation and cold tolerance. *Proc. Natl. Acad. Sci. U. S. A.* 105, 4945–4950. doi:10.1073/pnas.0801029105

Zhuang, X., Yang, C., Murphy, K. R., and Cheng, C. H. C. (2019). Molecular mechanism and history of non-sense to sense evolution of antifreeze glycoprotein gene in northern gadids. *Proc. Natl. Acad. Sci. U. S. A.* 116 (10), 4400–4405. doi:10.1073/pnas.1817138116

Zuther, E., Schaarschmidt, S., Fischer, A., Erban, A., Pagter, M., Mubeen, U., et al. (2019). Molecular signatures associated with increased freezing tolerance due to low temperature memory in Arabidopsis. *Plant Cell Environ.* 42, 854–873. doi:10.1111/pce.13502

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 16 September 2022
DOI 10.3389/fgene.2022.958217

Check for updates

OPEN ACCESS

EDITED BY
Dragan Perovic,
Julius Kühn-Institut, Germany

REVIEWED BY
Bobin Liu,
Yancheng Teachers University, China
Gursharn Singh Randhawa,
Sardar Bhagwan Singh Post Graduate Institute of Biomedical Science & Research, India
Ravinder Singh,
Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu, India
Upendra Kumar Devisetty,
GreenLight Biosciences, United States

*CORRESPONDENCE
Preetesh Kumari,
preetesh79@gmail.com

SPECIALTY SECTION
This article was submitted to Plant Genomics, a section of the journal Frontiers in Genetics

RECEIVED 31 May 2022
ACCEPTED 23 August 2022
PUBLISHED 16 September 2022

CITATION
Singh KP, Kumari P and Yadava DK (2022), Development of *de-novo* transcriptome assembly and SSRs in allohexaploid Brassica with functional annotations and identification of heat-shock proteins for thermotolerance. *Front. Genet.* 13:958217. doi: 10.3389/fgene.2022.958217

COPYRIGHT
© 2022 Singh, Kumari and Yadava. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Development of *de-novo* transcriptome assembly and SSRs in allohexaploid Brassica with functional annotations and identification of heat-shock proteins for thermotolerance

Kaushal Pratap Singh[1], Preetesh Kumari[2]* and Devendra Kumar Yadava[2]

[1]ICAR-Directorate of Rapeseed Mustard Research, Bharatpur, India, [2]Genetics Division, ICAR—Indian Agricultural Research Institute, New Delhi, India

Crop Brassicas contain monogenomic and digenomic species, with no evidence of a trigenomic Brassica in nature. Through somatic fusion (*Sinapis alba* + *B. juncea*), a novel allohexaploid trigenomic Brassica (H1 = AABBSS; 2*n* = 60) was produced and used for transcriptome analysis to uncover genes for thermotolerance, annotations, and microsatellite markers for future molecular breeding. Illumina Novaseq 6000 generated a total of 76,055,546 paired-end raw reads, which were used for *de-novo* assembly, resulting in the development of 486,066 transcripts. A total of 133,167 coding sequences (CDSs) were predicted from transcripts with a mean length of 507.12 bp and 46.15% GC content. The BLASTX search of CDSs against public protein databases showed a maximum of 126,131 (94.72%) and a minimum of 29,810 (22.39%) positive hits. Furthermore, 953,773 gene ontology (GO) terms were found in 77,613 (58.28%) CDSs, which were divided into biological processes (49.06%), cellular components (31.67%), and molecular functions (19.27%). CDSs were assigned to 144 pathways by a pathway study using the KEGG database and 1,551 pathways by a similar analysis using the Reactome database. Further investigation led to the discovery of genes encoding over 2,000 heat shock proteins (HSPs). The discovery of a large number of HSPs in allohexaploid Brassica validated our earlier findings for heat tolerance at seed maturity. A total of 15,736 SSRs have been found in 13,595 CDSs, with an average of one SSR per 4.29 kb length and an SSR frequency of 11.82%. The first transcriptome assembly of a meiotically stable allohexaploid Brassica has been given in this article, along with functional annotations and the presence of SSRs, which could aid future genetic and genomic studies.

KEYWORDS

allohexaploid Brassica, illumina Novaseq 6000, RNA-seq, *de-novo* assembly, functional annotations, SSRs, gene ontology

Introduction

Unlike naturally developed allohexaploid bread wheat (Gaebelein and Mason, 2018), trigenomic "allohexaploid Brassicas" with three genomes do not exist in nature (Mason and Batley, 2015). Such allohexaploid Brassicas have been successfully developed through somatic hybridization, and their potential is being examined, which is adapted to climatic change (Gupta et al., 2016; Wei et al., 2016). Higher ploidy levels are linked to speciation, gene augmentation, and genetic diversity (Leitch and Leitch, 2008; Soltis et al., 2009; Jiao et al., 2011; Tank et al., 2015) and provide genetic resources that can adapt to changing climatic conditions (Marchant et al., 2016). Although attempts to synthesize allohexaploid Brassica through interspecific hybridization have been made (Pradhan et al., 2010; Tian et al., 2010; Chen et al., 2011), such attempts have failed (Mason et al., 2012, 2014). However, wild diploid Brassicaceae members should be added to cultivated species as a source of unique genes for resistance/tolerance to a variety of fungal diseases, insects, pests, nematodes, heat, and drought (Sjodin and Glimelius, 1989; Kirti et al., 1995; Li et al., 2009; Garg et al., 2010; Singh et al., 2021). With this in mind, we developed stable allohexaploid Brassicas through somatic hybridization involving diploid *S. alba* (a Brassica coenospecies) and tetraploid *B. juncea* (an amphidiploid crop species); the resulting stable allohexaploids had a somatic chromosome number of $2n = 60$ (AABBSS) and a high level of male and female fertility (Kumari et al., 2018; 2020c). Two allohexaploids (H1 and H2) showed resistance to Alternaria blight and Sclerotinia stem rot and had high-temperature tolerance (upto 40°C). The H1 allohexaploid also had recombinant mitochondria and *B. juncea*-type chloroplasts, which improved resistance to Alternaria blight and Sclerotinia stem rot (Kumari et al., 2018; Kumari and Singh, 2019), as well as several other desirable traits like thermotolerance, which is partly dependent on ubiquitous heat shock proteins (HSPs) and allows normal cell functions to be maintained during hot spells (Park and Seo, 2015). Under stressful conditions, they can aid to stabilize and refold proteins (molecular chaperons) (Whitley et al., 1999; Sitia and Braakman, 2003; Huttner and Strasser, 2012). HSPs are divided into five primary groups based on their size (kDa): HSP100, HSP90, HSP70, HSP60, and small HSP (sHSP) (Wang et al., 2004; Kotak et al., 2007; Gupta et al., 2010). To respond to stresses, HSPs can be found in the cytosol, endoplasmic reticulum, chloroplast, mitochondria, and nucleus (Vierling, 1991; Boston et al., 1996). The transcriptome profile studies in a large number of crop plants have been conducted to report these heat shock proteins such as Arabidopsis (Rahmati Ishka et al., 2018), rice (Sarkar et al., 2014), maize (Shi et al., 2017), and rapeseed (Wang et al., 2018a).

NGS technologies such as Illumina, Ion Torrent, PcaBio, Nanopore, and others are now being used for RNA sequencing (RNA-seq) for a number of objectives, including genomic architecture study, molecular pathway elucidation, and the production of molecular markers such as SSRs (Shawn and Richard, 2016; Zhang et al., 2017). However, no such research has been done on any stable allohexaploid Brassica (*B. juncea* + *S. alba*) yet. On the other hand, the allotetraploid Brassica napus was frequently used for transcriptome analysis to decipher the genic constitution for important agronomic and yield-related traits. Yao et al. (2020) used a single molecule long-read isoform sequencing (Iso-Seq) technique to unravel the complex nature of transcriptome in *B. napus*. Out of 147,698 unique long-read isoforms identified, a total of 37,403 genes were annotated. Moreover, An et al. (2019) used a different strategy to uncover the genic architecture of *B. napus* by using 183 *B. napus* accessions along with their diploid progenitors, i.e., *B. rapa* (112 accessions) and *B. oleracea* (62 accessions) and five other Brassicaceae members. These accessions collectively represented the complete phenotypic diversity of *B. napus* species and shared a total of 372,546 high-quality SNPs. The identification of a higher number of SNPs in the A subgenome than in the C confirmed the higher level of nuclear diversity. The transcriptome/RNA-seq strategy is successfully used to identify the target candidate gene for a particular trait. Jian et al. (2019) identified 115 flowering time–related differentially expressed genes that were related to plant circadian clock/photoperiod, autonomous pathway, and hormone and vernalization pathways. They identified a total of 27 quantitative loci dispersed on eight different chromosomes of *B. napus*. These loci were identified for harboring 45 candidate genes for flowering time.

In light of the foregoing, the current work was designed to examine the transcriptome of a stable allohexaploid Brassica that we had previously produced through somatic hybridization. The allohexaploid contains unique genes that confer disease resistance and heat stress tolerance. Transcriptome analysis can be used to characterize these new genes. We analyzed gene ontology (GO) and performed functional annotations against common protein databases. The transcriptome information is also used to create genic SSRs. The current research added to the knowledge of the genomic architecture of a synthetic allohexaploid Brassica. The findings of this study will serve as a genetic resource for future research on gene identification and expression patterns, population genetics, pathway investigations, phylogenetics, and marker-assisted selection (MAS).

Materials and methods

Plant material, total RNA isolation, and quality check

The allohexaploid (H1; Kumari and Bhat, 2021) was used as the experimental plant material in this investigation (Kumari et al., 2018). After the commencement of flowering in all field-

grown plants, leaf samples were taken in triplicates from three individual plants in the evening for Total RNA isolation. Until RNA extraction, all samples were collected in liquid nitrogen and stored at −80°C. Total RNA was isolated from each sample using TRIzol Reagent (Invitrogen, CA, United States) according to the manufacturer's instructions. Following Total RNA isolation, the samples were quality checked and quantified using the NanoDrop 8000 (OD260/OD280) and Qubit 3.0 Fluorometer (Invitrogen Life Technologies, United States), respectively, and qualified using a 1% agarose gel electrophoresis to check for RNA degradation and contamination.

cDNA library construction, quality control, and RNA sequencing

All three RNA samples showed nanodrop ratio >2.0 and qubit concentrations 2170, 2240, and 2010 ng/μl, respectively. The equimolar concentration of Total RNA extracted from all three plant samples was combined to create the pooled sample and used for the downstream experiments. For the production of complementary DNA (cDNA) libraries (NEBNext II RNA Library Prep Kit for Illumina®), a total of 1 μg of Total RNA with a RIN (RNA integrity number) value greater than 7 was employed. Following the manufacturer's instructions, the ready-to-run final library was quantified using a Qubit 3.0 fluorometer using a dsDNA HS assay kit (Q32851; Thermofisher, United States). The library's insert size (391 bp) was determined using highly sensitive D1000 ScreenTapes5067-5582 on a TapeStation 4150 (Agilent Technologies, CA, United States) following the manufacturer's protocol. Following the manufacturer's directions, mRNA was purified from Total RNA using poly-T oligo-attached magnetic beads. The mRNA fragmentation buffer was added to break them into small fragments. Random primers were used to synthesize the cDNA strand from cleaved mRNA fragments during reverse transcription. DNA polymerase-I, dNTPs, buffer, and RNase H were used to make the second strand of cDNA. The freshly produced double-stranded cDNA was purified using the QIAquick PCR extraction kit (QIAGEN, Hilden, Germany) and washed with EB buffer. Adapters were ligated to the sequencing RNA on both ends and inserted adenine (A) nucleotide into the 3′ ends to repair the ends. AMPure XP beads were used to select fragments, which were then enriched by PCR to create a library for transcriptome sequencing. Nucleome Informatics Pvt. Ltd. in Hyderabad performed the RNA sequence library preparation and sequencing (India) (https://www.nucleomeinfo.com/). After pooling the qualified libraries according to their effective concentration and expected data volume, they were fed into an Illumina sequencer NovaSeq 6000 with S4 type Flow Cell (2 × 150 bp read length). After sequencing, the raw data in FASTQ format was used for *de-novo* assembly, microsatellite marker development, and functional annotations using bioinformatics tools. The raw data of the allohexaploid Brassica transcriptome sequence were submitted to the Sequence Read Archive (SRA) of the National Center for Biotechnology Information (NCBI) under accession number SRR14934389 and BioProject number PRJNA741791.

De-novo assembly development, redundancy removal, and coding sequence extraction

De-novo assembly was constructed from clean reads with a quality value or Phred score (Q) of 30 after removing low-quality bases and adapters from 5′ and 3′ ends, very short sequences, and low-quality reads to yield robust transcripts of the allohexaploid Brassica. Bioinformatics tools like FastQC and NGS QC toolkits were used to clean the reads. The clean raw reads were assembled *de-novo* into transcripts using the Trinity v2.11.0 software (Grabherr et al., 2011), with default parameters, and k-mer 25, a short-read assembly program (https://github.com/trinityrnaseq/trinityrnaseq/wiki). After a successful run of the software, the assembly was utilized to filter out and identify true transcripts. The evidentiary gene packages (tr2aacds.pl) were used to remove the spurious transcripts while keeping the CDS that was at least 90 bp long (http://arthropods.eugenes.org/EvidentialGene/trassembly.html). The package was run following pipeline: Perfect redundant removal-fastanrdb (from the exonerate programme); perfect fragment identification- CD-HIT-EST –c 1.0 (https://github.com/weizhongli/cdhit); BLASTN to find highly similar transcripts; CDS classification- evigene/rnaseq/asmrna dupfilter2. Pl. The software-generated coding sequences (CDSs) were then used for microsatellite development and functional annotations. The completeness of the transcriptome assembly was also evaluated by OrthoDB database of orthologs to define BUSCO sets for eukaryote clade (https://www.orthodb.org).

Functional annotations

OmicsBox v2.0 (https://www.biobam.com/omicsbox/) was used to mask the coding sequences for repeat sequence types, utilizing the RMBlast search engine and the Dfam v3.0 consensus repeat database, with the Brassicaceae family filter set to default values (Smit et al., 2013; Hubley et al., 2016). The masked CDSs were aligned with the combined protein database of *B. rapa* (GCF_000309985.2_CAAS_Brap_v3.01) and *Arabidopsis thaliana* (GCF_000001735.4_TAIR10.1) reference sequences downloaded from NCBI using local BLASTX at an expectation value of 1e-25 with default parameters. Through OmicsBox, the outcomes of the local blast search were used for downstream analyses, including mapping, annotations, gene

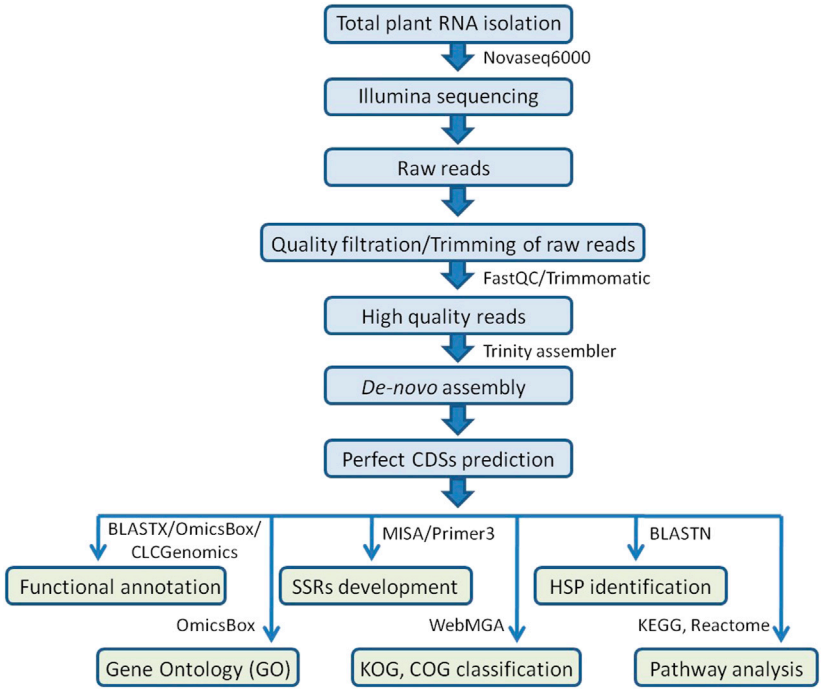


FIGURE 1
Workflow for Illumina sequencing, *de-novo* transcriptome assembly, functional annotations, SSR development, and identification of HSPs carried out in the allohexaploid Brassica.

ontology (GO), and pathway analysis (KEGG and Reactome). After BLASTX, the GO terms were classified into three categories: cellular component, molecular function, and biological process. The CDSs were also aligned using BLASTX (E 1e-25) against NCBI public databases, such as non-redundant protein sequences v5 (nr), reference proteins (refseq protein v5) using Brassicaceae, *B. juncea*, *B. rapa*, *B. nigra*, and *S. alba* as a taxonomic filter using OmicsBox and locally BLASTX against UniProtKB/Swiss-Prot v5 database on the CLC genomics workbench v20.0.4 (Qiagen, United States). Furthermore, utilizing the COG, KOG, Pfam (Stephen et al., 1990), Prk (Finn et al., 2010), and Tigrfam (Haft et al., 2003) protein databases (http://weizhong-lab.ucsd.edu/webMGA/), the protein translations of these CDSs were categorized and examined (Wu et al., 2011). To investigate the activities of putative genes against heat stress during seed development stages in allohexaploid Brassica, the CDSs were subjected to BLASTN with plant heat shock proteins (HSPs) retrieved from the NCBI Gene database. A total of 13,022 HSPs were downloaded, including A. thaliana, *H. syriacus*, *C. sativa*, *B. napus*, *B. rapa*, *H. annus*, *T. dicoccoides*, *P. sominferum*, and other plant genera included. The BLASTN tool was utilized through the CLC genomics workbench to identify the potential HSPs, with a 1e-25 expectation (E) value, a maximum number of hit sequences of 250, and word size of 11.

SSR loci identification and marker development

The MIcroSAtellite (MISA) identification tool Perl script (http://pgrc.ipk-gatersleben.de/misa/misa.html) was used to find probable microsatellite loci in coding sequences extracted from *de-novo* assembled transcripts. The parameters were adjusted to find perfect mono-, di-, tri-, tetra-, penta-, and hexa-nucleotide motifs with a minimum number of 10, 6, 5, 5, 5, and 5 repetitions, respectively. Primer3 v0.4.0 was used to create the primer pairs. For primer pair designing, the following parameters were used: PCR product size: 100–400 bp (optimum: 280 bp), GC content: 45–70% (optimum: 50%), Tm: 57–62°C (optimum: 60°C), primer size: 18–25 bp (optimum: 20 bp), GC content: 45–70% (optimum: 50%), and Tm: 57–62°C (optimum: 60°C) (Figure 1).

Results

De-novo transcriptome assembly

Illumina Novaseq 6000 runs generated a total of 76,055,546 raw paired-end reads, including low-quality sequences, adapter-primer sequences, and very short reads.

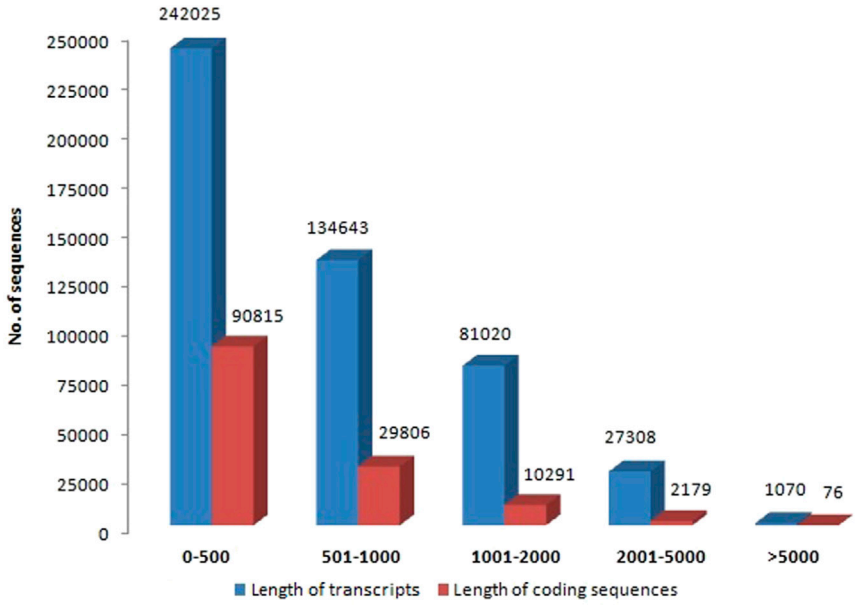


FIGURE 2
Length distribution of allohexaploid Brassica transcripts and coding sequences.

This was accomplished using cDNA generated from the total mRNA of allohexaploid Brassica. The raw reads showed a total GC content of 44.94% and a mean sequence length of 151 bp. At a Phred score (Q) of 30, low-quality bases, adapter-primer sequences, and very short sequences were eliminated. A total of 74,693,278 high-quality clean reads were obtained after a thorough quality check and data filtering, with 11,012,782 reads (98.01%) with Q20 and 10,583,124 reads (94.19%) with Q30. With 44.93% GC content, the clean reads totaled 11,236,399 kb in length. The BUSCO of assembly showed 93.3% completeness for conserved ortholog content. The raw reads (Accession No. SRR14934389) were submitted to the Sequence Read Archive (SRA) database at the National Center for Biotechnology Information (NCBI) in the United States. There were 486,066 transcripts generated from the paired-end reads, with a total length of 368,936,034 bp and an average length of 759.02 bp per transcript. The transcripts N50 and N70 values were 1,069 and 627 bp, respectively, with the longest transcript length of 33,251 bp. The transcripts could be divided into five categories based on their length: 1) 242,025 transcripts (49.79%) with a length of less than 500 bp; 2) 134,643 transcripts (27.7%) with a length range of 501–1,000 bp; 3) 81,020 transcripts (16.67%) with a length range of 1,001–2000 bp; 4) 27,308 transcripts (5.62%) with a length range of 2001–5000 bp; and 5) only 1,070 transcripts (0.22%) with length >5000 bp (Figure 2). From these transcripts, 133,167 coding sequences (CDSs) were predicted, with a total length of 67,531,134 bp and an average length of 507.12 bp, with a range of 93–14,688 bp. CDSs had a GC content of 46.15%, and their N50 and N70 values were 606 bp and 417 bp, respectively. Furthermore, 90,815 CDSs (68.2%) had lengths less than 500 bp; 29,806 CDSs (22.38%) had lengths between 501 and 1,000 bp; 10,291 CDSs (7.73%) had lengths between 1,001–2,000 bp; 2,179 CDSs (1.64%) had lengths between 2,001 and 5,000 bp; and 76 CDSs (0.06%) had lengths greater than 5,000 bp (Figure 2). The statistics of *de-novo* assembly and CDSs are given in Table. 1. The length distribution of CDSs is shown in Figure 3A.

TABLE 1 Summary of results obtained after the development of *de-novo* assembly of allohexaploid Brassica transcriptomes.

Parameters	Transcripts	CDSs
Sum bp	368936034	67531134
Number of sequences	486,066	133,167
Average	759.02	507.12
Largest sequence	33,251	14,688
N50	1,069 (99,977)	606 (30,386)
N60	824 (139312)	498 (42,738)
N70	627 (190765)	417 (57,616)
N80	473 (258626)	351 (75,274)
N90	343 (350035)	297 (96,113)
N100	173 (486066)	93 (133167)
N_count	0	0
Gaps	0	0

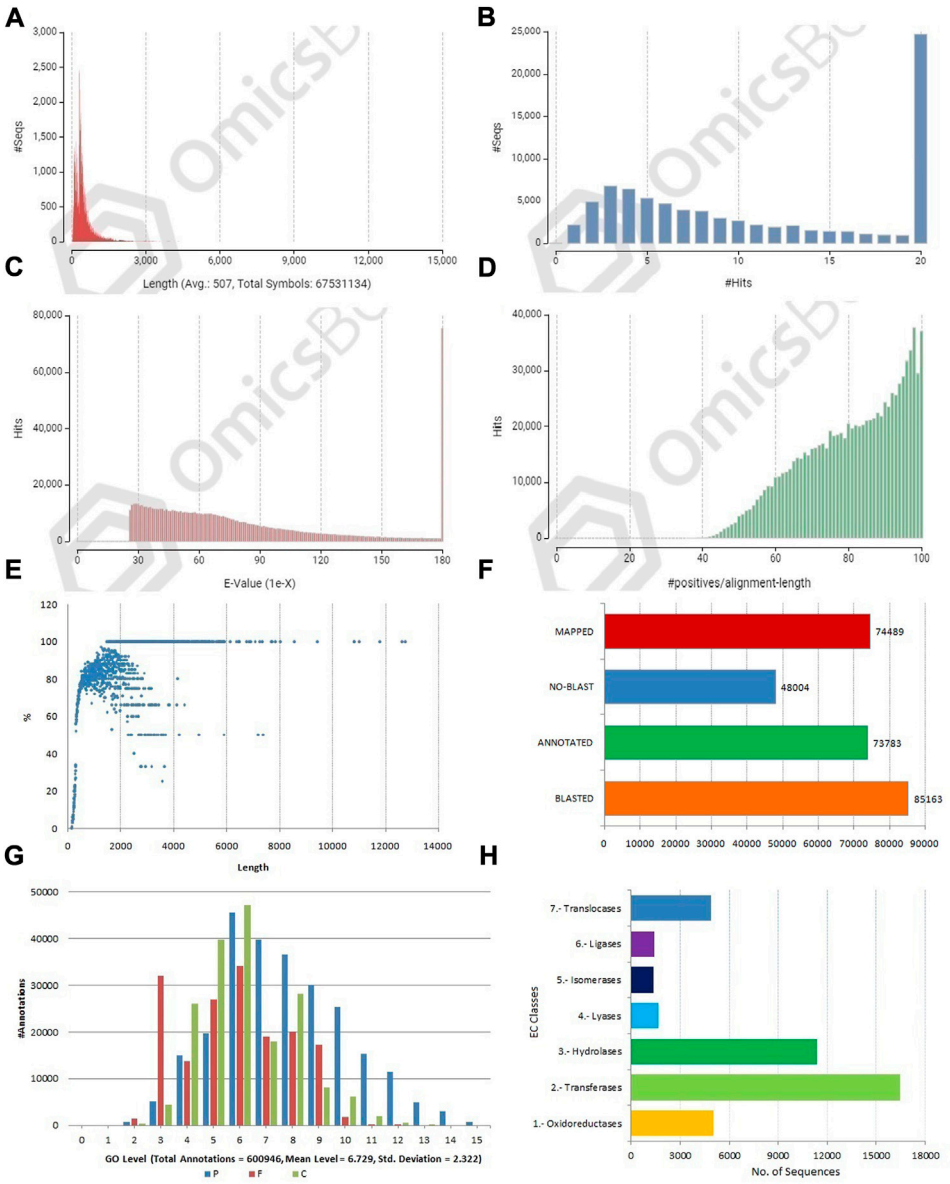


FIGURE 3
Structural and functional annotation of allohexaploid transcriptome assembly. **(A)** Sequence length distribution of coding sequences; **(B)** BLASTX hit distribution; **(C)** E-value distribution; **(D)** Sequence similarity distribution of BLASTX hits; **(E)** Annotation percent of coding sequences against length. **(F)** Three-step distribution of OmicsBox process (BLASTX, mapping, and annotation); **(G)** GO level distribution for biological process, molecular function, and cellular component [P-BP, F-MF, and C-CC]; **(H)** Enzyme code distribution for major classes.

Functional annotations of coding sequences

The CDSs were masked for repetitions by RMBlast search engine using Dfam v3.0 consensus database using Brassicaceae as species filter with OmicsBox tool. A total of 880,817 bp (1.30%) were masked by the search engine, with 1.0% belonging to simple repeats and 0.3% to low complexity sequences. The masked CDSs were used for annotation, and probable functions were assigned to them. The CDSs were locally BLASTX at an E value of 1e-25 against the reference protein sequences of *A. thaliana* and *B. rapa*. Between the query and the local protein database, the blast search found 85,163 (63.95%) unique hits. The remaining CDSs, on the other hand, showed no resemblance to the database (Figures 3B,F). CDS similarity was calculated using E-values ranging from 1e-26 to 1e-180 (Figure 3C). The sequence similarity distribution of BLASTX hits ranged from 36 to 100%, with the largest number of hits (37,672) falling in the

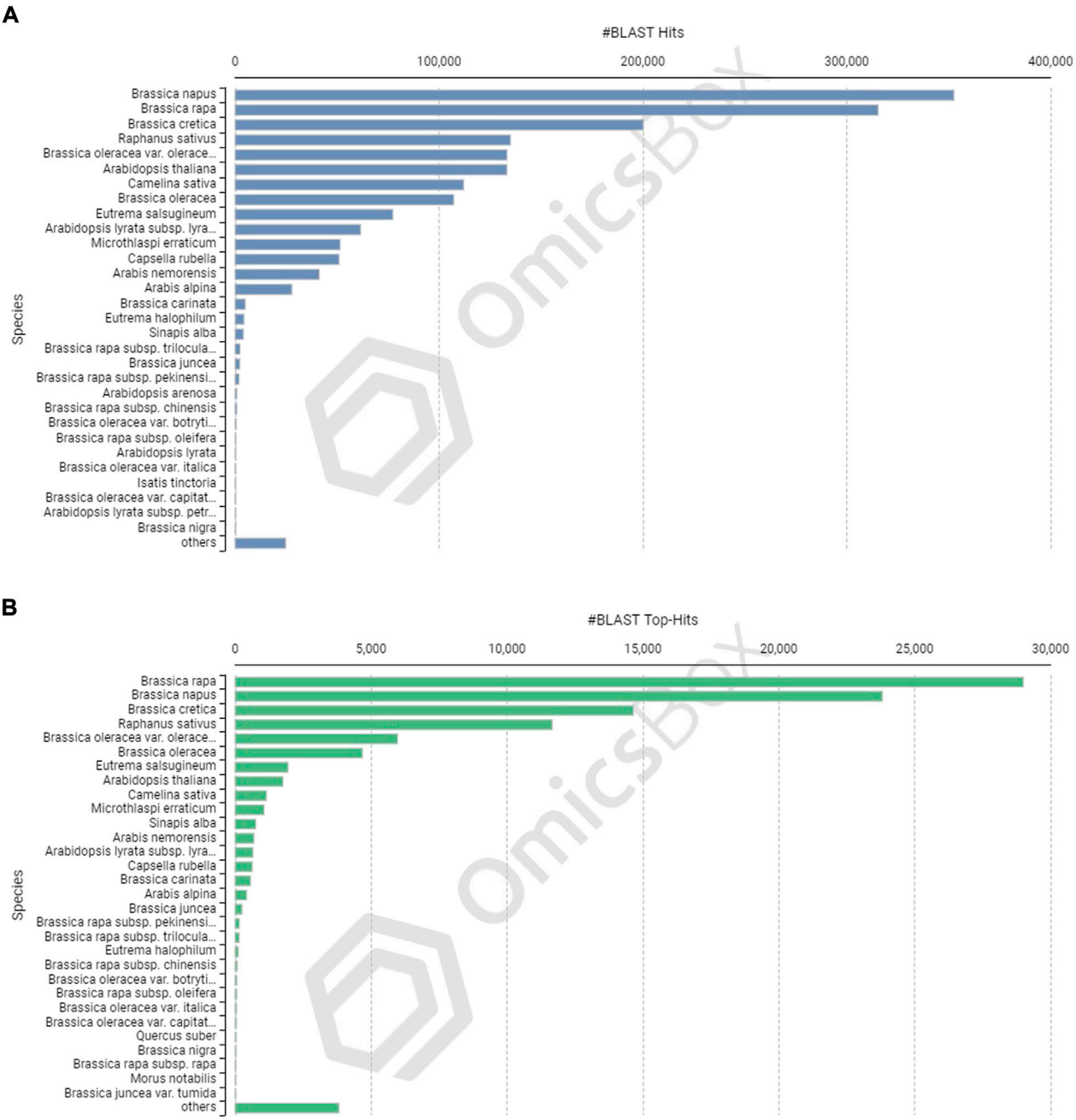


FIGURE 4
Species-based distribution of the BLASTX hits **(A)** and Top BLASTX hits **(B)** for each coding sequence of allohexaploid Brassica.

98% range. In 37,005 BLAST hits, there was an absolute similarity between query and database sequences. However, 56.16% of the hits had a resemblance of more than 80%, whereas only 43.84% of the hits had a similarity of 35%–80% (Figure 3D). The smallest CDS (153 bp) had just 1% annotation, whilst the largest CDS (12,741 bp) had 100% annotation (Figure 3E). A total of 74,489 and 73,783 CDSs were mapped and annotated with GO terms, respectively, according to the tag distribution analysis (Figure 3F). The CDSs had a top-hit similarity of more than 99% with Brassicaceae species, indicating that our transcriptome assembly had a reasonable coverage of homologous sequences. Based on BLASTX results, 600,946 GO level annotations were recorded in the biological process, molecular function, and cellular components categories.



FIGURE 5
Functional classification of gene ontology (GO) terms under biological process **(A)**, cellular component **(B)**, and molecular function **(C)** categories.

Only one GO term was found in a large percentage of CDSs, and the maximum number of GO terms (6,251) was found in CDSs with a length of 315 bp (Figure 3G). There were 42,225 CDSs identified and categorized into seven enzyme coding classes (Figure 3H). The BLASTX program was used to find homology between CDSs of allohexaploid Brassica and CDSs

from other related species in the NCBI Nr database, and it revealed that many plant species in the Brassicaceae family are related (Figures 4A and B).

Gene ontology classification

A total of 77,613 (58.28%) CDSs were assigned to gene ontology terms with a maximum and minimum of 63 and one GO IDs, respectively. A BLASTX search was used to classify the CDSs of allohexaploid Brassica into three functional categories (biological processes, cellular components, and molecular functions). All CDSs that fall under various functional categories were allocated a total of 953,773 GO terms. Biological processes were assigned the greatest number of GO terms (467,893; 49.06%), followed by cellular components (302,064; 31.67%), and molecular functions (183,816; 19.27%). Among the biological processes, organic substance metabolism (43,122; 9.22%) had the highest number of CDSs, followed by cellular metabolism (42,770; 9.14%), primary metabolism (40,506; 8.66%), nitrogen compound metabolism (35,746; 7.64%), biosynthetic process (22,240; 4.75%), and cellular process regulation (21,639; 4.62%). (Figure 5A). In the cellular component category, CDSs in the intracellular anatomical structure (57,234; 18.95%) were the largest group, followed by organelle (52,232; 17.29%), cytoplasm (46,712; 15.46%), and membrane (35,078; 11.61%) (Figure 5B). The CDSs involved in organic cyclic compound binding (30,441; 16.56%) were the most numerous among the molecular functions, followed by heterocyclic compound binding (30,333; 16.50%), protein binding (30,131; 16.39%), and ion binding (24,505; 13.33%). (Figure 5C).

Functional classification of CDSs by protein databases

To obtain the corresponding annotation information, all of the CDSs were BLAST against various protein databases, including Nr, Swiss-Prot, Refseq, COG, KOG, Pfam, Prk, and Tigrfam. A total of 87,781 CDSs (65.92%) were aligned to the Nr database, with about 59,413 CDSs (67.68%) showing more than 90% similarity. The CDSs that corresponded with the database in the case of the Refseq protein database yielded a total of 86,226 positive hits. Furthermore, using CLC genomics workbench, all CDSs were aligned against the Swiss-Prot protein database at an E-value of 1e-25. In CDSs, a total of 126,131 (92.46%) positive hits were found, with 27,704 CDSs sharing more than 90% similarity. Furthermore, BLASTP was performed on these CDSs against the Pfam, Prk, and Tigrfam protein databases, yielding 120,489, 36,476, and 39,368 positive hits, respectively.

The CDSs were annotated against various protein databases, such as COG for prokaryotes and KOG for eukaryotes, using the RPSBLAST software, which was run over the WebMGA web server, with an E-value of 1e-25. In total, 35,763 CDSs were assigned to the COG database's 24 functional classes. General functions (7,737; 21.63%), followed by posttranslational modification, protein turnover, and chaperones (4,254; 11.89%), translation, ribosomal structure, and biogenesis (2,803; 7.84%), carbohydrate transport and metabolism (2,241; 6.26%), unknown function (1,947; 5.45%), and nuclear structure (34; 0.09%) were the classes with the most CDSs (Figure 6A). The KOG database was used to functionally annotate 53,081 CDSs and divide them into 26 types. Signal transduction mechanisms had the largest cluster (7,045; 13.27%), followed by posttranslational modification, protein turnover, and chaperones (6,176; 11.64%), general function (5,948; 11.2%), translation, ribosomal structure, and biogenesis (3,299; 6.21%), unknown function (3,192; 6.01%), transcription (3,083; 5.81%), and unnamed proteins (3; 0.01%) (Figure 6B) (Table 2) (Supplementary Table S1).

Functional classification by Kyoto Encyclopedia of Genes and Genomes and Reactome pathways

The allohexaploid Brassica CDSs were examined further in the KEGG and Reactome pathway databases. A total of 29,810 (22.38%) significant matches in the database were assigned to 144 KEGG pathways, including the metabolism and genetic information processing translation, out of a total of 133,167 CDSs. In the metabolism category, carbohydrate metabolism (7,481; 25%) was the most common sub-category, followed by amino acid metabolism (4,594; 15%), lipid metabolism (3,461; 12%), xenobiotic biodegradation and metabolism (2,568; 9%), energy metabolism (2,370; 8%), metabolism of cofactors and vitamins (2,302; 8%), biosynthesis of other secondary metabolites (1,563; 5%), and nucleotide metabolism (3,461; 12%). (1,148; 4%). A single aminoacyl-tRNA biosynthesis pathway with 325 CDSs was discovered under the genetic information processing translation category (Figure 7A). *Homo sapiens* (Human) had the most pathways in the Reactome database (153; 9.87%), followed by *Mus musculus* (House mouse) (130; 8.38%), *Rattus norvegicus* (Brown rat) (130; 8.38%), *Gallus gallus* (Fowl) (128; 8.26%), *Danio rerio* (Zebrafish) (127; 8.19%), and *Canis familiaris* (Dog) (126; (4; 0.26%). Due to their close proximity, *M. musculus* and *R. norvegicus* demonstrated a total of 18,096 CDSs belonging to 130 pathways. However, in contrast, 126 pathways were identified to belong to *C. familiaris* and *S. scrofa* with 17,324 and 17,803 CDSs, respectively (Figure 7B) (Supplementary Table S2).

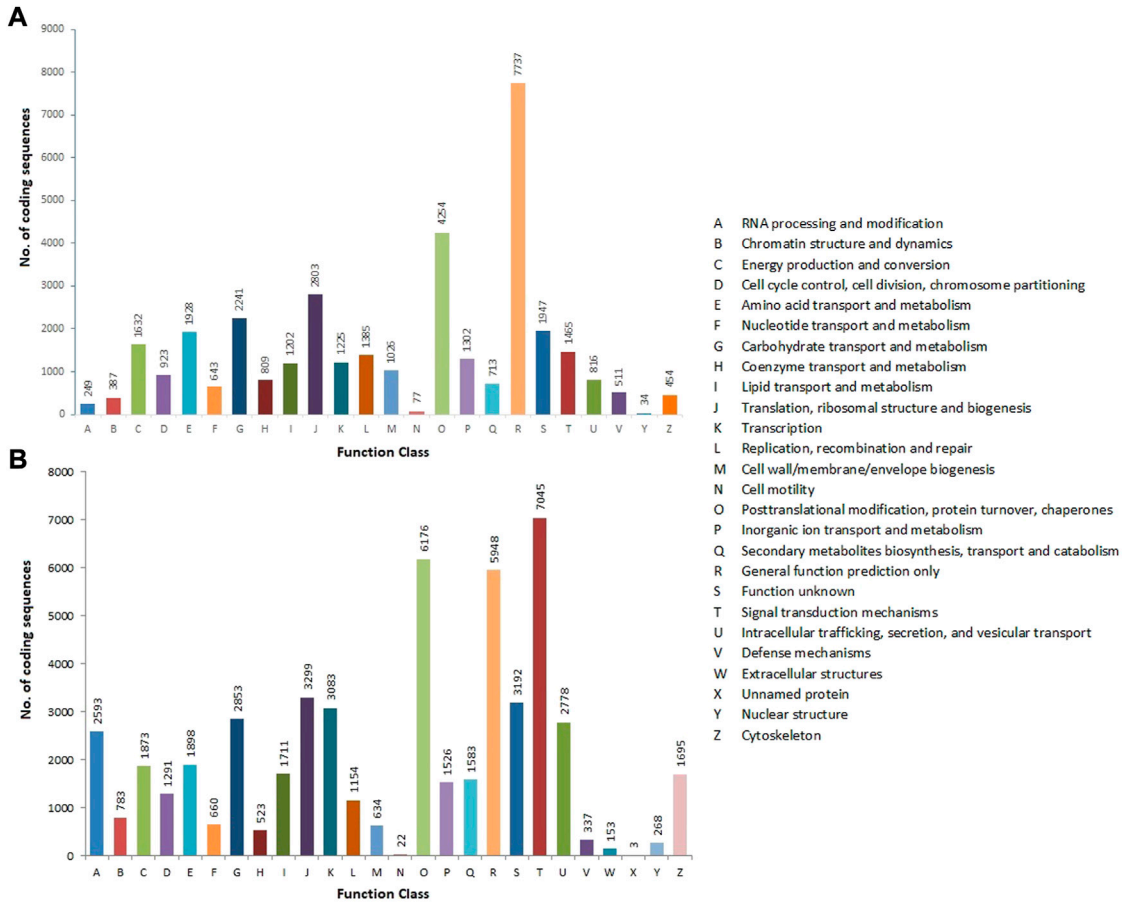


FIGURE 6
COG **(A)** and KOG **(B)** functional classification of allohexaploid Brassica coding sequences. The *y*-axis indicates the number of coding sequences in a specific functional cluster. The *x*-axis indicates the functional classes.

TABLE 2 BLAST hit percentage of annotated coding sequences in different protein databases.

Value	Total	Nr	Swiss-Prot	Refseq	COG	KOG	KEGG	GO	Pfam	Prk	Tigrfam	Average
No.	133,167	87,781	126,131	86,226	35,763	53,081	29,810	77,613	120,489	36,476	39,368	69,273.8
%	100	65.92	94.72	64.75	26.86	39.86	22.39	58.28	90.48	27.39	29.56	52.02

Genes putatively related to heat stress tolerance

The downloaded HSPs are 13,279,051 bp in length, with an average size of 1,019.74 bp. HSPs had the greatest and smallest gene sizes of 11,682 and 30 bp, respectively, with an N50 value of 1,413 bp (n = 3,240). BLASTN results indicated a total of 2,012 CDSs in allohexaploid Brassica, with over 95% similarity with all of the plants' downloaded HSPs. However, the vast majority of CDSs (99.85%) share absolute homology with these HSPs. The 100% similarity between CDSs and HSPs further confirmed the highly conserved character of these proteins across plant taxa. The molecular weights of the downloaded HSPs that showed similarity with CDSs ranged from 14.7 to 70 kDa. The bulk of HSPs, however, belonged to the small HSP category, with MW ranging from 14.7 to 26.5 kDa. These HSPs show up in a variety of cell sites, such as the cytosol, peroxisome, chloroplast, and mitochondria. The sHSPs were categorized into class I–class VI categories based on their involvement in the cell, out of the

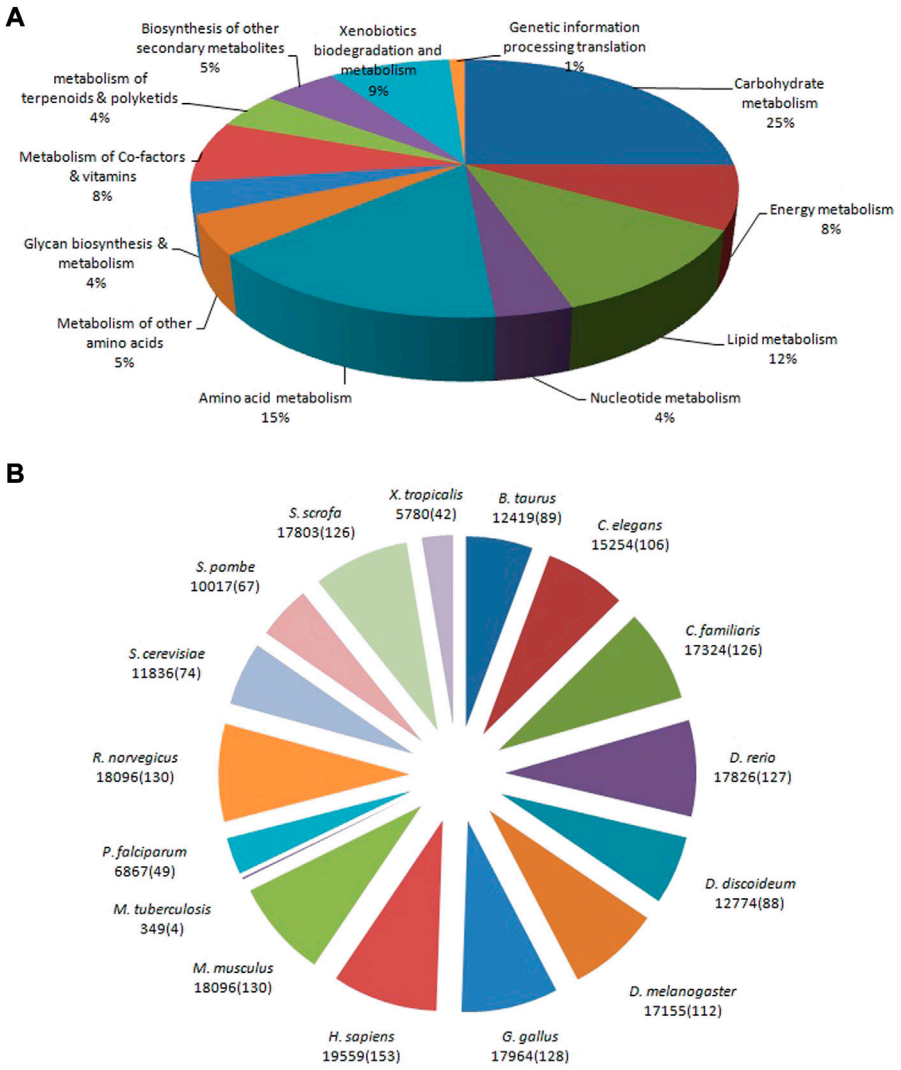


FIGURE 7
Pathway assignment based on the Kyoto Encyclopedia of Genes and Genomes: KEGG **(A)** and Reactome **(B)** databases.

total discovered HSPs of allohexaploid (Supplementary Table S3).

SSRs identification and development

The SSR prediction and primer designing were performed in CDSs of assembly developed using MISA and Primer3 software. MISA has analyzed a total of 133,167 CDSs for SSR prediction, but only 13,595 sequences were found to have SSRs. From 133,167 CDSs, 15,736 potential SSRs were identified, with 1,814 CDSs having multiple microsatellite markers. There were 1,407 compound microsatellite markers among the 15,736 SSRs. Microsatellite markers were detected in 11.82% of allohexaploid Brassica CDSs, with an average of one SSR every 4.29 kb of CDS length. Mononucleotide repeats were detected in the highest number of microsatellites (7,232 or 45.96%), followed by di-, tri-, tetra-, penta-, and hexa-nucleotide repeats in 2,058 (13.08%), 6,272 (39.86%), 71 (0.45%), 48 (0.30%), and 55 SSRs (0.35%), respectively (Figure 8A). The most common mononucleotide motif was T, which was discovered in 4,007 (25.46%) SSRs, followed by A, G, and C, which were found in 3,158 (20.07%), 41 (0.26%), and 26 (0.16%) SSRs, respectively. TC (496; 3.15%), CT (424; 2.69%), AG (362; 2.30%), GA (234; 1.49%), TA (187; 1.19%), AT (155; 0.99%), TG (68; 0.43%), AC (47; 0.30%), CA (43; 0.27%), GT (33; 0.21%), CG (5), and GC (4) were the most common dinucleotide motifs in SSRs. The trinucleotide motifs were found in 60 different combinations,

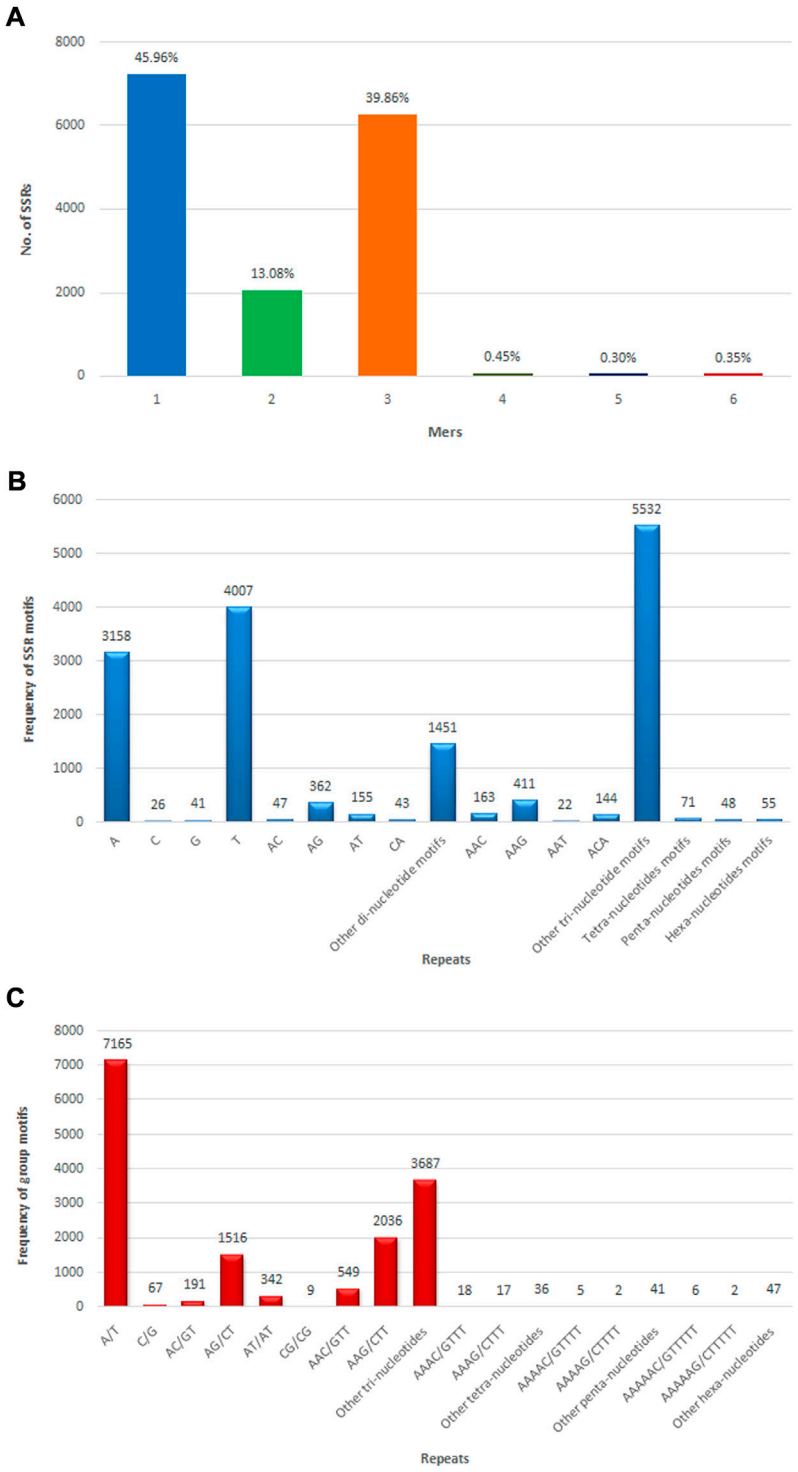


FIGURE 8
Numbers of different types of SSRs identified in allohexaploid Brassica transcriptomes. **(A)** Different repeat motif distribution in total SSRs; **(B)** Top motif distribution in SSRs; **(C)** Abundance of different grouped repeat motifs in total SSRs.

with GAA (627; 3.99%) being the most common, followed by AAG (411; 2.61%), AGA (316; 2.01%), TCT (302; 1.92%), and TTA and TAG (9) being the least common. There were 71, 48, and 55 microsatellites with tetra-, penta-, and hexa-nucleotide motifs, respectively (Figure 8B).

The most abundant grouped motif in the mononucleotide repeats was A/T (7,165; 45.53%), followed by C/G (67; 0.42%). Among the dinucleotide repeats, the most frequent repeat motif was AG/CT (1,516; 9.63%), followed by AT/AT (342; 2.17%), AC/GT (191; 1.21%), and CG/CG (9; 0.06%). Among the trinucleotide repeat motifs, the most abundant repeat motifs was AAG/CTT (2,036; 12.94%), followed by AGG/CTT (1,165; 7.40%), ATC/ATG (988; 6.28%), ACC/GGT (564; 3.58%), AAC/GTT (549; 3.49%), AGC/CTG (464; 2.95%), CCG/CGG (195; 1.24%), ACG/CGT (120; 0.76%), ACT/AGT (99; 0.63%), and AAT/ATT (92; 0.58%). The tetra-, penta-, and hexa-nucleotide grouped motifs were present in 71 (0.45%), 48(0.30%), and 55 (0.35%) microsatellites, respectively (Figure 8C) (Supplementary Table S1). The details included in the table for microsatellites are unique ID, positions of microsatellite flanking regions in the CDSs, sequence, length of forward and reverse primers, melting temperature, and expected PCR product size (Supplementary Table S4).

Discussion

The cultivated members of Brassicaceae existing in nature are either diploid (*B. rapa*, *B. nigra*, *B. oleracea*) or tetraploid/amphidiploids (*B. junea*, *B. napus*, *B. carinata*) (Yang et al., 2016). However, due to a considerable benefit over existing cultivated Brassica, such as combining genetic diversity and traits from all three crops and related genomes, which leads to allelic heterosis from additional genomes, various attempts have been made to generate an allohexaploid Brassica. As a result, the Brassicaceae family has been extensively studied for developing synthetic polyploids (hexaploids) using cultivated and wild members *via* embryo rescue, followed by chromosome doubling *via* colchicine treatment (Meng et al., 1998; Rahman, 2001; Pradhan et al., 2010; Gupta et al., 2016; Mwathi et al., 2017; Mwathi et al., 2020) and protoplast fusion used as an alternative approach for overcoming from pre- and post-fertilization barriers (Primard et al., 1988; Lelivelt et al., 1993; Gaikwad et al., 1996; Wang et al., 2005). However, in previously established hexaploid Brassicas, genome stability was a serious concern; therefore, they were either sterile or aneuploids (Mason et al., 2012; Mwathi et al., 2019). Polyploids enhanced the gene pool of cultivated plant species, allowing for better tolerance to environmental and biological challenges and the production of novel allelic gains (Osborn, 2004; Hu et al., 2019; Yin et al., 2020). When *B. juncea* is stressed by heat or dryness, the seed set is severely impeded (Anand et al., 2010; Youssefi et al., 2011). The novel allohexaploid (H1; *B. juncea* + *S. alba*) is promising for the future due to its high temperature tolerance of up to 40.2°C during seed developmental stages, as well as resistance to two major diseases, including *Alternaria brassicae* blight, which is not present in all cultivated Brassica varieties (Kumari et al., 2018). Apart from them, due to the restoration of regular meiosis and amphidiploid brassicas, the unique allohexaploid (H1) is possible. Because of its crossability with cultivated diploid and allotetraploid Brassicas, it is considered an "elite" breeding material. Polyploid plants are more tolerant to abiotic stresses than their diploid counterparts (Saleh et al., 2008; Chandra and Dubey, 2010; Van et al., 2011; Manzaneda et al., 2012). Stress-related genes are preserved during the polyploidization process in *Arabidopsis thaliana*, according to a genome-wide investigation of gene expression (Casneuf et al., 2006). The number of CDSs discovered from transcripts in this study was larger than the total number of unigenes identified in *B. juncea* (77,750) and *S. alba* (47,972), the parents of the H1 allohexaploid (Bhardwaj et al., 2015; Zhang et al., 2016). Furthermore, the GC percentage of CDSs was greater than the average GC percent of transcriptome assemblies from *B. juncea* (41.92) and *S. alba* (46.63). (Sinha et al., 2015; Zhang et al., 2016). Due to the lack of a reference transcriptome assembly to assess the quality of CDSs, the N50 value was employed instead, which was found to be adequate (Mundry et al., 2012). The *de-novo* transcriptome assembly was annotated against protein databases (Nr, Swiss-Prot, Refseq, COG, KOG, KEGG, GO, Pfam, Prk, and Tigrfam) to provide a comprehensive result for genetic investigation. The protein-coding potential of sequences was confirmed by a maximum of about 94% positive hits of allohexaploid CDSs against the public protein database. Due to incomplete sequences or inadequate information on the *S. alba*, *B. juncea*, and related species in the databases, certain genes cannot be linked with any functional annotation. Unmatched sequences, on the other hand, are nevertheless a valuable resource for allohexaploid use. To determine the evolutionary link, the CDSs of allohexaploid plants revealed a high level of absolute sequence similarity with Brassicaceae plant species, indicating that the assemblies are homologues. Similarly, *Cyamopsis tetragonoloba* (Fabaceae) has the strongest phylogenetic relationship to leguminous plants in a transcriptome analysis (Rawal et al., 2017; Tanwar et al., 2017; Qurainy et al., 2019).

The GO terms are a valuable resource for determining gene function (Ashburner et al., 2000). A total of 77,613 CDSs have been assigned to the 953,773 GO terms that have been divided into three functional classes (biological process, cellular components, and molecular function). Many GO terms (e.g., response to abiotic and biotic stimulus, immune response, stress response, and response to other organisms) have substantial significance to allohexaploid's wide pathogen tolerance and high temperature and present new avenues for future research. The allohexaploid brassica (H1) and *S. alba* (one of their parents) have been reported to be resistant to a variety of fungal diseases, insect pest, drought, and heat tolerance (Kumari

et al., 2018; Kumari et al., 2020a; Kumari et al., 2020c). The genetic mechanism underlying all of these resistance responses, however, is still unknown. We used the RPSBLAST tool to annotate CDSs against the COG and KOG protein databases using the WebMGA web server. KOG identifiers (IDs) are used to identify orthologous and paralogous proteins among eukaryotes that have assigned roles to novel genes (Liu et al., 2016). In the COG and KOG databases, a total of 26.86% and 39.86% of CDSs were classified into 24 and 26 functional classes, respectively. A larger number of COG and KOG classifications revealed that the CDSs in allohexaploid Brassica comprise a diverse set of genes. Many functional classes linked with allohexaploid stress responses were identified in these databases, including defense mechanisms, chaperons, secondary metabolite production, and glucose transport and metabolism (Lopez et al., 2017). The Kyoto Encyclopedia of Genes and Genomes (KEGG) is a curated database for gene functional annotations, with the results displayed in the form of graphics of various biochemical pathways and an additional table of orthologous groups that depicted the evolutionary link between genes (Ogata et al., 1999). In the current study, the majority (98.9%) of total CDSs (29,810) assigned to KEGG were found to be connected to metabolic pathways. Antibiotic, Brassinosteroid, starch and sucrose metabolism, insect hormone, cutin, suberin and wax biosynthesis, glucosinolate, and secondary metabolite biosynthesis were all detected in the allohexaploid Brassica CDSs and were putatively associated with the wide range of adaptations during harsh stress conditions (Beck et al., 2007; Peng et al., 2015; Zielniok et al., 2015; Kumari et al., 2020a). Reactome is an open-source pathway database for curated animal proteins that is freely accessible. NCBI, Ensembl, UniProt, KEGG (Gene and Compound), ChEBI, PubMed, and GO are widely used to cross-reference the content of this database. The Reactome database graphically depicts the enriched pathways (Cai et al., 2020).

HSPs are implicated in disease and insect pest invasion and stress conditions such as heat, alkalinity, drought, low temperature, and UV light (Boston et al., 1996; Bhattarai et al., 2007; Breiman, 2014). Resistant protein stability, immunity regulation, and programmed cell death are all regulated by the HSP70 and sHSP families (Kim et al., 2007; Ooijen et al., 2010; Liu and Whitham, 2013). Protein homeostasis, or maintaining appropriate concentration, conformation, and subcellular location of proteins in the cytosol, peroxisome, chloroplast, and mitochondria, was found to be aided by these sHSP (Hoppe and Cohen, 2020). Restricted tobacco etch potyvirus (TEV) mobility, enhanced DNA methylation, an ATP-independent chaperone that prevents protein aggregation and protects against cell stresses, serves as co-chaperones, and other molecular roles were assigned to these HSPs. HSPs with MW greater than 20.0 kDa were found in chloroplasts or mitochondria, while those with MW less than 20.0 kDa were found in peroxisomes. HSPs with molecular weights of 57 and 70 kDa were found in the cytosol. The presence of HSPs in the cytosol confers resistance to biotic and abiotic stresses (Vierling, 1991; Boston et al., 1996). HSP70 in the cytosol is required for the hypersensitive response (HR) to pathogen attack and non-host resistance (Kanzaki et al., 2003). Gao et al. (2021) provided heat shock treatment at 40 and 60°C to analyze gene expression patterns for the high temperature tolerance in *B. napus*. The DEGs examination study identified a total of 442 genes in seeds treated with high temperature that belongs to posttranslational modifications, protein turnover, chaperons, carbohydrate transport, metabolic pathways, and secondary metabolite biosynthesis pathways. Out of all these DEGs, they have identified only 6 sHSP and 22 transcription factor genes that were involved in heat stress tolerance. Wang et al. (2018b) conducted a study on *Lentinula edodes* by transcriptome and proteome analysis to identify proteins related to thermotolerance. They have identified various types of heat shock proteins, such as HSP40, HSP70, and HSP90, in association with tryptophan and IAA pathways. Yu et al. (2014) identified genes that worked under heat-stress conditions at the time of seed development in *B. napus* through transcriptome profiling. The study of DEGs identified many upregulated genes, including heat transcription factors (13 HSFs), heat shock proteins (91 HSPs; DnaJ/Hsp40, Hsp60/10, Hsp70, Hsp90, Hsp101, and sHsp), and heat-related marker genes such as ROF2, DREEB2a, MBF1c, and Hsa32. Wang et al. (2016) used Chinese cabbage (*Brassica rapa* ssp. chinensis) strains for comparative transcriptional analysis to reveal the heat-responsive genes. Out of 625 DEGs identified in their study, two HSPs, i.e., Bra034104 (HSP70-1) and Bra030036 (HSP) with some genes from the WRKY gene family (Bra015372, WRKY7; Bra017561, WRKY8, and Bra006178, WRKY75) were reported to be expressed under high temperature conditions. Guo et al. (2019) identified a total of 23 heat shock transcription factors and 61 heat shock proteins (HSP100/ClpB, HSP90, HSP70, and sHSPs) that were upregulated upon the heat treatment of four different strains of Chinese kale (*Brassica alboglabra*). Moreover, Lohani et al. (2021) used comparative transcriptome profiling of heat stressed and non-stressed pollen and pistil (stigma and style) to identify DEGs for heat tolerance in *B. napus*. They have identified a variety of proteins involved in heat tolerance, i.e., heat stress transcription factors (HSFs) and heat shock proteins (HSPs)/chaperons. They have identified small heat shock proteins such as 15.7, 17.6, 17.8, 18.5, 21, 22, 23.6, 26.5, and 70 kDA in pollen and pistil, as we recorded in our studies. The presence of a diverse set of HSPs in the CDSs of allohexaploid Brassica (H1) confirmed our prior findings of tolerance to temperatures as high as 40.2°C and resistance to the leaf blight pathogen *A. brassicae* (Kumari et al., 2018; Kumari et al., 2020b).

The transcriptomes are a valuable resource for developing microsatellites from the genome's most conserved regions (Zou

et al., 2013). The most promising technique to study genetic variations and marker-assisted breeding (MAB) is to employ microsatellite markers (Qi et al., 2016). Because their detection is dependent on sequencing technology, sequence completeness, mining software, and input parameters, the frequency and distribution of microsatellite markers in transcribed areas may differ between studies (Aggarwal et al., 2006; Blair and Hurtado, 2013). During the development of microsatellite markers, mononucleotides were discovered to be the most common repeat motif. In SSRs of *S. alba* transcriptomes, trinucleotides were the most prevalent repeat motif, followed by di- and mono-nucleotides, according to Zhang et al. (2016), whereas in *B. juncea* transcriptomic SSRs, dinucleotides (60.3%), and trinucleotides (38.6%) were the most abundant repeat motif types (Dhaka et al., 2017). In another investigation of microsatellite findings, trinucleotides were the most common pattern across 19 Brassicaceae species, followed by di-, tetra-, penta-, and hexa-nucleotides. Similarly, our findings for Brassicaceae and angiosperms met expectations since the most common di- and trinucleotide repeat motifs were AG/CT and AAG/CTT, respectively (Morgante et al., 2002; Victoria et al., 2011; Lopez et al., 2017). We found that the microsatellite frequency level (4.29 kb/SSR) of *S. alba* is greater than the genomic (5.88 kb/SSR) and genic (4.95 kb/SSR) SSRs previously reported (Zhang et al., 2016; Kumari et al., 2020a).

Data availability statement

The data availability statement should be replaced with following link: https://www.ncbi.nlm.nih.gov/sra/PRJNA741791; Accession No. SRR14934389.

Author contributions

KPS: Developed *de-novo* transcriptome assembly and SSRs, performed functional annotations, drafted and finalized manuscript; PK: Designed the research, Developed allohexaploid Brassicas, *in-vivo* cultivation and sample collection, laboratory work, drafted and finalized the manuscript; DKY: Provided laboratory and field facilities for research work. All authors read and approved the final version of manuscript for submission..

Funding

PK received funding from (File No. YSS/2015/001849) the Science and Engineering Research Board, Ministry of Science and Technology, New Delhi, India, for funding the research work (File No. SR/WOS-A/LS-373/2018) and the Department of Science and Technology, Ministry of Science and Technology, New Delhi, Govt. of India, India, for fellowship; KS received Research Associateship [File No. 09/1247(0001)/2019-EMR-I] from HRDG-Council of Scientific and Industrial Research, New Delhi.

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/fgene.2022.958217/full#supplementary-material

References

Aggarwal, R. K., Hendre, P. S., Varshney, R. K., Bhat, P. R., Krishnakumar, V., and Singh, L. (2006). Identification, characterization and utilization of EST-derived genic microsatellite markers for genome analyses of coffee and related species. *Theor. Appl. Genet.* 114, 359–372. doi:10.1007/s00122-006-0440-x

An, H., Qi, X., Gaynor, M. L., Yue, H., Sarah, C. G., Makenzie, E. M., et al. (2019). Transcriptome and organellar sequencing highlights the complex origin and diversification of allotetraploid Brassica napus. *Nat. Commun.* 10, 2878. doi:10.1038/s41467-019-10757-1

Anand, A., Nagarajan, S., Kishore, N., and Verma, A. (2010). Impact of high temperature at pod development stage on yield and quality of Brassica juncea cultivars under controlled conditions. *Indian J. Agric. Sci.* 80 (12), 1043–1047.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: Tool for the unification of biology. The gene ontology consortium. *Nat. Genet.* 25, 25–29. doi:10.1038/75556

Beck, E. H., Fettig, S., Knane, C., Hartig, K., and Bhattarai, T. (2007). Specific and unspecific responses of plants to cold and drought stress. *J. Biosci.* 32, 501–510. doi:10.1007/s12038-007-0049-5

Bhardwaj, A. R., Joshi, G., Kukreja, B., Malik, V., Arora, P., Pandey, R., et al. (2015). Global insights into high temperature and drought stress regulated genes by RNA-Seq in economically important oilseed crop Brassica juncea. *BMC Plant Biol.* 15, 9. doi:10.1186/s12870-014-0405-1

Bhattarai, K. K., Li, Q., Liu, Y., Kumar, D., and Kaloshian, I. (2007). The MI-1-mediated pest resistance requires Hsp90 and Sgt1. *Plant Physiol.* 144, 312–323. doi:10.1104/pp.107.097246

Blair, M. W., and Hurtado, N. (2013). EST-SSR markers from five sequenced cDNA libraries of common bean (Phaseolus vulgaris L.) comparing three bioinformatic algorithms. *Mol. Ecol. Resour.* 13, 688–695. doi:10.1111/1755-0998.12099

Boston, R. S., Viitanen, P. V., and Vierling, E. (1996). Molecular chaperones and protein folding in plants. *Plant Mol. Biol.* 32, 191–222. doi:10.1007/BF00039383

Breiman, A. (2014). Plant Hsp90 and its co-chaperones. *Curr. Protein Pept. Sci.* 15, 232–244. doi:10.2174/1389203715666140331115603

Cai, G., Yang, X., Chen, T., Jin, F., Ding, J., and Wu, Z. (2020). Integrated bioinformatics analysis of potential pathway biomarkers using abnormal proteins in clubfoot. *Peer J.* 8, e8422. doi:10.7717/peerj.8422

Casneuf, T., De Bodt, S., Raes, J., Maere, S., and Van de Peer, Y. (2006). Nonrandom divergence of gene expression following gene and genome duplications in the flowering plant *Arabidopsis thaliana*. *Genome Biol.* 7, R13. doi:10.1186/gb-2006-7-2-r13

Chandra, A., and Dubey, A. (2010). Effect of ploidy levels on the activities of delta(1)-pyrroline-5-carboxylate synthetase, superoxide dismutase and peroxidase in Cenchrus species grown under water stress. *Plant Physiol. biochem.* 48, 27–34. doi:10.1016/j.plaphy.2009.10.003

Chen, S., Nelson, M. N., Chèvre, A. M., Jenczewski, E., Li, Z., Mason, A. S., et al. (2011). Trigenomic bridges for Brassica improvement. *Crit. Rev. Plant Sci.* 30 (6), 524–547. doi:10.1080/07352689.2011.615700

Dhaka, N., Mukhopadhyay, A., Paritosh, K., Gupta, V., Pental, D., and Pradhan, A. K. (2017). Identification of genic SSRs and construction of a SSR-based linkage map in Brassica juncea. *Euphytica* 213 (15). doi:10.1007/s10681-016-1814-z

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, 211–222. doi:10.1093/nar/gkp985

Gaebelein, R., and Mason, A. S. (2018). Allohexaploids in the genus Brassica. *Crit. Rev. Plant Sci.* 37 (5), 422–437. doi:10.1080/07352689.2018.1517143

Gaikwad, K., Kirti, P. B., Sharma, A., Prakash, S., and Chopra, V. L. (1996). Cytogenetical and molecular investigations on somatic hybrids of Sinapis alba and Brassica juncea and their backcross progeny. *Plant Breed.* 115 (6), 480–483. doi:10.1111/j.1439-0523.1996.tb00961.x

Gao, G., Hu, J., Zhang, X., Zhang, F., Li, M., and Wu, X. (2021). Transcriptome analysis reveals genes expression pattern of seed response to heat stress in Brassica napus L. *Oil Crop Sci.* 6 (2), 87–96. doi:10.1016/j.ocsci.2021.04.005

Garg, H., Atri, C., Sandhu, P. S., Kaur, B., Renton, M., Banga, S. K., et al. (2010). High level of resistance to Sclerotinia sclerotiorum in introgression lines derived from hybridization between wild crucifers and the crop Brassica species B. napus and B. juncea. *Field Crops Res.* 117 (1), 51–58. doi:10.1016/j.fcr.2010.01.013

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, R., Wang, X., Han, X., Li, W., Liu, T., Chen, B., et al. (2019). Comparative transcriptome analyses revealed different heat stress responses in high- and low-GS Brassica alboglabra sprouts. *BMC Genomics* 20, 269. doi:10.1186/s12864-019-5652-y

Gupta, M., Atri, C., Agarwal, N., and Banga, S. S. (2016). Development and molecular genetic characterization of a stable Brassica allohexaploid. *Theor. Appl. Genet.* 129, 2085–2100. doi:10.1007/s00122-016-2759-2

Gupta, S. C., Sharma, A., Mishra, M., Mishra, R. K., and Chowdhuri, D. K. (2010). Heat shock proteins in toxicology: How close and how far? *Life Sci.* 86, 377–384. doi:10.1016/j.lfs.2009.12.015

Haft, D. H., Selengut, J. D., and White, O. (2003). The TIGRFAMs database of protein families. *Nucleic Acids Res.* 31 (1), 371–373. doi:10.1093/nar/gkg128

Hoppe, T., and Cohen, E. (2020). Organismal protein homeostasis mechanisms. *Genetics* 215 (4), 889–901. doi:10.1534/genetics.120.301283

Hu, D., Zhang, W., Zhang, Y., Chang, S., Chen, L., Chen, Y., et al. (2019). Reconstituting the genome of a young allopolyploid crop, Brassica napus, with its related species. *Plant Biotechnol. J.* 17 (6), 1106–1118. doi:10.1111/pbi.13041

Hubley, R., Finn, R. D., Clements, J., Eddy, S. R., Jones, T. A., Bao, W., et al. (2016). The Dfam database of repetitive DNA families. *Nucleic Acids Res.* 44, 81–89. doi:10.1093/nar/gkv1272

Huttner, S., and Strasser, R. (2012). Endoplasmic reticulum-associated degradation of glycoproteins in plants. *Front. Plant Sci.* 3, 67. doi:10.3389/fpls.2012.00067

Jian, H., Zhang, A., Ma, J., Wang, T., Yang, B., Shuang, L. S., et al. (2019). Joint QTL mapping and transcriptome sequencing analysis reveal candidate flowering time genes in Brassica napus L. *BMC Genomics* 20, 21. doi:10.1186/s12864-018-5356-8

Jiao, Y., Wickett, N., Ayyampalayam, S., Chanderbali, A. S., Landherr, L., Ralph, P. E., et al. (2011). Ancestral polyploidy in seed plants and angiosperms. *Nature* 473, 97–100. doi:10.1038/nature09916

Kanzaki, H., Saitoh, H., Ito, A., Fujisawa, S., Kamoun, S., Katou, S., et al. (2003). Cytosolic HSP90 and HSP70 are essential components of INF1-mediated hypersensitive response and non-host resistance to Pseudomonas cichorii in Nicotiana benthamiana. *Mol. Plant Pathol.* 4, 383–391. doi:10.1046/j.1364-3703.2003.00186.x

Kim, H. J., Hwang, N. R., and Lee, K. J. (2007). Heat shock responses for understanding diseases of protein denaturation. *Mol. Cells* 23, 123–131.

Kirti, P. B., Mohapatra, T., Khanna, H., Prakash, S., and Chopra, V. L. (1995). Diplotaxis catholica+ Brassica juncea somatic hybrids: Molecular and cytogenetic characterization. *Plant Cell Rep.* 14, 593–597. doi:10.1007/BF00231945

Kotak, S., Larkindale, J., Lee, U., Von Koskull-Doring, P., Vierling, E., and Scharf, K. D. (2007). Complexity of the heat stress response in plants. *Curr. Opin. Plant Biol.* 10, 310–316. doi:10.1016/j.pbi.2007.04.011

Kumari, P., and Bhat, S. R. (2021). Allohexaploid (H1)(IC0628060; INGR19102), an allohexaploid (Brassica juncea + Sinapis alba) Germplasm resistant to Alternaria brassicae, Sclerotinia sclerotiorum, and tolerant to temperature. *Indian J. Plant Genet. Resour.* 34 (1), 160–161.

Kumari, P., Bist, D. S., and Bhat, S. R. (2018). Stable, fertile somatic hybrids between Sinapis alba and Brassica juncea show resistance to Alternaria brassicae and heat stress. *Plant Cell Tissue Organ Cult.* 133, 77–86. doi:10.1007/s11240-017-1362-9

Kumari, P., Singh, K. P., Bisht, D., and Kumar, S. (2020b). Somatic hybrids of sinapis alba + Brassica juncea: Study of backcross progenies for morphological variations, chromosome constitution and reaction to Alternaria brassicae. *Euphytica* 216, 93. doi:10.1007/s10681-020-02629-3

Kumari, P., and Singh, K. P. (2019). Characterization of stable somatic hybrids of Sinapis alba and Brassica juncea for Alternaria blight, Sclerotinia sclerotiorum resistance and heat tolerance. *Indian Res. J. Ext. Educ.* 19 (2), 99–103.

Kumari, P., Singh, K. P., Kumar, S., and Yadava, D. K. (2020c). Development of a yellow-seeded stable allohexaploid brassica through inter-generic somatic hybridization with a high degree of fertility and resistance to Sclerotinia sclerotiorum. *Front. Plant Sci.* 11, 575591. doi:10.3389/fpls.2020.575591

Kumari, P., Singh, K. P., and Rai, P. K. (2020a). Draft genome of multiple resistance donor plant Sinapis alba: An insight into SSRs, annotations and phylogenetics. *PLoS ONE* 15 (4), e0231002. doi:10.1371/journal.pone.0231002

Leitch, A. R., and Leitch, I. J. (2008). Genomic plasticity and the diversity of polyploid plants. *Science* 320 (5875), 481–483. doi:10.1126/science.1153585

Lelivelt, C. L. C., Leunissen, E. H. M., Frederiks, H. J., Helsper, J. P. F. G., and Krens, F. A. (1993). Transfer of resistance to the beet cyst nematode (*Heterodera schachtii* Schm.) from Sinapis alba L. (white mustard) to the Brassica napus L. gene pool by means of sexual and somatic hybridization. *Theor. Appl. Genet.* 85, 688–696. doi:10.1007/bf00225006

Li, A. M., Wei, C. X., Jiang, J. J., Zhang, Y. T., Snowdon, R. J., and Wang, Y. (2009). Phenotypic variation in progenies from somatic hybrids between Brassica napus and Sinapis alba. *Euphytica* 170, 289–296. doi:10.1007/s10681-009-9979-3

Liu, H., Sun, M., Du, D., Pan, H., Cheng, T., Wang, J., et al. (2016). Whole-transcriptome analysis of differentially expressed genes in the ray florets and disc florets of Chrysanthemum morifolium. *BMC Genomics* 17, 398. doi:10.1186/s12864-016-2733-z

Liu, J. Z., and Whitham, S. A. (2013). Over-expression of a soybean nuclear localized type-III DnaJ domain-containing HSP40 reveals its roles in cell death and disease resistance. *Plant J.* 74, 110–121. doi:10.1111/tpj.12108

Lohani, N., Singh, M. B., and Bhalla, P. L. (2021). RNA-seq highlights molecular events associated with impaired pollen-pistil interactions following short-term heat stress in Brassica napus. *Front. Plant Sci.* 11, 622748. doi:10.3389/fpls.2020.622748

Lopez, L., Wolf, E. M., Pires, J. C., Edger, P. P., and Koch, M. A. (2017). Molecular resources from transcriptomes in the Brassicaceae family. *Front. Plant Sci.* 8, 1488. doi:10.3389/fpls.2017.01488

Manzaneda, A. J., Rey, P. J., Bastida, J. M., Weiss-Lehman, C., Raskin, E., and Mitchell-Olda, T. (2012). Environmental aridity is associated with cytotype segregation and polyploidy occurrence in Brachypodium distachyon (Poaceae). *New Phytol.* 193, 797–805. doi:10.1111/j.1469-8137.2011.03988.x

Marchant, B. D., Soltis, D. E., and Soltis, P. S. (2016). Patterns of abiotic niche shifts in allopolyploids relative to their progenitors. *New Phytol.* 212, 708–718. doi:10.1111/nph.14069

Mason, A. S., and Batley, J. (2015). Creating new interspecific hybrid and polyploid crops. *Trends Biotechnol.* 33 (8), 436–441. doi:10.1016/j.tibtech.2015.06.004

Mason, A. S., Nelson, M. N., Takahira, J., Cowling, W. A., Alves, G. M., Chaudhuri, A., et al. (2014). The fate of chromosomes and alleles in an allohexaploid Brassica population. *Genetics* 197 (1), 273–283. doi:10.1534/genetics.113.159574

Mason, A. S., Yan, G., Cowling, W. A., and Nelson, M. N. (2012). A new method for producing allohexaploid Brassica through unreduced gametes. *Euphytica* 186, 277–287. doi:10.1007/s10681-011-0537-4

Meng, J., Shi, S., Gan, L., Li, Z., and Qu, X. (1998). The production of yellow-seeded Brassica napus (AACC) through crossing interspecific hybrids of B. campestris (AA) and B. carinata (BBCC) with B. napus. *Euphytica* 103, 329–333. doi:10.1023/A:1018646223643

Morgante, M., Hanafey, M., and Powell, W. (2002). Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. *Nat. Genet.* 30, 194–200. doi:10.1038/ng822

Mundry, M., Bornberg-Bauer, E., Sammeth, M., and Feulner, P. G. D. (2012). Evaluating characteristics of de novo assembly software on 454 transcriptome data: A simulation approach. *PLoS One* 7 (2), e31410. doi:10.1371/journal.pone.0031410

Mwathi, M. W., Gupta, M., Atri, C., Banga, S. S., Batley, J., and Mason, A. S. (2017). Segregation for fertility and meiotic stability in novel Brassica Allohexaploids. *Theor. Appl. Genet.* 130, 767–776. doi:10.1007/s00122-016-2850-8

Mwathi, M. W., Gupta, M., Martinez, D. Q., Pradhan, A., Batley, J., and Mason, A. S. (2020). Fertile allohexaploid Brassica hybrids obtained from crosses between *B. oleracea* and B. juncea via ovule rescue and colchicine treatment of cuttings. *Plant Cell Tissue Organ Cult.* 140, 301–313. doi:10.1007/s11240-019-01728-x

Mwathi, M. W., Schiessl, S. V., Batley, J., and Mason, A. S. (2019). Doubled-haploid" allohexaploid Brassica lines lose fertility and viability and accumulate genetic variation due to genomic instability. *Chromosoma* 128, 521–532. doi:10.1007/s00412-019-00720-w

Ogata, H., Goto, S., Sato, K., Fujibuchi, W., Bono, H., and Kanehisa, M. (1999). Kegg: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 27 (1), 29–34. doi:10.1093/nar/27.1.29

Ooijen, G. V., Lukasik, E., Van Den Burg, H. A., Vossen, J. H., Cornelissen, B. J., and Takken, F. L. (2010). The small heat shock protein 20 RSI2 interacts with and is required for stability and function of tomato resistance protein I-2. *Plant J.* 63, 563–572. doi:10.1111/j.1365-313X.2010.04260.x

Osborn, T. C. (2004). The contribution of polyploidy to variation in Brassica species. *Physiol. Plant.* 121 (4), 531–536. doi:10.1111/j.1399-3054.2004.00360.x

Park, C. J., and Seo, Y. S. (2015). Heat shock proteins: A review of the molecular chaperones for plant immunity. *Plant Pathol. J.* 31 (4), 323–333. doi:10.5423/PPJ.RW.08.2015.0150

Peng, X., Teng, L., Yan, X., Zhao, M., and Shen, S. (2015). The cold responsive mechanism of the paper mulberry: Decreased photosynthesis capacity and increased starch accumulation. *BMC Genomics* 16, 898. doi:10.1186/s12864-015-2047-6

Pradhan, A., Plummer, J. A., Nelson, M. N., Cowling, W. A., and Yan, G. (2010). Successful induction of trigenomic hexaploid Brassica from a triploid hybrid of B. napus L. and B. nigra (L.) Koch. *Euphytica* 176, 87–98. doi:10.1007/s10681-010-0218-8

Primard, C., Vedel, F., Mathieu, C., Pelletier, G., and Chevre, A. M. (1988). Interspecific somatic hybridization between Brassica napus and Brassica hirta (Sinapis alba L.) *Theor. Appl. Genet.* 75, 546–552. doi:10.1007/bf00289119

Qi, W., Lin, F., Liu, Y., Huang, B., Cheng, J., Zhang, W., et al. (2016). High-throughput development of simple sequence repeat markers for genetic diversity research in Crambe abyssinica. *BMC Plant Biol.* 16, 139. doi:10.1186/s12870-016-0828-y

Qurainy, A. F., Alshameri, A., Gaafar, A. R., Khan, S., Nadeem, M., Alameri, A. A., et al. (2019). Comprehensive stress-based de-novo transcriptome assembly and annotation of guar (Cyamopsis tetragonoloba (L.) Taub.): An important industrial and forage crop. *Int. J. Genomics* 2019, 7295859. doi:10.1155/2019/7295859

Rahman, M. H. (2001). Production of yellow-seeded Brassica napus through inter-specific crosses. *Plant Breed.* 120, 463–472. doi:10.1046/j.1439-0523.2001.00640.x

Rahmati Ishka, M., Brown, E., Weigand, C., Tillett, R. L., Schlauch, K. A., Miller, G., et al. (2018). A comparison of heat-stress transcriptome changes between wild type Arabidopsis pollen and a heat sensitive mutant harboring a knockout of cyclic nucleotide gated cation channel 16 (cngc 16). *BMC Genomics* 19 (1), 549. doi:10.1186/s12864-018-4930-4

Rawal, H., Kumar, S., Sv, A. M., Solanke, A. U., Nigam, D., Saxena, S., et al. (2017). High quality unigenes and microsatellite markers from tissue specific transcriptome and development of a database in cluster bean (Cyamopsis tetragonoloba, L. Taub). *Genes* 8 (11), 313. doi:10.3390/genes8110313

Saleh, B., Allario, T., Dambier, D., Ollitrault, P., and Morillon, R. (2008). Tetraploid citrus rootstocks are more tolerant to salt stress than diploid. *C. R. Biol.* 331, 703–710. doi:10.1016/j.crvi.2008.06.007

Sarkar, N. K., Kim, Y. K., and Grover, A. (2014). Coexpression network analysis associated with call of rice seedlings for encountering heat stress. *Plant Mol. Biol.* 84 (1-2), 125–143. doi:10.1007/s11103-013-0123-3

Shawn, E. L., and Richard, M. M. (2016). Advancements in next-generation sequencing. *Annu. Rev. Genomics Hum. Genet.* 17, 95–115. doi:10.1146/annurev-genom-083115-022413

Shi, J., Yan, B., Lou, X., Ma, H., and Ruan, S. (2017). Comparative transcriptome analysis reveals the transcriptional alterations in heat resistant and heat sensitive sweet maize (Zea mays L.) varieties under heat stress. *BMC Plant Biol.* 17 (1), 26. doi:10.1186/s12870-017-0973-y

Singh, K. P., Kumari, P., and Rai, P. K. (2021). Current status of the disease-resistant gene(s)/QTLs, and strategies for improvement in Brassica juncea. *Front. Plant Sci.* 12, 617405. doi:10.3389/fpls.2021.617405

Sinha, S., Raxwal, V. K., Joshi, B., Jagannath, A., Agarwal, S. K., Goel, S., et al. (2015). De novo transcriptome profiling of cold-stressed siliques during pod filling stages in Indian mustard (Brassica juncea L.) *Front. Plant Sci.* 6, 932. doi:10.3389/fpls.2015.00932

Sitia, R., and Braakman, I. (2003). Quality control in the endoplasmic reticulum protein factory. *Nature* 426, 891–894. doi:10.1038/nature02262

Sjödin, C., and Glimelius, K. (1989). Brassica naponigra, a somatic hybrid resistant to Phoma lingam. *Theor. Appl. Genet.* 77 (5), 651–656. doi:10.1007/BF00261238

Smit, J., Van Eerde, D. H. A. A., and Bakker, A. (2013). A conceptualisation of whole-class scaffolding. *Br. Educ. Res. J.* 39 (5), 817–834. doi:10.1002/berj.3007

Soltis, D. E., Albert, V. A., Leebens-Mack, J., Bell, C. D., Paterson, A. H., Zheng, C., et al. (2009). Polyploidy and angiosperm diversification. *Am. J. Bot.* 96 (1), 336–348. doi:10.3732/ajb.0800079

Stephen, F. A., Warren, G., Webb, M., Eugene, W. M., and David, J. L. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215 (3), 403–410. doi:10.1016/S0022-2836(05)80360-2

Tank, D. C., Eastman, J. M., Pennell, M. W., Soltis, P. S., Soltis, D. E., Hinchliff, C. E., et al. (2015). Nested radiations and the pulse of angiosperm diversification: Increased diversification rates often follow whole genome duplications. *New Phytol.* 207 (2), 454–467. doi:10.1111/nph.13491

Tanwar, U. K., Pruthi, V., and Randhawa, G. S. (2017). RNA-seq of guar (Cyamopsis tetragonoloba, L. Taub.) leaves: De-novo transcriptome assembly, functional annotation and development of genomic resources. *Front. Plant Sci.* 8, 91. doi:10.3389/fpls.2017.00091

Tian, E., Jiang, Y., Chen, L., Zou, J., Liu, F., and Meng, J. (2010). Synthesis of a Brassica trigenomic allohexaploid (B. carinata × B. rapa) de novo and its stability in subsequent generations. *Theor. Appl. Genet.* 121, 1431–1440. doi:10.1007/s00122-010-1399-1

Van, L. K., Franca, S. C., Vansteenkiste, H., Van Huylenbroeck, J., Steppe, K., and Van Labeke, M. C. (2011). Influence of ploidy level on morphology, growth and drought susceptibility in Spathiphyllum wallisii. *Acta Physiol. Plant.* 33, 1149–1156. doi:10.1007/s11738-010-0643-2

Victoria, F. C., da Maia, L. C., and de Oliveira, A. C. (2011). *In-silico* comparative analysis of SSR markers in plants. *BMC Plant Biol.* 11, 15. doi:10.1186/1471-2229-11-15

Vierling, E. (1991). The roles of heat shock proteins in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 42, 579–620. doi:10.1146/annurev.pp.42.060191.003051

Wang, A., Hu, J., Huang, X., Li, X., Zhou, G., and Yan, Z. (2016). Comparative transcriptome analysis reveals heat responsive genes in Chinese cabbage (Brassica rapa ssp. chinensis). *Front. Plant Sci.* 7, 939. doi:10.3389/fpls.2016.00939

Wang, G. Z., Ma, C. J., Luo, Y., Zhou, S. S., Zhou, Y., Ma, X. L., et al. (2018a). Proteome and transcriptome reveal involvement of heat shock proteins and indoleacetic acid metabolism process in Lentinula edodes thermotolerance. *Cell. Physiol. biochem.* 50, 1617–1637. doi:10.1159/000494784

Wang, T., Hou, L., Jian, H., Di, F., Li, J., and Liu, L. (2018b). Combined QTL mapping, physiological and transcriptomic analyses to identify candidate genes involved in Brassica napus seed aging. *Mol. Genet. Genomics* 293 (6), 1421–1435. doi:10.1007/s00438-018-1468-8

Wang, W., Vinocur, B., Shoseyov, O., and Altman, A. (2004). Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends Plant Sci.* 9, 244–252. doi:10.1016/j.tplants.2004.03.006

Wang, Y. P., Sonntag, K., Rudloff, E., and Chen, J. M. (2005). Intergeneric somatic hybridization between Brassica napus L. And sinapis alba L. *J. Integr. Plant Biol.* 47, 84–91. doi:10.1111/j.1744-7909.2005.00009.x

Wei, Z., Wang, M., Chang, S., Wu, C., Liu, P., Meng, J., et al. (2016). Introgressing subgenome components from Brassica rapa and B. carinata to B. juncea for broadening its genetic base and exploring inter-subgenomic heterosis. *Front. Plant Sci.* 7, 1677. doi:10.3389/fpls.2016.01677

Whitley, D., Goldberg, S. P., and Jordan, W. D. (1999). Heat shock proteins: A review of the molecular chaperones. *J. Vasc. Surg.* 29, 748–751. doi:10.1016/s0741-5214(99)70329-0

Wu, S., Zhu, Z., Fu, L., Niu, B., and Li, W. (2011). WebMGA: A customizable web server for fast metagenomic sequence analysis. *BMC Genomics* 12, 444. doi:10.1186/1471-2164-12-444

Yang, J., Liu, D., Wang, X., Ji, C., Cheng, F., Liu, B., et al. (2016). The genome sequence of allopolyploid Brassica juncea and analysis of differential homoeolog gene expression influencing selection. *Nat. Genet.* 48, 1225–1232. doi:10.1038/ng.3657

Yao, S., Liang, F., Gill, R. A., Huang, J., Cheng, X., Liu, Y., et al. (2020). A global survey of the transcriptome of allopolyploid Brassica napus based on single-molecule long-read isoform sequencing and Illumina-based RNA sequencing data. *Plant J.* 103, 843–857. doi:10.1111/tpj.14754

Yin, L., Zhu, Z., Luo, X., Huang, L., Li, Y., Mason, A. S., et al. (2020). Genome-wide duplication of allotetraploid Brassica napus produces novel characteristics and extensive ploidy variation in self-pollinated progeny. *G3* 10 (10), 3687–3699. doi:10.1534/g3.120.401493

Youssefi, A., Nshanian, A., and Aziz, M. (2011). Evaluation of influences of drought stress in terminal growth duration on yield and yield components of different spring Brassica oilseed species. *Am. Euras. J. Agric. Environ. Sci.* 11 (3), 406–410.

Yu, E., Fan, C., Yang, Q., Li, X., Wan, B., Dong, Y., et al. (2014). Identification of heat responsive genes in Brassica napus siliques at the seed-filling stage through transcriptional profiling. *PLoS ONE* 9 (7), e101914. doi:10.1371/journal.pone.0101914

Zhang, X., Liu, T., Duan, M., Song, J., and Li, X. (2016). De novo transcriptome analysis of Sinapis alba in revealing the glucosinolate and phytochelatin pathways. *Front. Plant Sci.* 7, 259. doi:10.3389/fpls.2016.00259

Zhang, Y., Zhang, X., Wang, Y. H., and Shen, S. K. (2017). De-novo assembly of transcriptome and development of novel EST-SSR markers in Rhododendron rex Lévl. through Illumina sequencing. *Front. Plant Sci.* 8, 1664. doi:10.3389/fpls.2017.01664

Zielniok, K., Szkoda, K., Gajewska, M., and Wilczak, J. (2015). Effect of biologically active substances present in water extracts of white mustard and coriander on antioxidant status and lipid peroxidation of mouse C2C12 skeletal muscle cells. *J. Anim. Physiol. Anim. Nutr.* 100, 988–1002. doi:10.1111/jpn.12412

Zou, D., Chen, X., and Zou, D. (2013). Sequencing, de novo assembly, annotation and SSR and SNP detection of sabaigrass (Eulaliopsis binata) transcriptome. *Genomics* 102, 57–62. doi:10.1016/j.ygeno.2013.02.014

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 19 October 2022
DOI 10.3389/fgene.2022.955295

Check for updates

OPEN ACCESS

EDITED BY
Mohan Lal,
North East Institute of Science and Technology (CSIR), India

REVIEWED BY
Hantao Wang,
Cotton Research Institute (CAAS), China
Sindhu Sareen,
Indian Institute of Wheat and Barley Research (ICAR), India

*CORRESPONDENCE
Frank Ordon,
frank.ordon@julius-kuehn.de

SPECIALTY SECTION
This article was submitted to Plant Genomics, a section of the journal Frontiers in Genetics

RECEIVED 28 May 2022
ACCEPTED 28 July 2022
PUBLISHED 19 October 2022

CITATION
Deblieck M, Szilagyi G, Andrii F, Saranga Y, Lauterberg M, Neumann K, Krugman T, Perovic D, Pillen K and Ordon F (2022), Dissection of a grain yield QTL from wild emmer wheat reveals sub-intervals associated with culm length and kernel number. *Front. Genet.* 13:955295. doi: 10.3389/fgene.2022.955295

COPYRIGHT
© 2022 Deblieck, Szilagyi, Andrii, Saranga, Lauterberg, Neumann, Krugman, Perovic, Pillen and Ordon. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Dissection of a grain yield QTL from wild emmer wheat reveals sub-intervals associated with culm length and kernel number

Mathieu Deblieck[1], Gergely Szilagyi[2], Fatiukha Andrii[3], Yehoshua Saranga[2], Madita Lauterberg[4], Kerstin Neumann[4], Tamar Krugman[3], Dragan Perovic[1], Klaus Pillen[5] and Frank Ordon[1]*

[1]Julius Kühn Institute (JKI), Federal Research Centre for Cultivated Plants, Institute for Resistance Research and Stress Tolerance, Quedlinburg, Germany, [2]The Hebrew University of Jerusalem, Robert H. Smith Faculty of Agriculture, Food and Environment, Rehovot, Israel, [3]University of Haifa, Institute of Evolution, Haifa, Israel, [4]Research Group Genebank Documentation, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Seeland, Germany, [5]Department of Plant Breeding, Martin-Luther-University Halle-Wittenberg, Institute of Agricultural and Nutritional Sciences, Halle, Germany

Genetic diversity in wheat has been depleted due to domestication and modern breeding. Wild relatives are a valuable source for improving drought tolerance in domesticated wheat. A QTL region on chromosome 2BS of wild emmer wheat (*Triticum turgidum* ssp. *dicoccoides*), conferring high grain yield under well-watered and water-limited conditions, was transferred to the elite durum wheat cultivar Uzan (*T. turgidum* ssp. *durum*) by a marker-assisted backcross breeding approach. The 2B introgression line turned out to be higher yielding but also exhibited negative traits that likely result from trans-, cis-, or linkage drag effects from the wild emmer parent. In this study, the respective 2BS QTL was subjected to fine-mapping, and a set of 17 homozygote recombinants were phenotyped at BC_4F_5 generation under water-limited and well-watered conditions at an experimental farm in Israel and at a high-throughput phenotyping platform (LemnaTec-129) in Germany. In general, both experimental setups allowed the identification of sub-QTL intervals related to culm length, kernel number, thousand kernel weight, and harvest index. Sub-QTLs for kernel number and harvest index were detected specifically under either drought stress or well-watered conditions, while QTLs for culm length and thousand-kernel weight were detected in both conditions. Although no direct QTL for grain yield was identified, plants with the sub-QTL for kernel number showed a higher grain yield than the recurrent durum cultivar Uzan under well-watered and

Abbreviations: CL, culm length; CKN, calculated kernel number; DPH, days from planting to heading; GY, grain yield; HI, harvest index; MSpL, Main Spike Length; MSpTKW, Main Spike Thousand Kernel Weight; Mbp, Million base Pairs; MSpSp, Main Spike Spikelets; MSpSpSp, Main Spike Seeds per Spikelet; NIL, near-isogenic line; PCA, principal component analysis; QTL, quantitative trait loci; SNP, single nucleotide polymorphism; Spp, spikes per plan; TKW, thousand kernel weight; TotDM, total dry matter; WL, water-limited; WW, well-watered.

mild drought stress conditions. We, therefore, suggest that this sub-QTL might be of interest for future breeding purposes.

KEYWORDS

wild emmer, grain yield, drought tolerance, culm length, kernel number, GBS, 15K-iSelect

Introduction

The meteorological concept of drought can be defined as "an extended period of time characterized by a deficiency in a region's water supply" (Below et al., 2007). Stress might be defined as an altered physiological condition that alters an equilibrium, leading to a strain, i.e., a biochemical and/or physical change, which can lead to injury, disease, or aberrant physiology of the plant (Gaspar et al., 2002). Plants can experience drought stress either when water supply becomes limited or when the transpiration rate becomes too high (Fahad et al., 2017). Drought stress effects on plants are, e.g., the reduction of water content, diminished leaf water potential and turgor loss, closure of stomata, decrease in cell enlargement/growth due to low turgor pressure, and arrest of photosynthesis, which can finally result in the disturbance of metabolism and death. Increased periods of drought were shown to have major negative effects on the yield (Lesk et al., 2016; Fahad et al., 2017). Potential yield losses differ according to the magnitude of stress and the developmental stage of the plant. Extreme yield losses were reported to be up to 92% (Farooq et al., 2014). Plant responses to drought stress can be divided into long- and short-term responses on the biochemical, molecular, and finally physiological levels. These responses lead to different strategies to deal with drought stress, i.e., escape, tolerance, avoidance, and recovery (Chen et al., 2015; Fang and Xiong, 2015).

Wheat (*Triticum* spp.) is among the top five grown crops in the world (Oyewole 2016; FAO 2019) and provides about 20% of the total calories consumed by humans (Poole et al., 2021). Future plant production faces the challenge of feeding the growing global population, which is predicted to reach about 9.8 billion people by 2050 (UN 2017). This challenge is increased by climate change, because per degree-Celsius increase in the global mean temperature wheat yield will be decreased by 6.0% (Fahad et al., 2017; Zhao et al., 2017). Developing high-yielding cultivars under drought with yield stability between environments is, thus, of prime importance (Langridge and Fleury, 2011).

The genetic diversity of tetraploid and hexaploid wheat has been depleted by the limited number of allopolyploidization events, domestication, and modern breeding (Dempewolf et al., 2017). Using crop wild relatives (CWR) such as wild emmer wheat in pre-breeding turned out to be an efficient tool to exploit their genetic diversity (Xie and Nevo, 2008; Huang et al., 2016; Dempewolf et al., 2017). A significant increase in the use of CWR has been noted since 1980 (Dempewolf et al., 2017). However, since abiotic stress tolerance is complex, the introgression of quantitative trait loci (QTL) from CWR aiming to confer tolerance to abiotic stress into elite cultivars is difficult (Xie and Nevo, 2008; Dempewolf et al., 2017). Traits that are advantageous for the CWR's fitness can be detrimental for breeding purposes. An increased culm length, for instance, might allow the plant to compete for light (Lane et al., 2000) or to promote the distribution of the plants' own seeds in the ecosystem—even if this leads to a trade-off in grain yield (GY). Avoiding trans-, cis-, or linkage drag effects from the CWR is, therefore, one of the most difficult issues in pre-breeding.

In 2017, the near-isogenic line NIL-U-2B-1 carrying a wild emmer QTL region on chromosome 2BS was shown to produce more GY than its recurrent elite durum parent, under water-limited and well-watered conditions (Merchuk-Ovnat et al., 2016; Merchuk-Ovnat et al., 2017). It also turned out that negative traits such as an increased culm length or delayed heading were introduced, possibly due to trans-, cis-, or linkage drag effects from the CWR (Merchuck-Ovnat et al., 2016). A recent analysis of NIL-U-2B-1 revealed that more than two-thirds of chromosome 2B and additional fragments of chromosomes 2A, 3A, and 5A were also introgressed from the donor wild emmer parent (Deblieck et al., 2020). Consequently, the original RIL mapping population, previously used for mapping the QTL on chromosome 2B, was re-genotyped with the 15k iSelect chip (TraitGenetics) to narrow down the previously identified QTL regions by a higher marker density of ~4,000 SNPs (Fatiukha et al., 2021), which is much higher than the density used for the original mapping (Peleg et al., 2008; Peleg et al., 2009). The existence of the corresponding 20 cM large QTL-region for GY on chromosome 2BS was again confirmed under both water-limited and well-watered conditions (Fatiukha et al., 2021). Therefore, this study aimed at the fine mapping of a 15.67 cM region of this 20 cM large QTL, to identify the smallest sub-QTL region affecting grain yield, and to diminish the observed trans-, cis-, or linkage drag effects from the CWR. For this purpose, we used the iSelect marker Tdurum_contig27976_414 and wsnp_Ex_c6537_1133876, flanking the corresponding QTL-interval (Fatiukha et al., 2021) and 10 additional markers within the QTL, to establish a population of sub-NILs, carrying segmental chromosomal substitutions of the target region.

Materials and methods

Plant material

All plant materials were developed at the Hebrew University of Jerusalem (HUJI) in Rehovot (31°54′N, 34°47′E, 54 m above

sea level). In past works, recombinant inbred lines (RILs), derived from a cross between durum wheat (cv. Langdon) and wild emmer wheat (acc# G18-16) were phenotyped under contrasting water availabilities and used for QTL mapping (Peleg et al., 2009; Fatiukha et al., 2021). Subsequently, the wild emmer QTL allele on chromosome 2BS conferring higher grain yield and harvest index (HI) was introgressed into the elite tetraploid durum wheat cultivar Uzan to develop the near-BC_3F_3 isogenic line (NIL) NIL-U-2B-1 as previously described (Merchuk-Ovnat et al., 2016; Deblieck et al., 2020). In this study, BC_3F_3 NIL-U-2B-1 (pollinator) was crossed with Uzan (Supplementary Figure S1) and BC_4F_3 plants were screened with the respective flanking molecular markers to identify heterozygous recombinants. The heterozygous recombinant BC_4F_3 plants were selfed and BC_4F_4 single seed descendants were again genotyped to identify homozygous recombinant plants for seed multiplication. Finally, BC_4F_5 seeds of 17 homozygous recombinant plants were used for phenotypic experiments in 2017, 2018, and 2019.

Development of molecular markers and genotyping by sequencing (GBS)

The 15k iSelect data of G18-16, Langdon, Uzan, and NIL-U-2B-1 were previously published (Merchuk-Ovnat et al., 2016; Soleimani et al., 2020). Tdurum_contig27976_414 (32.88 cM) and wsnp_Ex_c6537_1133876/IAAV980 (48.55 cM) which flank the QTL-interval on chromosome 2B (Fatiukha et al., 2021) were converted into competitive allele-specific PCR (KASP) markers (Wilkinson et al., 2016). In addition to this, a total of 10 other PCR-based molecular markers, including KASP, cleaved amplified polymorphic sides (CAPS), and simple sequence repeats (SSRs) were developed along the QTL-interval at a distance of ~1 cM to genotype and identify heterozygous recombinant BC_4F_3 sub-NILs. All molecular markers were tested and validated on the DNA of F_7 lines of the original mapping population (Peleg et al., 2008; Peleg et al., 2009; Fatiukha et al., 2021). Furthermore, genotyping by sequencing (GBS) was applied to NIL-U-2B-1, the parental lines, the wild emmer wheat acc. G18-16, the durum elite parent Uzan, and homozygous sub-NILs with recombination close to the marker Gene-1741_103, which was previously described to be the nearest marker to the target QTL interval (Fatiukha et al., 2021). This additional step served to identify new recombination events within the recombinant sub-NILs to further improve fine-mapping efficiency.

GBS libraries were prepared as described previously (Elshire et al., 2011) and sequenced (150 bp paired-end, Illumina MiSeq). The sequencing produced millions of reads. These were de-multiplexed according to the barcodes and the adapters/barcodes using the CASAVA pipeline 1.8 (Illumina, Inc.). Trim Galore software from Babraham Bioinformatics (2012) was used for adapter and quality trimming of the amplified genomic sequences. After this first filtering, the trimmed sequences were aligned to the draft genome sequence of the wild emmer acc. Zavitan (v1) (Avni et al., 2017) using BWM-MEM (version: 0.7.7.-r1140) (Li, 2013), and variant-calling was performed with the samtool and bcftools (version: 0.1.19–96) (Li et al., 2009). High-quality bi-allelic SNPs were filtered, and the imputation of missing SNPs was conducted with Beagle (Browning and Browing, 2016). Aligned sequencing reads were used for SNP detection after a quality check (Q score >20). Multi-allelic SNPs, SNPs with minor allele frequency (MAF) < 5%, missing values ≥5%, or heterozygosity ≥90% were further excluded and a high-quality SNP genotyping dataset was compiled. SNPs which could clearly be assigned to a unique position on the physical genome of wild emmer were kept for analysis. GBS data were finally evaluated with the GenoTypeMapper (GTM) (Deblieck et al., 2020).

Plant growth conditions and experimental design

A total of five plants per genotype were phenotyped under contrasting water-limited (WL) and well-watered (WW) conditions from 2017–2019.

In 2017 and 2018, plants were grown at the HUJI experimental farm at Rehovot. Seedlings were first placed in moist germination paper at 4C° in a dark vernalization room for 2 weeks and then transplanted at the beginning of December into an insect-proof screen house rain-protected by a polyethylene top. The soil was a brown-red sandy loam (Rhodoxeralf) composed of 76% sand, 16% clay, and 8% silt (Merchuk-Ovnat et al., 2016). The different water regimes were simulated in a factorial (genotype x irrigation regime) split plot block design. Each block consisted of two main plots (for the two irrigation regimes), split into subplots for genotypes. Each subplot consisted of a single row with five plants, 10 cm apart (50-cm long plots). In each bed, two 40 cm spaced rows were planted with 100 cm between each pair of rows (Supplementary Figure S2). Seasonal rainfall was simulated by applying water once or twice a week from planting in December to heading in April/May, leading to a total seasonal water application of 350 mm (for WL conditions) and 650 mm (for WW conditions) in 2017. However, drought stress appeared to be mild in 2017. To increase drought stress, 384 mm and 201 mm of water were applied to the segmental sub-NILs in 2018. Plants that grew poorly in a certain area of the screen house were excluded from the experiment, keeping three replicates per genotype.

In July 2019, one plant per genotype was grown per pot on the high throughput phenotyping (HTP) facility (LemnaTec-129 Scanalyzer 3D) (http://www.lemnatec.com) of the Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) in

TABLE 1 Phenotypic traits.

Trait[a]	Description
Calculated kernel number (CKN)	(GY/TKW)x1000
Culm length (CL)	Height from soil to the base of the spike in centimeter (cm)
Days planting to heading (DPH)	Days from planting to heading of 50% of the plants per plot in 2017
Grain yield (GY)	Grain yield in gram (g)
Harvest index (HI)	Ratio between grain yield and total dry weight
Thousand kernel weight (TKW)	TKW of all spikes (g)—including the main spike
Total biomass (TBM)	Total biomass per plant (g)
Main spike length (MSpL)	Main ear length without awns (cm)
Main spike thousand kernel weight (MSpTKW)	TKW of the main spikes (g)
Main spike spikelets (MSpSp)	Number of spikelets of the main ear
Main spike seeds per spikelet (MSpSpSp)	Number of seeds per spikelet of the main spike
Spikes per plant	Number of spikes per plant

[a]Please note that for each genotype, five plants were analyzed under drought stress and controlled conditions. Each of the trait was collected per replicate/plant. Subsequently, the means of the traits were calculated per genotype and conditions. For more details, please read the material and methods section of the article.

Gatersleben in Germany (Supplementary Figure S3). To avoid the confounding effect of flowering time on drought evaluation, sub-NILs with a much later flowering time were not included in the HTP experiment. The controlled greenhouse is described by Neumann et al. (2015) along with the potting procedure. In short, pots (2 L) were filled in a standardized way by weighing in the same amount of soil as a standard substrate ("Klasmann Substrate no. 2" (http://www.klasmann-deilmann.com)) allowing a slow development of drought as described by Neumann et al. (2015). After sowing, 7 g of fertilizer was added to each pot (19% N, 9% P_2O_5, and 10% K_2O) to supply the plants with nutrients throughout the life cycle. After 10 days, the plants were thinned out to one plant per pot. Directly after the daily image recording, each pot was weighed individually and watered to a previously defined target weight according to Dhanagond et al. (2019). Plants were grown on the platform from sowing until maturity and watered daily to the corresponding target weights.

The aim of the experiment was to mimic as good as possible the average temperatures and slowly progressing drought conditions of the screen house experiment conducted in Israel. Drought thresholds for severe drought stress were defined as 20% plant-available water (PAW) based on the results in barley (Dhanagond et al., 2019), while mild drought was defined as 30% PAW. No drought stress was applied until 30 days after sowing (30 DAS). The temperature during this pre-drought phase was set to 12°C at night and 16°C during the day. Supplementary greenhouse lights provided light for 15 h per day. To simulate drought stress, irrigation of plants intended for stress treatment was changed 31 DAS from 90% PAW to 30% PAW (mild drought threshold) and the temperature was raised to 20°C during the day and 16°C at night. At 62 DAS, the temperature was further increased to 24°C during the day and 20°C at night. From 64 DAS, watering was reduced to 20% PAW (severe drought level). This drought level and temperature regime lasted until maturity. At maturity, the plants were subjected to a detailed assessment of agronomic traits. Furthermore, the raw images were inspected, and the date of heading (BBCH 55) was determined for each plant.

Phenotypic data, statistical analysis, and QTL-detection

Table 1 summarizes the type of traits that were recorded at maturity for each single plant (replicate) at the HUJI experimental farm and HTP platform (LemnaTec-129 Scanalyzer 3D) in Gatersleben. All statistical analyses were performed using R (version 3.4.1). The data of each trait, replication, year, and treatment were inspected separately. First, extreme outliers were filtered out, if they fell outside of an interval of plus or minus three times the standard deviation of the mean. Then, quantile–quantile (QQ) plots and density plots were used to evaluate whether the data were normally distributed. Data points which were not normally distributed at the QQ-plots residuals were removed manually. Subsequently, arithmetic means of the traits were calculated for each year/treatment and Shapiro–Wilk tests (Shapiro and Wilk, 1965) were carried out to confirm that data were normally distributed and suitable for further statistical analyses and tests. T-tests and ANOVA were applied to prove for each trait and/or year whether and to what extent they differ under both conditions. Descriptive statistics, i.e., density and correlation plots were calculated to analyze how the traits changed and correlate under the respective irrigation regimes. Finally, mean values of each trait, year, and treatment were used separately to calculate QTLs for each irrigation regime with the software MultiQTL2.6 (http://www.multiqtl.com).

TABLE 2 Genotypic data of the recombinant inbred lines.

	Marker***	*Tdurum_contig27976_414*	*Kukri_rep_c69803_82*	*BS00010055_51*	*Kukri_c62277_80*	GBS_96013158	GBS_97013837	*Gene-1741_103*	GBS_100504657	GBS_10264774	*Tdurum_10414990*8	*Tdurum_contig68806_677*	GBS_113979606	GBS_115312227	*Rac875_c21378_474*	*Kukri_c46621_143*	*wsnp_Ex_c6537_11338763*
	bp**	NA	77156243	80465477	96011570	96013158	97013837	97707169	100504657	102647748	10414986	111425302	113979606	115312227	122506008	128064567	138099492
	cM***	32.88	35.44	36.48	39.36			40.75			42.15	43.95			46.81	47.86	48.55
	1663G																
	1688C																
	1488A																
	1767E																
	1174B																
	1735E***																
	1121B***																
	1004B***																
	1539A***																
	1695C																
	1431G																
	1761D***																
	1145C***																
	1336E																
	1403D																
	1353C																
	1557G																

*15k iSelect markers that were used to develop PCR-based molecular markers are written in italics. Genotypic information that was obtained via GBS, is marked with "GBS" **. Physical positions were obtained from the wild emmer acc. Zavitan (v1) (Avni et al., 2017). Please note that no physical position was obtained for the marker Tdurum_contig27976_414. *** Genetic positions were obtained from the original mapping population (Fatiukha et al., 2021). **** Plants that showed recombination events between Kukri_c6227780 and Rac875_c2138_474 were subjected to GBS (please read material and methods section for more detailed information).

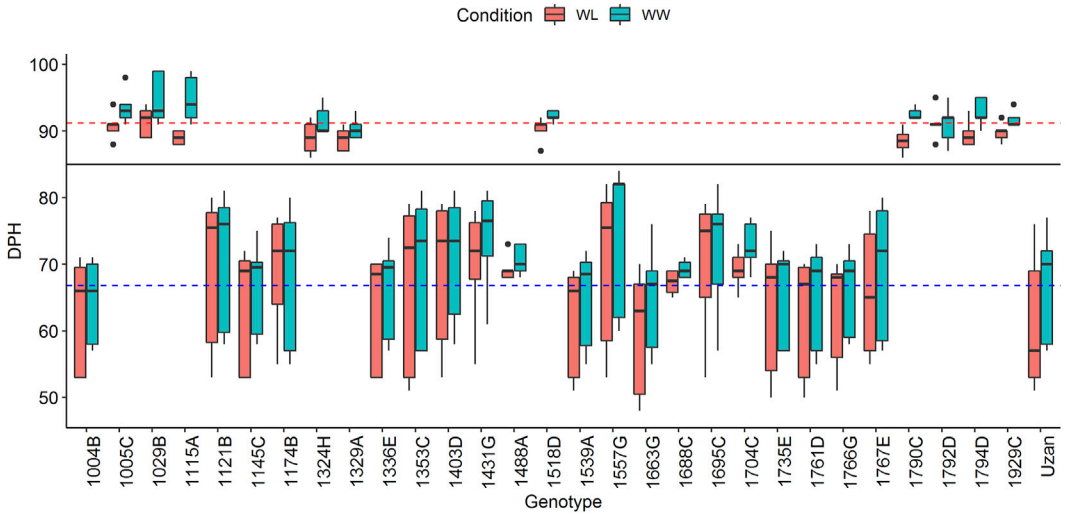


FIGURE 1
Days from planting to heading (DPH) of the complete set of 29 segmental sub-NILs in 2017. The black line indicates that plants cluster into two groups according to their DPH (for more details, see text). Blue and red dashed lines mark the mean DPH time of 66.79 and 91.22 days, respectively. Plants above the black line show delayed heading toward the elite parent Uzan and were, therefore, not considered for the phenotypic experiments in 2018 and 2019.

Results

Marker development

A total of 10 PCR-based molecular markers were developed at a distance of 1–2 centimorgan (cM) along the QTL-interval of 15.67 cM (Supplementary Tables S1-3). Within the interval, six additional polymorphic regions were identified by applying GBS to a subset of plants that showed recombinations close to the marker Gene-1741_103, thus, in between Kukri_c6227780 and Rac875_c2138_474 (Table 2). All markers share the same order in the genetic map and sequenced genome of the wild emmer acc. Zavitan (Table 2, Supplementary Table S1, Deblieck et al., 2020). For Tdurum_contig27976_414, no physical position could be annotated.

Phenotypic data of homozygous recombinants sub-NILs

A total of 600 BC_4F_3 plants were derived from a cross of Uzan with NIL-U-2B-1 (Merchuk-Ovnat et al., 2016). Eighty-two heterozygous recombinant BC_4F_3 plants were identified after genotyping with the flanking Tdurum_contig27976_414 (32.88 cM) and wsnp_Ex_c6537_1133876/IAAV980 (48.55) markers (Supplementary Table S1). Subsequently, a sample of six to eight descendent plants of each heterozygote recombinant BC_4F_3 plant was screened and a total of 96 BC_4F_4 homozygote recombinants were obtained. After further genotyping the BC_4F_4 plants with the 10 markers (Supplementary Table S1, Supplementary Table S2, Supplementary Table S3), 29 of them with representative recombinations along the QTL interval were selected and used for seed multiplication. Finally, 10 plants per genotype, i.e., 290 BC_4F_5 plants, were subjected to phenotyping in 2017. Two very distinct groups of genotypes with different numbers of days from planting to heading (DPH) were observed (Figure 1). One group comprised 17 plants and required a mean time of 66.8 days for heading, whereas the other group comprised 12 plants and flowered 24.4 days later, thus, at 91.2 days. Grouping of these data was additionally confirmed with a Tukey's test (Tukey, 1949) and independent segregation of DPH was confirmed with a chi-square test (Pearson, 1900). A comparison of GBS data of these plants revealed that sub-NILs with delayed DPH-values, i.e., 1029B, 1115A, 1329A, and 1929C, share a single 669 kbp fragment on chromosome 2A ranging from 39980886 to 40649713 BP, which is absent in early flowering sub-NILs, e.g., 1004B, 1539A, 1121B, 1663G, 1766G, 1761D, 1735F, 1336E, 1145C, and Uzan (Supplementary Table S4). As indicated previously, this region harbors the photoperiod sensitive gene *PpdA-1* from the wild emmer parent (Takenaka and Kawahara, 2012), ranging from 40487317–40489398 bp (Deblieck et al., 2020). Remarkably, this huge effect was not observed in the original mapping population (Fatiukha et al., 2021) although Langdon and Uzan share the same GBS-marker alleles along the PpdA-1 locus (Supplementary Table S4).

DPH correlates with grain yield (GY) and many important yield-related traits and might increase/decrease the respective

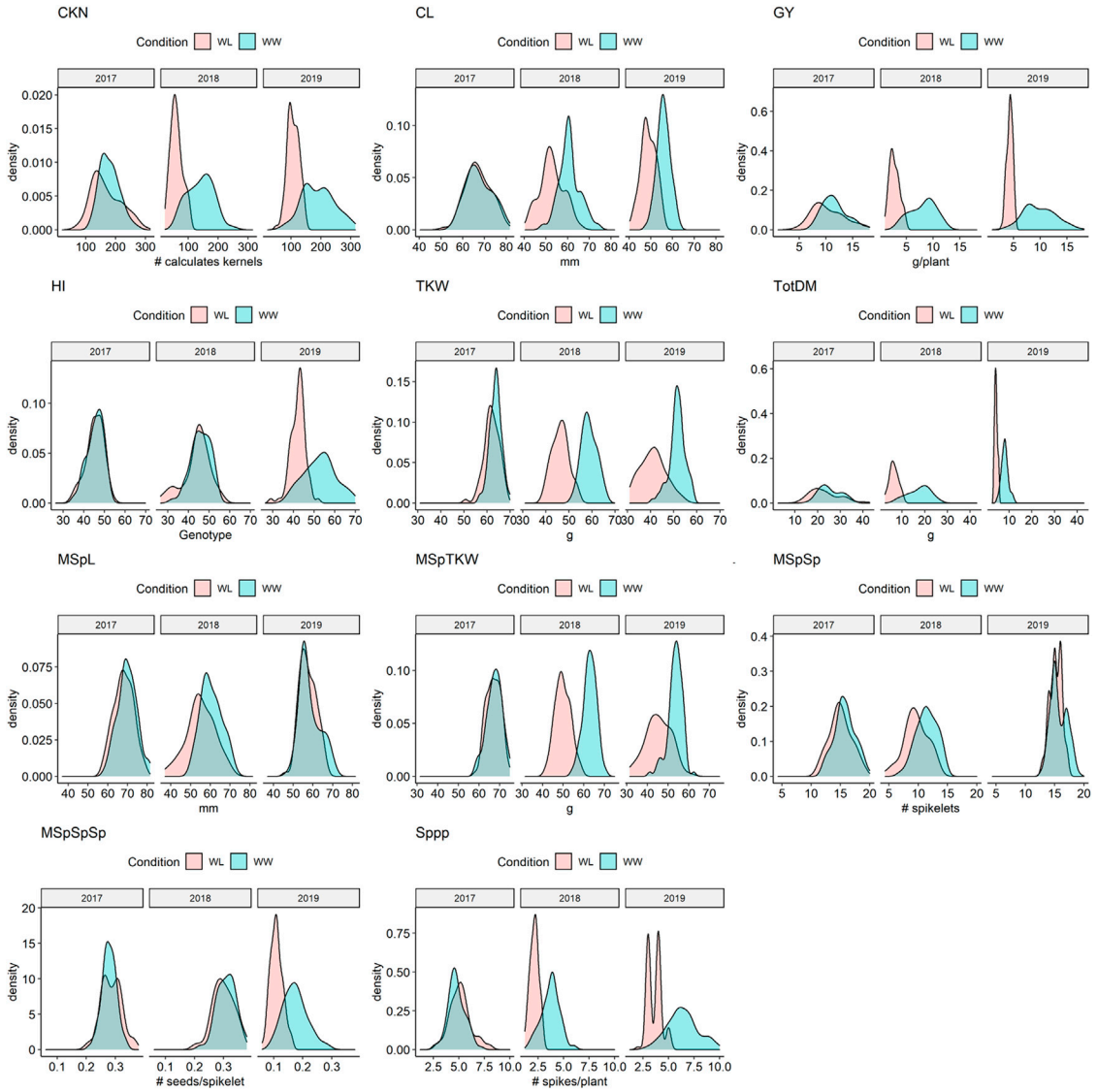


FIGURE 2
Density plots of selected traits under water-limited (WL) and well-watered conditions (WW) in 2017, 2018, and 2019. CKN = Calculated kernel number, CL = Culm Length, GY = Grain yield, HI = Harvest index, TKW = Thousand kernel weight, TotDM = Total Dry Matter, MSpL = Main Spike Length, MSpTKW = Main Spike thousand kernel weight, MSpSp = Main Spike Spikelets, MSpSpSp = Seeds per spikelet (of the main spike), Sppp = Spikelets per plant.

QTL-LOD scores (Fatiukha et al., 2021). We, therefore, decided to exclude plants with increased DPH values from the phenotypic experiments in 2018 and 2019 (Figure 1). In addition to this, two sub-NILs did not deliver enough phenotypic data in 2018. For these genotypes, less than three replicates per genotype/treatment and year were available, which was considered unreliable and not representative, leading to a final set of 17 segmental sub-NILs. For each of these sub-NIL, 10 plants were subjected to phenotyping, i.e., phenotypic data were resolved for 170 plants in 2017, 2018, and 2019 (Table 2).

Almost all traits showed a normal distribution in both environments and all 3 years (Supplementary Table S5, Figure 2 and Supplementary Figure S4). However, in a few cases such as the main spike length (MSpL), the number of spikes per plant (Sppp), and grain yield per spike (GYpSp), phenotypic data were not normally distributed, and therefore, (log-) transformed (Hackett, 1997). These data are clearly highlighted, e.g., in Supplementary Table S5. ANOVA revealed that the means of all traits, except for the number of main spike spikelets (MSpSp), were different between the 3 years (Supplementary Table S6).

While in 2017 the drought stress effect was mild but significant for most of the traits except for culm length (CL), thousand kernel weight of the main spike (MSpTKW), and

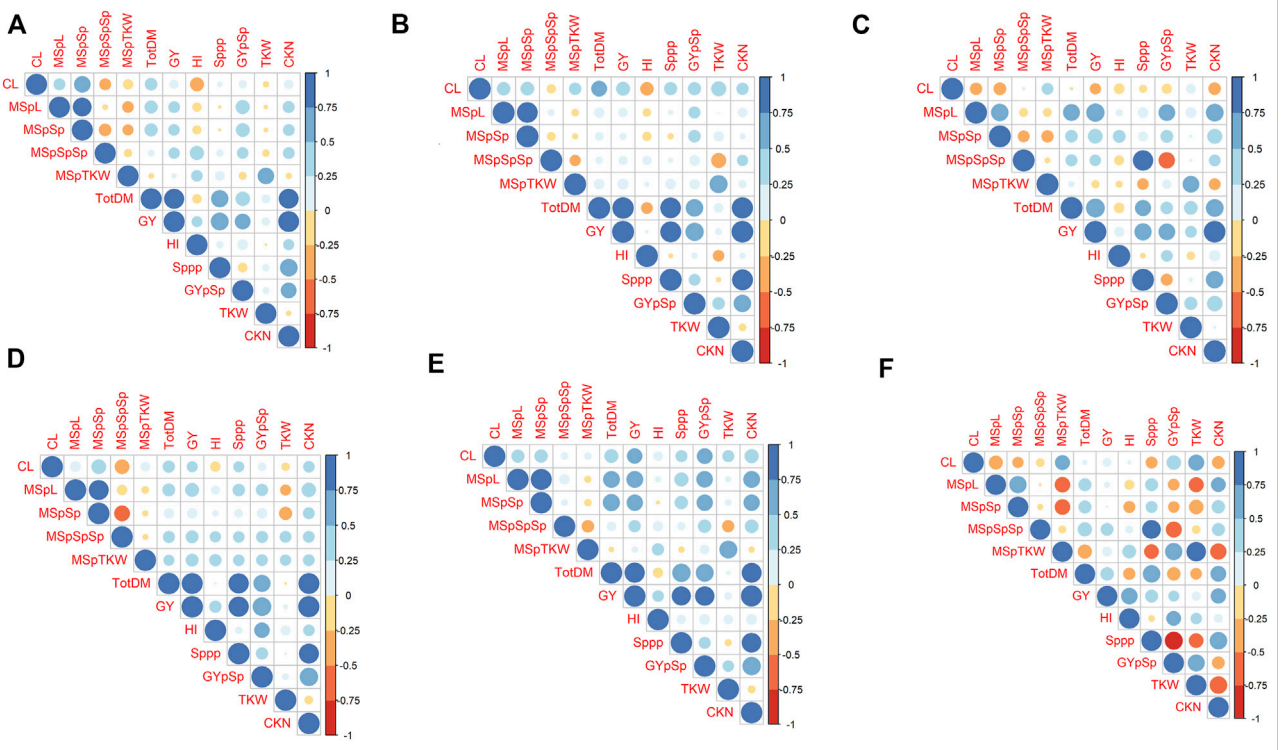


FIGURE 3
Correlation analysis of phenotypic data, obtained under water-limited (WL) and well-watered conditions (WW) in 2017, 2018 and 2019. Results for 2017, 2018 and 2019 were illustrated from the left to the right. The subplots **(A–C)** illustrate results that were obtained under WW conditions, **(D–F)** illustrate results for WL-conditions. CKN = Calculated kernel number, CL = Culm Length, GY = Grain yield, HI = Harvest index, TKW = Thousand kernel weight, TotDM = Total Dry Matter, MSpL = Main Spike Length, MSpTKW = Main Spike thousand kernel weight, MSpSp = Main Spike Spikelets, MSpSpSp = Seeds per spikelet (of the main spike), Sppp = Spikelets per plant.

harvest index (HI), the effect of drought stress was very clear for all traits in 2018 and 2019 (Figure 2, Supplementary Figure S4, Supplementary Table S5).

In general, plants under drought stress have less GY, less MSpTKW, a reduced calculated kernel number (CKN), and shorter culms than under WW conditions (Figure 2 and Supplementary Figure S4, Supplementary Table S5, Supplementary Table S6). Traits such as CKN, MSpSp, MSpL, and the number of spikes per plant (Sppp) correlate most positively with GY under WW and WL conditions (Figure 3). In addition to this, HI correlates more negatively with CL and slightly more positively with CKN under WW conditions.

These drought stress effects were observed independently in 2018 and 2019. Genotypic and phenotypic data of all years and both water regimes are given in Supplementary Table S7.

Sub-QTL-detection

Sub-QTLs under well-watered and water-limited conditions were calculated using data from 2017, 2018, and 2019, respectively. As mentioned previously, CL, MSpTKW, and HI did not show significant differences between drought and control conditions in 2017. Therefore, the data of these traits were not considered to calculate drought-specific QTLs. Table 3 shows some general information about the QTLs obtained, while Table 4 shows the logarithm of odds (LOD) values and dominance (d) effects of significant QTLs ($p < 0.05$) along the whole target interval. While QTLs for CL and TKW could be detected in both environments, QTLs for CKN or HI appear under drought stress or control conditions, respectively.

The QTLs for HI, CKN, and CL co-localize at the very distal parts of the QTL interval, leading to a biased set of segmental sub-NILs (i.e., northern or southern recombinants) which lack or harbor the respective part of the QTL (Table 2, Table 4 and Figure 4). Sub-NILs from the upper part of the wild emmer QTL tend to have shorter culms, an increased number of kernels, and slightly higher GY than sub-Nils than the lower part of the QTL (Table2; Figure 4).

Furthermore, different QTLs related to traits of the main spike, e.g., for MSpSP, MSpTKW, and MSpL, were detected under WL and WW conditions between the markers

TABLE 3 Significant QTLs detected in 2017, 2018, and 2019.

	Year	2017					2018					2019					2017–2019				
Condition	**Trait**	**LOD**	**P**	**Closest marker**	**PEV**	**Subst. Effect**	**LOD**	**P**	**Closest marker**	**PEV**	**Subst. Effect**	**LOD**	**P**	**Closest marker**	**PEV**	**Subst. Effect**	**LOD**	**P**	**Closest marker**	**PEV**	**Subst. Effect**
WL	CKN	2.39	0.01	*Kukri_rep_c69803_82*	0.50	24.60	1.96	0.06	*NOT SIGNIFICANT*	0.45	−13.67	1.81	0.08	*NOT SIGNIFICANT*	0.40	−17.85	2.59	0.01	*wsnp_Ex_c6537_1133876*	0.55	−13.00
	CL	*No significant difference between drought and control*					1.31	0.17	*NOT SIGNIFICANT*	0.37	−3.31	2.86	0.01	*wsnp_Ex_c6537_1133876*	0.54	3.88	2.88	0.01	*Kukri_rep_c69803_82*	0.65	−3.45
	HI						0.62	0.49	*NOT SIGNIFICANT*	0.17	−2.43	0.32	0.82	*NOT SIGNIFICANT*	0.10	1.17	1.33	0.19	*NOT SIGNIFICANT*		1.89
	MSpSp	0.46	0.66	*NOT SIGNIFICANT*	0.12	-1.01	0.70	0.46	*NOT SIGNIFICANT*	0.20	−0.88	2.55	0.02	*Kukri_c46621_14*	0.51	−1.03	2.24	0.03	*Tdurum_contig_68806_677*	0.60	−1.07
	MSpTKW	*No significant difference between drought and control*					2.95	0.01	*wsnp_Ex_c6537_1133876*	0.60	3.59	2.14	0.04	*Kukri_c46621_14*	0.45	6.47	1.14	0.20	*NOT SIGNIFICANT*	0.37	−3.38
	TKW	0.78	0.38	*NOT SIGNIFICANT*	0.20	2.41	3.86	0.00	*wsnp_Ex_c6537_1133876*	0.70	4.64	2.48	0.01	*wsnp_Ex_c6537_1133876*	0.49	6.04	2.67	0.01	*Kukri_rep_c69803_82*	0.61	−4.16
	CKNLOG	2.37	0.01	*Kukri_rep_c69803_82*	0.50	0.14															
WW	CL	3.01	0.01	*wsnp_Ex_c6537_1133876*	0.56	7.63	1.94	0.06	*NOT SIGNIFICANT*	0.46	4.14	2.34	0.00	*wsnp_Ex_c6537_1133876*	0.47	2.82	2.98	0.01	*Kukri_rep_c69803_82*	0.66	−4.20
	GYpSp	1.27	0.14	*NOT SIGNIFICANT*	0.29	−0.20	0.41	0.71	*NOT SIGNIFICANT*	0.12	−0.11	1.39	0.11	*NOT SIGNIFICANT*	0.32	−0.22	1.64	0.06	*Kukri_rep_c69803_82*	0.40	0.16
	HI	3.67	0.00	*wsnp_Ex_c6537_1133876*	0.63	−5.36	1.50	0.11	*NOT SIGNIFICANT*	0.38	3.59	1.61	0.09	*NOT SIGNIFICANT*	0.36	7.91	2.69	0.01	*BS00010055_51*	0.56	6.00
	MSpL	1.00	0.25	*NOT SIGNIFICANT*	0.25	−2.92	0.58	0.53	*NOT SIGNIFICANT*	0.18	−2.77	2.41	0.02	*Tdurum_contig_68806_677*	0.55	−6.66	1.78	0.06	*Kukri_c6227780*	0.50	−3.73
	MSpSp	0.46	0.61	*NOT SIGNIFICANT*	0.12	−0.91	0.49	0.64	*NOT SIGNIFICANT*	0.16	−0.81	2.90	0.02	*Kukri_c46621_14*	0.55	−1.44	1.81	0.08	*NOT SIGNIFICANT*	0.55	−1.02
	MSpSpSp	1.65	0.08	*NOT SIGNIFICANT*	0.36	-0.02	2.75	0.02	*Kukri_rep_c69803_82*	0.64	0.05	0.61	0.50	*NOT SIGNIFICANT*	0.17	−0.02	2.33	0.03	*Gene-1741_103*	0.52	0.02
	TKW	1.15	0.17	*NOT SIGNIFICANT*	0.35	1.89	2.82	0.01	*Kukri_rep_c69803_82*	0.59	−4.08	1.62	0.04	24.90	0.38	−2.21	2.38	0.04	*Kukri_rep_c69803_82*	0.53	−1.96
	MSpLLOG											2.44	0.01	*Tdurum_contig_68806_677*	0.55	−0.11					

LOD, Logarithm of odds. The sign of the QTL, substitution effect d represents the effect of the G18-16 on the respective trait. It is the difference of the mean of homozygote Sub-NILs, with the G18-16 and Uzan-allele, respectively. CKN, calculated kernel number; CL, culm length; GY, grain yield; HI, harvest index; MSPL, main spike length; MSpSp, Main spike spikelets; MSpSpSp, Main Spike seeds per spikelet; MSpTKW, main spike thousand kernel weight; TKW, thousand kernel weight; TotDM, total dry matter; P, Probability (based on a permutation test with 1,000 repeats); PEV, Percentage of explained variance. Significant QTLs were highlighted in gray. In a few cases, the respective data of the year were not normal distributed and log transformed. These data were analyzed separately and written in italics. Log transformation was not successful in 2018 for the SPPP under WW and WL, conditions.

10.3389/fgene.2022.955295

TABLE 4 QTL LOD-values along the target interval.

Condition	Year	iSelect	LOD-Values	*Tdurum_contig27976_414*	*Kukri_rep_c69803_82*	*BS00010055_51*	*Kukri_c62277_80*	GBS _96013158	GBS _97013837	*Gene-1741_103*	GBS _100504657	GBS _10264774	*Tdurum_contig30989_79*	*Tdurum_contig68806_677*	GBS_113979606	GBS_115312227	*Rac875_c21378_474*	*Kukri_c46621_143*	*wsnp_Ex_c6537_1133876*
		BP		77.16		80.47	96.01	96.01	97.01	97.70	100.50	102.65	10.41	111.43	113.98	115.31	122.51	128.06	138.10
		cM		*35.44*		*36.48*	*39.36*			*40.75*			*42.15*	*43.95*			*46.81*	*47.86*	*48.55*
WW	2017	CL	LOD	3.01	0.45	0.16	0.01	0.01	0.01	0.01	0.01	0.00	0.11	0.28	0.88	1.40	1.40	3.01
			eff.(d)	−7.63	−3.60	−2.11	0.40	0.55	0.55	0.55	0.55	−0.15	1.73	2.87	5.14	5.83	5.83	7.63
		HI	LOD	3.67	1.91	1.91	1.12	1.12	1.02	1.02	1.02	0.35	0.35	0.40	1.23	2.71	2.71	3.67
			eff.(d)	5.36	4.35	4.35	3.60	3.60	3.36	3.36	3.36	2.02	2.02	−2.26	−3.94	−4.94	−4.94	−5.36
	2018	CL	LOD	1.66	1.66	1.52	1.21	1.21	1.16	1.16	1.16	0.39	0.39	0.39	1.12	1.94	1.94	1.94
			eff.(d)	−3.90	−3.90	−3.79	−3.58	−3.58	−3.38	−3.38	−3.38	−2.04	−2.04	2.15	3.70	4.14	4.14	4.14
		MSpSpSp	LOD	2.75	1.73	1.54	1.17	1.23	1.23	1.23	1.23	0.65	0.65	0.05	0.09	0.52	0.52	1.65
			eff.(d)	0.05	0.04	0.04	0.03	0.03	0.03	0.03	0.03	0.02	0.02	−0.01	−0.01	−0.02	−0.02	−0.04
		TKW	LOD	2.82	2.82	1.63	0.66	0.66	0.40	0.40	0.40	0.21	0.21	0.10	0.25	1.34	1.34	1.34
			eff.(d)	−4.08	−4.08	−3.30	−2.39	−2.39	−1.83	−1.83	−1.83	−1.32	−1.32	0.96	1.60	3.11	3.11	3.10
	2019	CL	LOD	2.22	−1.43	−3.54	−0.42	−0.40	−0.42	−0.42	−0.44	0.01	0.97	1.19	1.63	1.63	1.56	2.34
			eff.(d)	−2.82	−1.70	−2.23	−0.90	−0.90	−0.84	−0.84	−0.84	0.17	2.00	2.25	2.69	2.69	2.45	2.82
		MSpL	LOD	0.60	0.00	0.84	0.84	1.30	1.30	1.30	1.98	1.98	2.41	1.99	1.47	1.79	1.79	1.79
			eff.(d)	3.50	0.17	−4.24	−4.24	−4.99	−4.99	−4.99	−5.82	-5.82	-6.66	−5.92	−5.67	−5.67	−5.67	−5.67
		MSpSp	LOD	2.64	0.91	0.91	0.03	0.03	0.00	0.00	0.07	0.07	0.61	0.64	1.34	2.90	2.90	2.90
			eff.(d)	1.48	0.92	0.92	0.18	0.18	−0.03	−0.03	−0.27	-0.27	−0.76	−0.80	−1.16	−1.44	−1.44	−1.44
		MSpLOG	LOD	0.64	0.00	0.83	0.83	1.30	1.30	1.30	1.99	1.99	2.44	2.05	1.53	1.81	1.81	1.81
			eff.(d)	0.06	0.00	−0.07	−0.07	−0.08	−0.08	−0.08	−0.10	−0.10	-0.11	−0.10	−0.10	−0.10	−0.10	-0.10

(Continued on following page)

10.3389/fgene.2022.955295

TABLE 4 (*Continued*) QTL LOD-values along the target interval.

Condition	Year	iSelect	LOD-Values	*Tdurum_contig27976_414*	*Kukri_rep_c69803_82*	*BS00010055_51*	*Kukri_c62277_80*	GBS _96013158	GBS _97013837	*Gene-1741_103*	GBS _100504657	GBS _10264774	*Tdurum_contig30989_79*	*Tdurum_contig68806_677*	GBS_113979606	GBS_115312227	*Rac875_c21378_474*	*Kukri_c46621_143*	*wsnp_Ex_c6537_1133876*
		BP			77.16	80.47	96.01	96.01	97.01	97.70	100.50	102.65	10.41	111.43	113.98	115.31	122.51	128.06	138.10
		cM			35.44	36.48	39.36			40.75			42.15	43.95			46.81	47.86	48.55
	2017 to 2019	CL	LOD	2.98	2.37	1.59	1.09	1.14	1.14	1.14	1.14	0.80	0.80	0.12	0.71	1.20	1.20	1.45
			eff.(d)	−4.20	−3.77	−3.20	−2.91	−2.86	−2.86	−2.86	−2.86	−2.41	−2.41	1.05	2.58	2.91	2.91	3.09
		HI	LOD	2.28	2.69	2.69	1.81	1.81	1.82	1.82	1.82	1.61	1.61	0.08	0.19	1.30	1.30	1.49
			eff.(d)	5.81	6.00	6.00	5.52	5.52	5.32	5.32	5.32	4.99	4.99	−1.31	−2.15	-4.68	−4.68	−4.84
		MSpSpSp	LOD	0.41	0.58	0.77	0.77	2.33	2.33	2.33	2.33	1.04	1.04	0.19	0.13	0.05	0.16	0.41
			eff.(d)	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.00	-0.01	−0.01
		TKW	LOD	2.38	2.38	0.89	0.82	0.82	0.34	0.34	0.34	0.13	0.13	0.13	0.54	1.49	1.49	1.56
			eff.(d)	−1.96	−1.96	−1.30	−1.33	−1.33	−0.86	−0.86	−0.86	−0.52	−0.52	0.55	1.18	1.65	1.65	1.65
WL	2017	CKN	LOD	2.39	2.39	1.97	0.15	0.15	0.00	0.00	0.03	0.03	0.11	0.16	0.36	0.47	0.47	0.88
			eff.(d)	24.60	24.60	22.23	7.04	7.04	0.99	0.99	2.97	2.97	−5.94	−7.24	−11.44	−11.91	−11.91	−15.73
		CKNLOG	LOD	2.37	2.37	1.99	0.20	0.20	0.01	0.01	0.05	0.05	0.11	0.16	0.39	0.47	0.47	0.90
			eff.(d)	0.14	0.14	0.13	0.05	0.05	0.01	0.01	0.02	0.02	−0.03	−0.04	−0.07	−0.07	−0.07	−0.09
	2018	MSpTKW	LOD	2.95	1.37	1.38	0.72	0.72	0.40	0.40	0.40	0.14	0.14	0.19	0.78	1.43	1.43	2.95
			eff.(d)	−3.59	−2.72	−2.84	−2.18	−2.18	−1.61	−1.61	−1.61	−0.95	−0.95	1.16	2.42	2.82	2.82	3.59
		TKW	LOD	3.86	2.69	2.29	1.04	1.04	0.28	0.28	0.28	0.17	0.17	0.21	0.92	2.60	2.60	3.86
			eff.(d)	−4.64	−4.25	−3.96	−3.08	−3.08	−1.63	−1.63	−1.63	−1.27	−1.27	1.48	3.11	4.20	4.20	4.64
	2019	CL	LOD	2.86	0.47	0.47	0.02	0.04	0.04	0.04	0.04	0.00	0.30	0.30	0.53	1.39	1.39	2.86
			eff.(d)	−3.88	−1.85	−1.85	−0.38	−0.56	−0.56	−0.56	−0.56	0.12	1.49	1.49	2.12	3.00	3.00	3.88

(Continued on following page)

10.3389/fgene.2022.955295

TABLE 4 (*Continued*) QTL LOD-values along the target interval.

Condition	Year	iSelect	LOD-Values	*Tdurum_contig27976_414*	*Kukri_rep_c69803_82*	*BS00010055_51*	*Kukri_c62277_80*	GBS _96013158	GBS _97013837	*Gene-1741_103*	GBS _100504657	GBS _10264774	*Tdurum_contig30989_79*	*Tdurum_contig68806_677*	GBS_113979606	GBS_115312227	*Rac875_c21378_474*	*Kukri_c46621_143*	*wsnp_Ex_c6537_1133876*
		BP			**77.16**	**80.47**	**96.01**	**96.01**	**97.01**	**97.70**	**100.50**	**102.65**	**10.41**	**111.43**	**113.98**	**115.31**	**122.51**	**128.06**	**138.10**
		cM			***35.44***	***36.48***	***39.36***			***40.75***			***42.15***	***43.95***			***46.81***	***47.86***	***48.55***
2017to 2019		MSpSp	LOD		1.67	0.28	0.24	0.24	0.24	0.22	0.22	0.46	0.46	1.10	1.10	1.45	2.55	2.55	2.55
			eff.(d)		0.87	0.41	−0.38	−0.38	−0.38	−0.35	−0.35	−0.49	−0.49	−0.74	−0.74	−0.90	−1.03	−1.03	−1.03
		MSpTKW	LOD		1.44	0.23	0.16	0.11	0.12	0.12	0.12	0.15	0.15	0.63	0.63	0.86	2.14	2.14	2.14
			eff.(d)		−5.47	−2.46	−1.99	1.72	1.74	1.74	1.74	1.89	1.89	3.87	3.87	4.81	6.47	6.47	6.47
		TKW	LOD		2.48	0.52	0.52	0.00	0.01	0.01	0.01	0.01	0.00	0.32	0.32	0.63	2.37	2.37	2.48
			eff.(d)		−6.04	−3.19	−3.19	−0.26	−0.36	−0.36	−0.36	−0.36	−0.01	2.51	2.51	3.74	6.03	6.03	6.04
		CKN	LOD		2.59	2.45	2.45	0.34	0.34	0.11	0.11	0.11	0.02	0.33	0.33	0.81	1.58	1.58	2.59
			eff.(d)		13.31	12.77	12.77	5.81	5.81	3.20	3.20	3.20	1.46	−5.54	−5.54	−9.30	−11.08	−11.08	−13.00
		CL	LOD		2.88	1.94	1.85	0.68	0.68	0.55	0.55	0.55	0.31	0.31	0.23	0.73	1.32	1.32	1.47
			eff.(d)		−3.45	−2.91	−2.80	−1.96	−1.96	−1.70	−1.70	−1.70	−1.28	−1.28	1.17	2.16	2.51	2.51	2.57
		MSpSp	LOD		0.22	0.23	0.23	0.09	0.24	0.24	0.24	0.41	0.41	2.24	1.74	1.10	0.76	0.25	0.25
			eff.(d)		0.36	0.36	0.36	−0.24	−0.37	−0.37	−0.37	−0.48	−0.48	−1.07	−0.91	−0.78	−0.71	−0.38	−0.38
		TKW	LOD		2.67	2.45	2.28	0.77	0.77	0.41	0.41	0.41	0.18	0.31	0.31	0.48	1.55	1.55	1.55
			eff.(d)		−4.16	−3.93	−3.77	−2.57	−2.57	−1.87	−1.87	−1.87	−1.25	1.64	1.64	2.21	3.32	3.32	3.32

[a]LOD, logarithm of odds; d, sign of the QTL, substitution effect d represents the effect of G18-16 on the respective trait. It is the difference of the mean of homozygote sub-NILs, with the G18-16 and Uzan allele, respectively, CKN, calculated kernel number; GY, grain yield; HI, harvest index; MSPL, main spike length; TKW, thousand kernel weight; P, Probability (based on a permutation test with 1,000 repeats); PEV, Percentage of explained variance. No physical position was obtained for the marker Tdurum_contig27976_414. *** Genetic positions were obtained from the original mapping population (Fatiukha et al., 2021). Significant QTLs, with positive or negative effects on the trait and a LOD >2.0 are colored in yellow and blue.

10.3389/fgene.2022.955295

Gene_1741_1 and Tdurum_contig_68806_677 (Table 4). Sub-NILs with the wild emmer allele of the marker Tdurum_contig_68806_677 appear to have a smaller main ear with fewer main ear spikelets (MSpSp) than the elite parent Uzan (Figure 5). Differentiating again between those sub-NILs which carry the upper part of the QTL interval and the G18-16 allele of the marker Tdurum_contig27976_414, but not the G18-16 allele of the flanking marker Gene-1741_103, revealed that these lines (i.e., 1663G, 1688C, 1488A, 1767E, and 1174B) again have a significantly higher HI, GY, and more CKN under WW conditions in all 3 years and under mild drought stress conditions in 2017 (Figure 6 and Table 4, Table 2). The

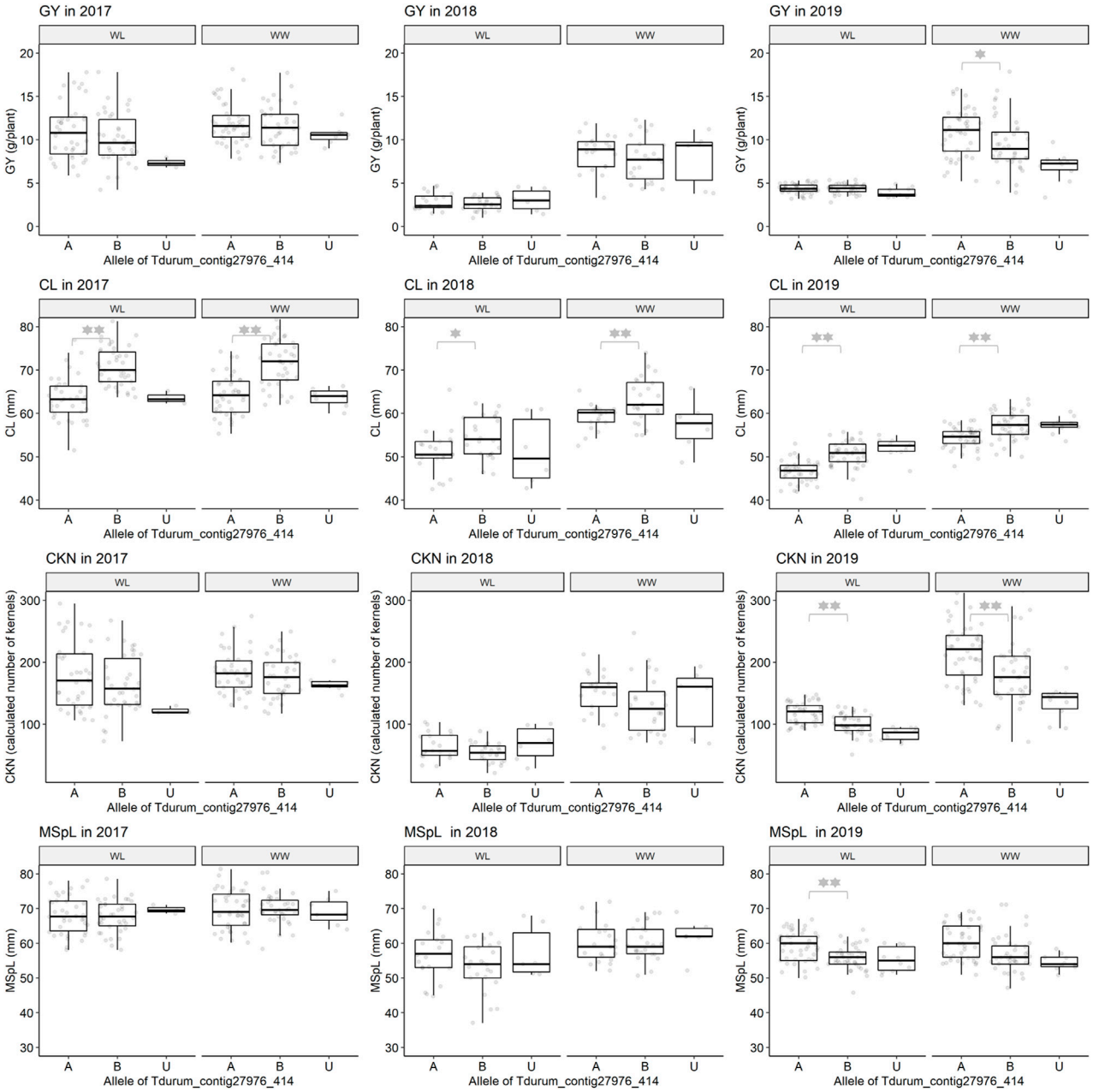


FIGURE 4
Grain yield (GY), culm length (CL), calculated kernel number (CKN), and main spike length (MSpL) of segmental sub-NILs with G18-16 or Uzan alleles of the iSelect marker Tdurum_contig27976_414. A and B Alleles were derived from G18-16 or Uzan, respectively. Values obtained for the elite parent Uzan (U) were illustrated separately. Significant differences between groups of genotypes with the different alleles of Tdurum_contig27976_414 were calculated with a two-sided *t*-test. *p*-values ≤ 0.05 or 0.01 were marked with one or two stars, respectively.

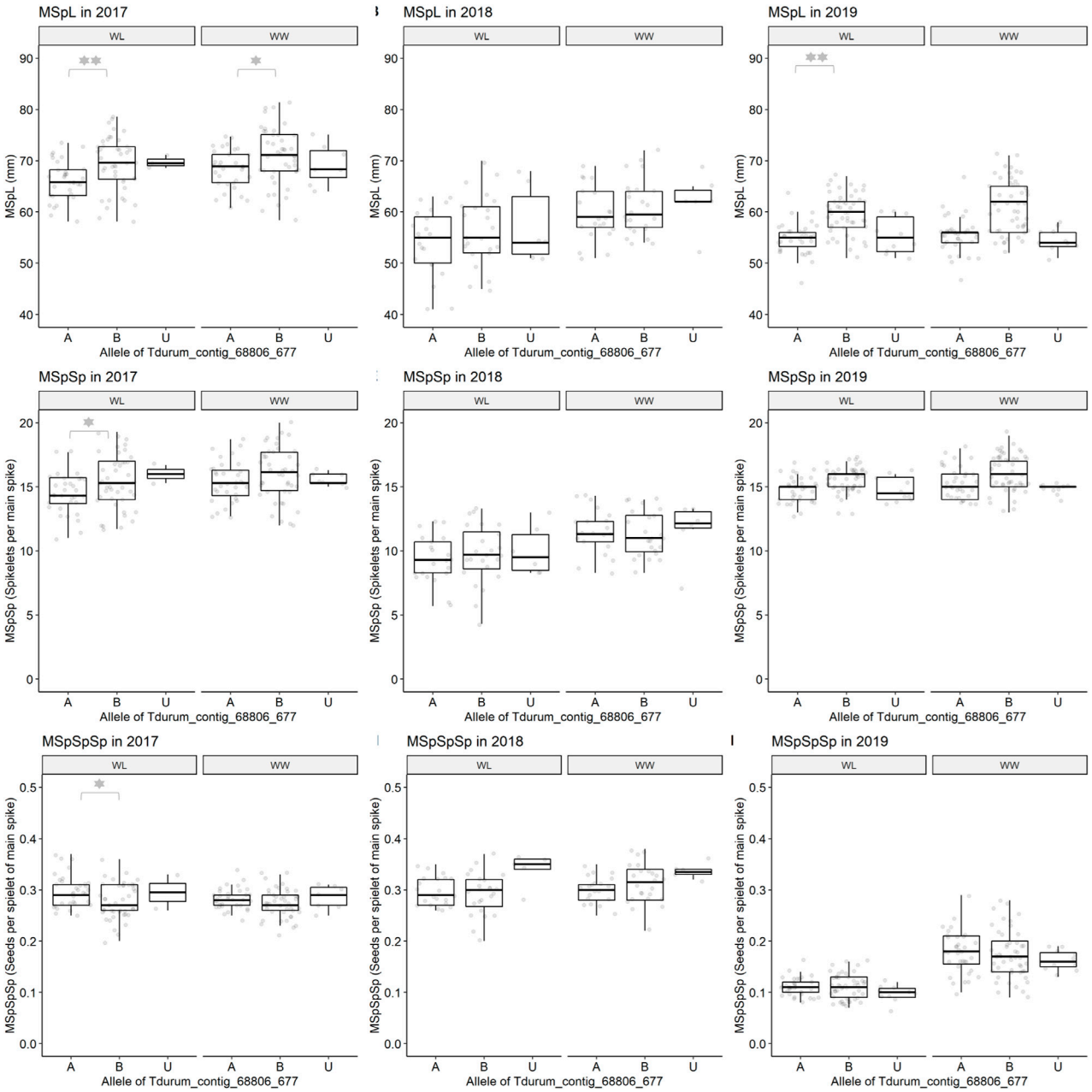


FIGURE 5
Main spike length (MSpL), main spike spikelets (MSpSp), and main spike seeds per spikelet (MSpSpSp) of segmental sub-NILs, carrying different alleles of the iSelect marker Tdurum_contig_68806_677. A and B Alleles were derived from G18-16 or Uzan, respectively. Values obtained for the elite parent Uzan (Uz) were illustrated separately. Significant differences between groups of genotypes with the different alleles of Tdurum_contig68806_677 were calculated with a two-sided *t*-test. *p*-values ≤ 0.05 or 0.01 were marked with one or two stars, respectively.

average GY of these lines under WW conditions was 12.05 g (2017), 9.48 g (2018), and 12.07 g (2019), whereas Uzan yielded 10.60 g, 7.92 g, or 6.92 g (Table 5). The average GY of these lines under WW conditions in 2017, 2018, and 2019 was thus 12, 16, and 42% higher than of Uzan, respectively (Table 5).

Discussion and conclusion

In general, the following observations should be considered when comparing the results of this article to the previous studies by Peleg et al. (2009), Merchuk-Ovnat et al. (2016) and Fatiukha et al. (2021). First, it should be noted that the genetic background

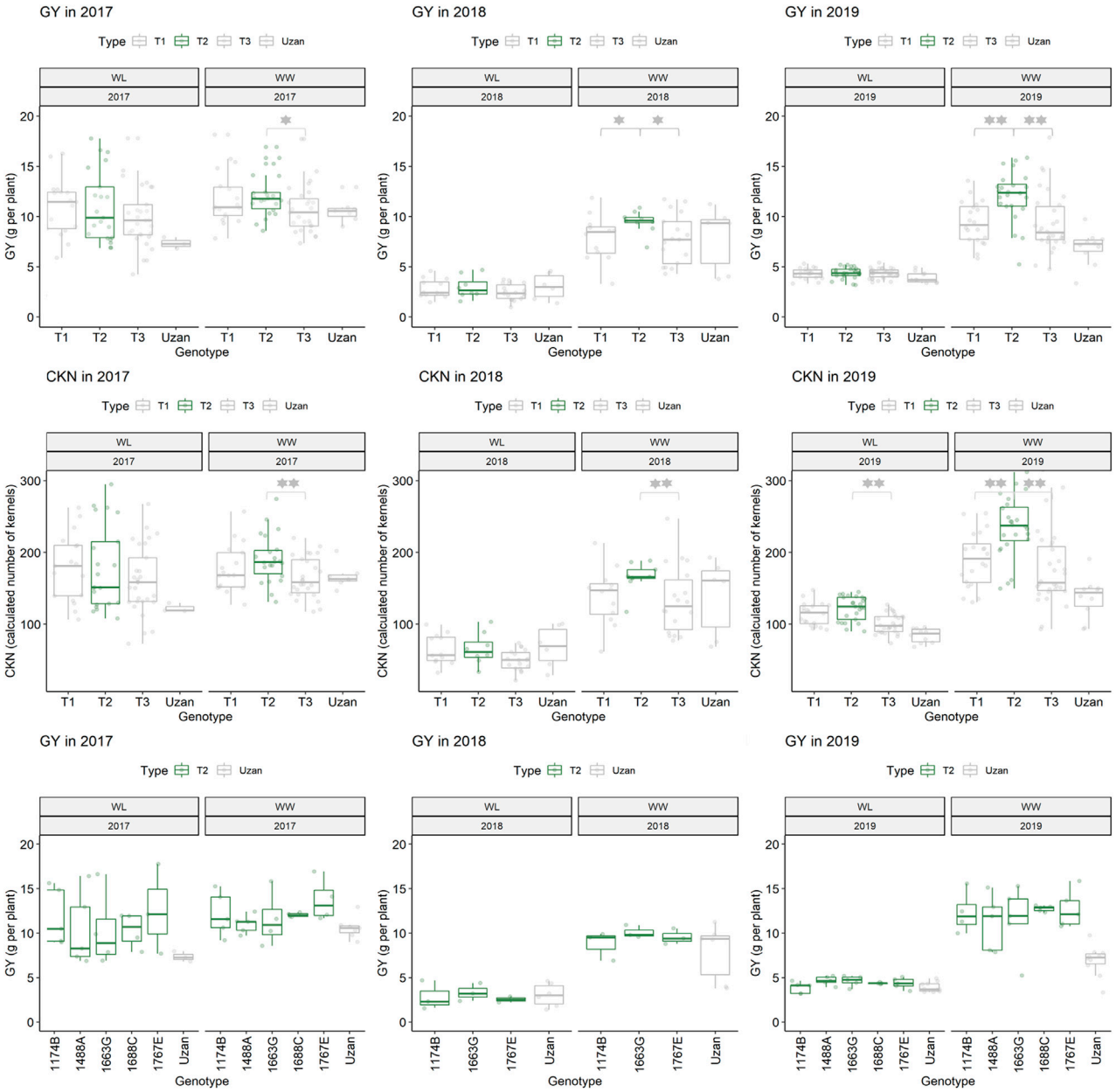


FIGURE 6
Grain yield (GY) and calculated kernel number (CKN) of segmental sub-NILs, carrying different alleles of the iSelect marker Tdurum_contig27976_414 and Gene-1741_103. T1: sub-NILs with the wild emmer allele from Tdurum_contig27976_414 and Gene-1741_103. T2 (green): sub-NILs with the wild emmer allele of Tdurum_contig27976_414 and elite parent allele of Gene-1741_103. T3: sub-NILs with the elite parent allele of Tdurum_contig27976_414 and Gene-1741_103. Uzan is the name of the respective elite parent. Figure C shows the GY values of each of the sub-NIL that belongs to T2. Significsant differences between the T1, T2, and T3 groups of genotypes were calculated with a two-sided *t*-test. *p*-values ≤ 0.05 or 0.01 were marked with one or two stars, respectively.

of the NIL-U2B-1 comes from cv. Uzan and not from cv. Langdon, the original parent of the mapping population (Peleg et al., 2009; Fatiukha et al., 2021). Although the effect of this QTL in NIL-U-2B-1 was confirmed previously (Merchuk-Ovnat et al., 2016), the different genomic backgrounds of Uzan against which the effect of the QTL region was determined in this study might have an impact on the detected QTLs. Second, the drought stress conditions that were used in 2018 and 2019 were probably more severe than in previous studies. Peleg et al. (2009) used 350 mm and Merchuk-Ovnat et al. (2016) used

TABLE 5 Mean values of segmental sub-NILs, carrying different alleles of the iSelect marker Tdurum_contig27976_414 and Gene-1741_103.

	Condition**	Year	Genotype[a]	Traits **											
				MSpSp	CL	MSpL	MSpSpSp	MSpTKW	TotDM	GY	HI	Sppp	GYpSp	TKW	CKN
Mean values	WL	2017	T1	14.40	63.71	66.68	0.32	67.96	22.83	10.96	48.72	5.28	2.03	62.38	178.67
			T2	14.81	63.49	69.38	0.29	66.36	23.51	11.00	46.53	5.06	2.06	61.85	178.45
			T3	15.22	70.43	68.18	0.27	66.01	23.67	9.89	42.27	5.23	1.87	61.92	160.88
			Uzan	16.00	63.60	69.73	0.30	65.51	16.75	7.35	44.41	4.63	1.59	59.96	122.46
		2018	T1	9.47	50.54	55.08	0.32	48.28	6.72	2.83	41.99	2.18	1.30	43.82	64.72
			T2	10.44	51.68	58.67	0.31	48.33	6.56	2.96	45.66	2.23	1.34	44.72	65.41
			T3	8.88	53.88	52.67	0.28	52.31	5.61	2.46	44.16	1.97	1.23	49.31	49.52
			Uzan	10.05	51.37	57.33	0.34	47.97	6.63	3.03	47.02	2.17	1.35	44.37	68.48
		2019	T1	15.30	47.99	55.60	0.10	43.90	4.17	4.32	41.55	3.55	1.24	38.17	113.90
			T2	15.92	45.55	62.28	0.11	40.48	4.32	4.38	42.18	3.92	1.14	36.73	121.92
			T3	14.90	51.27	56.45	0.11	46.94	4.24	4.33	42.09	3.66	1.21	43.42	100.27
			Uzan	14.80	52.10	55.40	0.10	51.40	3.63	3.94	41.82	2.90	1.37	46.85	84.15
	WW	2017	T1	15.12	64.36	67.53	0.29	68.55	24.71	11.83	49.09	4.55	2.53	63.57	177.60
			T2	15.53	64.06	71.46	0.28	66.05	25.78	12.05	46.79	4.77	2.49	63.22	190.80
			T3	15.95	70.39	70.33	0.26	67.05	25.15	10.69	41.52	4.71	2.24	63.41	164.74
			Uzan	15.60	63.67	69.24	0.29	69.45	21.31	10.60	49.84	4.75	2.26	63.44	167.01
		2018	T1	10.83	58.30	59.08	0.34	62.36	16.52	7.87	47.93	3.54	2.20	56.54	138.91
			T2	11.33	60.67	61.78	0.33	62.38	20.13	9.48	47.00	4.07	2.32	57.36	165.42
			T3	11.71	63.36	61.17	0.29	63.85	17.14	7.71	43.44	3.89	2.12	60.59	133.52
			Uzan	11.60	57.25	62.00	0.34	60.92	16.48	7.92	47.22	3.35	2.32	56.33	139.43
		2019	T1	16.20	54.83	56.65	0.16	52.37	7.45	9.38	58.98	6.25	1.51	49.36	189.82
			T2	16.58	54.03	63.58	0.17	53.24	8.77	12.07	55.45	6.79	1.76	51.72	235.38
			T3	15.03	57.69	57.67	0.18	54.48	8.31	9.37	49.60	6.40	1.50	52.79	177.00
			Uzan	14.80	57.28	54.60	0.16	54.70	6.86	6.92	45.51	5.30	1.41	51.70	137.42

(Continued on following page)

10.3389/fgene.2022.955295

TABLE 5 (*Continued*) Mean values of segmental sub-NILs, carrying different alleles of the iSelect marker Tdurum_contig27976_414 and Gene-1741_103.

	Condition**	Year	Genotype[a]	Traits **											
				MSpSp	CL	MSpL	MSpSpSp	MSpTKW	TotDM	GY	HI	Sppp	GYpSp	TKW	CKN
p-values of a two sided t-test ***	WL	2017	T1 vsT2	0.54	0.90	0.12	0.02	0.16	0.77	0.97	0.06	0.57	0.74	0.71	0.99
			T2 vsT3	0.50	0.00	0.43	0.02	0.72	0.94	0.25	0.00	0.55	0.05	0.92	0.26
		2018	T1 vsT2	0.33	0.66	0.30	0.43	0.97	0.88	0.79	0.33	0.81	0.74	0.5	0.95
			T1 vsT3	0.13	0.39	0.09	0.03	0.01	0.36	0.27	0.63	0.12	0.37	0.00	0.09
		2019	T1 vsT2	0.02	0.00	0.00	0.44	0.06	0.44	0.7	0.74	0.05	0.09	0.25	0.11
			T2 vsT3	0.00	0.00	0.00	0.43	0.00	0.65	0.74	0.93	0.14	0.16	0.00	0.00
	WW	2017	T1 vs. T2	0.50	0.84	0.02	0.13	0.05	0.51	0.78	0.00	0.36	0.70	0.66	0.23
			T2 vsT3	0.46	0.00	0.45	0.01	0.35	0.62	0.04	0.00	0.78	0.00	0.78	0.01
		2018	T1 vsT2	0.55	0.05	0.30	0.41	0.99	0.03	0.05	0.6	0.07	0.22	0.38	0.07
			T2 vsT3	0.65	0.08	0.80	0.00	0.16	0.05	0.02	0.05	0.51	0.08	0.00	0.02
		2019	T1 vsT2	0.38	0.25	0.00	0.70	0.51	0.00	0.00	0.06	0.18	0.00	0.04	0.00
			T2 vsT3	0.00	0.00	0.00	0.26	0.17	0.24	0.00	0.00	0.36	0.01	0.19	0.00

[a]T1: sub-NILS, with the wild emmer allele from Tdurum_contig27976_414 and Gene-1741_103. T2: sub-NILS, with the wild emmer allele of Tdurum_contig27976_414 and elite parent allele of Gene-1741_103. T3: sub-NILS, with the elite parent allele of Tdurum_contig27976_414 and Gene-1741_103. Uzan: the name of the respective elite parent. ** GY, grain yield; CL , culm length, Sppp = Spikelets per plant, MSpL = main spike length, MSpSp = Main Spike Spikelets, MSpSpSp = Seeds per spikelet of the main spike, MSpTKW, main spike thousand kernel weight; CKN, calculated kernel number; CL, culm length; HI , harvest index; WW, Well-watered treatment; WL , Water-limited treatment *** *p*-values below 0.05 are highlighted in gray.

10.3389/fgene.2022.955295

360–290 mm of seasonal water application to mimic drought stress. In this study, 350 mm of seasonal water application was used only once in 2017 resulting in mild drought stress symptoms. About 201 mm of seasonal water application was used in 2018, and in 2019, water application was reduced from 64 DAS to 20% PAW to simulate severe drought stress conditions (Dhanagond et al., 2019). Average GY was significantly reduced under drought stress to 10.39 g, 2.75 g, and 4.33 g per plant in 2017, 2018, and 2019, respectively. Lower GY in 2018 and 2019 is very likely a consequence of increased drought stress. Furthermore, ANOVA analysis revealed that the means of all traits were different between all 3 years (Supplementary Table S6). This is probably due to different reasons. On the one hand, the temperatures in the screen houses in 2018 were very high, i.e., above 45°C (Supplementary Table S8). These high temperatures suppressed tillering and induced early heading so that the plants were shorter, and yield was lower than in 2017. On the other hand, the experimental setup in 2019 differed significantly from 2017 to 2018. Plants in the pot experiments at the HTP in 2019 appeared to be smaller, to have reduced biomass, and less TKW (Supplementary Table S6, Supplementary Figure S4). This may be due to environmental factors, e.g., the size of the pots and light intensity. However, as mentioned in the results section, the drought stress effect in 2017 was mild but significant for most of the traits, except for CL, MSpTKW, and HI, and very clear for all traits in 2018 and 2019 (Figure 2, Supplementary Figure S4, Supplementary Table S5).

In contrast to the previous studies by Peleg et al. (2009) and Fatiukha et al. (2020), no QTL for GY could be confirmed, but different segments of the introgressed wild emmer wheat region on chromosome 2B of NIL-U-2B-1 showed impacts on CL, CKN, and GY. The sub-NILs with the upper part of the emmer wheat QTL-interval had a shorter CL, more CKN, and a slightly higher GY, while sub-NILs with an introgression at the lower part of the QTL had longer stems, less CKN, and reduced GY (Figure 4; Table 4, Table 5). Furthermore, the central part of the QTL interval has a negative impact on MSPL and the number of MSPSP. Removing sub-NILs with this MSPL-QTL from the set of sub-NILs that harbor the upper part of the QTL interval leads to five sub-NILs (i.e., sub-NIL 1663G, 1688C, 1488A, 1767E, and 1174B) with significantly ($p < 0.001$) higher HI, CKN, and GY than Uzan (Table2, Table 5, Figure 6). The increased number of kernels seems to have a decisive effect on the GY of these five lines. This effect was significant under mild drought stress in 2017 and controlled conditions in 2017 and 2019, but not under more severe drought stress conditions in 2018 and 2019. It can, therefore, be stated that the upper part of the introgressed fragment on chromosome 2B in the sub-NILs 1663G, 1688C, 1488A, 1767E, and 1174B might be of value for future breeding programs. The observed increased GY under mild drought stress conditions in 2017 is consistent with the results of previous studies (Peleg et al., 2009; Merchuk-Ovnat et al., 2016; Fatiukha et al., 2021). It might, therefore, be hypothesized that this region has a yield stabilizing effect under mild drought stress conditions. However, further trials with these sub-NILs would be useful to characterize up to which degree of drought stress the effect of this introgression remains advantageous.

The identification of candidate genes from the wild emmer parent that might be responsible for the increased kernel number remains difficult because a large fragment of about 50 million bp (Mbp) was introgressed upstream of Tdurum_contig27976_414 in NIL-U-2B-1 (Table 2, Deblieck et al., 2020). In addition to that, it is not clear which trait(s) are associated with the increased CKN of these plants. Plants with the upper part of the QTL region have an increased MSPL and an increased Sppp (T2 in Table 5). These differences are not significant (Table 5), and therefore, any further conclusions at this point of time remain speculative. However, it might be worth mentioning the presence of the *Ppd-B1* gene from wild emmer on chromosome 2B in NIL-U-2B-1, which has previously been demonstrated to specifically influence the number of seeds per spikelet and not DPH in durum wheat (Arjona et al., 2018). Alternative QTLs in the same region, but close to *Ppd-B1*, have also been identified in wheat in previous studies (Gao et al., 2015; Shi et al., 2017).

Referring to the QTL interval that might be related to MSpL, MSpSp, or MSpSpSp, 111 high-confidence genes were annotated between Gene-1741_103 and Tdurum_contig_68806_677 (Table 2, Table 4, Supplementary Table S9). One of these genes, the ethylene responsive factor (ERF) (TRIDC2BG016990), is a homolog of the well-analyzed Frizzy panicle (FZP) gene (LOC4344233) or Branched Silkless gene (BD1) in rice and maize, respectively (Colombo et al., 1998; Komatsu et al., 2001; Komatsu et al., 2003; Wang et al., 2020; Li et al., 2021). Wheat *fzp* lines were recently shown to have an increased number of spikelets, longer spike length, and reduced TKW (Li et al., 2021). Interestingly, these traits can be attributed to the elite parent allele (Table 4). It might, therefore, be hypothesized that a favorable FZP allele was selected in the course of evolution and introgressed into cultivars such as Uzan. Additional experiments, such as sequence and/or expression analysis, are required to prove or reject this hypothesis.

Finally, a CL-QTL was detected at the lower part of the QTL interval (Table 4). In accordance with this result, Zanke et al. (2014) confirmed that the iSelect marker wsnp_Ex_c6537_11338763 and markers downstream of it are associated with an increased plant height. The authors mention the existence of a gene in wheat that is orthologous to the GIBBERELLIN INSENSITIVE DWARF1 (GID1)-like receptor in rice (Zanke et al., 2014). This region was transferred from G18-16 into the NIL (Deblieck et al., 2020) and might have an impact on the detected CL-QTL. Remarkably, Merchuk-Ovnat et al. (2016) described that an alternative NIL, NIL-U-2B-3, carrying a smaller introgression of G18-16 on chromosome 2B than NIL-U-2B-1, downstream of Ku_c7740_879 (Supplementary Table S1, Merchuck-Ovnat et al., 2016) did not show such a strong increase in culm length relative to Uzan (Supplementary Table S5, Merchuck-Ovnat et al., 2016). Aligning Ku_c7740_879 to the reference genome of Zavitan (Avni et al., 2017) revealed that the marker is located at

164,96 Mbp on chromosome 2BS. We, therefore, hypothesize that the CL-QTL has a size of approximately 26,87 Mbp and ranges from 138.09 (Table 2) to 164,96 Mbp (Merchuck-Ovat et al. ,2016) on chromosome 2BS. This interval contains about 201 annotated high-confidence genes and was transferred into NIL-U-2B-1, but not into NIL-U-2B-3 (Merchuck-Ovnat et al., 2016; Avni et al., 2017). Aligning the rice GID1 gene (LOC4338764) against chromosome 2B of the reference genome of Zavitan (Avni et al., 2017), revealed a homolog candidate gene (TRIDC2BG021600) at position 146.52 Mbp within the interval. However, further fine-mapping of the introgressed G18-16 fragment that flanks wsnp_Ex_c6537_11338763 is required to further narrow down this QTL and to confirm or confute this candidate gene. In addition to this, it should also be taken into consideration that the reduced height gene 4 (Rht4) was described to be located on the long arm of chromosome 2B (Zanke et al., 2014; Wu et al., 2021). Aligning the primer sequences of the most significantly linked SSR marker WMC317 to the reference genome of Zavitan (Avni et al., 2017) revealed that it is located at 762 Mbp and was not transferred from G18-16 into NIL-U-2B-1 (Deblieck et al., 2020). Therefore, the observed differences in plant height cannot be attributed to Rht4.

The main tasks of this study were to identify the smallest sub-QTL region affecting grain yield and to diminish the effects of additional introgressions from the wild emmer parent in NIL-U-2B-1. Both goals were achieved. On the one hand, the delayed DPH and increased CL of NIL-U-2B-1 could be attributed to trans- and linkage drag effects of G18-16 introgressions on chromosome 2A or 2B, respectively. On the other hand, a GY and kernel number increasing G18-16 introgression was identified upstream of Tdurum_contig27976_414. The effect of this on the 50 million Mbp region was confirmed under controlled conditions in all years and under mild drought stress in 2017 (Table 5; Figure 6). Unfortunately, it flanks the QTL interval that was subjected to fine mapping and could not be further narrowed down. However, since the size of the region is much smaller than the original introgression in NIL-U-2B-1 (>400 Mbp) (Deblieck et al., 2020) and since the five selected sub-NILs with this fragment (sub-NILs 1663G, 1688C, 1488A, 1767E, and 1174B) do not show delayed DPH or increased CL, these genotypes might already be crossed with new elite cultivars to establish new NILs. In this respect, the effect of wild emmer introgression might be analyzed in different genomic backgrounds, environments, and years.

In addition to this, it might be useful to further examine the locus in the background of the cv. Uzan. A new fine-mapping approach might, therefore, be accomplished by again backcrossing 1663G, 1688C, 1488A, 1767E, or 1174B with the recurrent parent. Since the region upstream of Tdurum_contig27976_414 is located almost at the end of the chromosome (Table 2; Deblieck et al., 2020), recombinations should appear frequently in this region and fine mapping should be feasible. Finally, repeated trials with different levels of drought stress, Uzan, and the five sub-NILs might reveal to what extent of drought stress the actual yield-increasing effect of this wild emmer fragment persists.

Data availability statement

The datasets presented in this study can be found in online repositories. The data can be found using the following link: https://www.ebi.ac.uk/ena/browser/view/PRJEB53939.

Author contributions

MD developed the molecular markers, evaluated the phenotypic data, and wrote the manuscript. FO, YS, TK, and DP conceived the basic idea of this research project. YS, TK, FA, KN, ML, and GS established the plant material and phenotyped the plants. YS, KN, ML, KP, TK, and FA helped to evaluate the phenotypic data and to improve the manuscript.

Funding

This work was financially supported by the German Federal Ministry of Food and Agriculture (BMEL) based on the decision Hebrew Parliament Hebrew Federal Republic of Germany (FKZ: 2813IL03) and the Israel Ministry of Agriculture and Rural Development, Chief Scientist Foundation (Grants 837-0079-10 and 837-0162-14). GS is indebted to the Israel Ministry of Foreign Affairs for a post-doctoral fellowship award. YS is the incumbent of the Haim Gvati Chair in Agriculture..

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/fgene.2022.955295/full#supplementary-material

References

Arjona, J. M., Royo, C., Dreisigacker, S., Ammar, K., and Villegas, D. (2018). Effect of ppd-A1 and Ppd-B1 allelic variants on grain number and thousand kernel weight of durum wheat and their impact on final grain yield. *Front. Plant Sci.* 9, 888. doi:10.3389/fpls.2018.00888

Avni, R., Nave, M., Barad, O., Baruch, K., Twardziok, S. O., Gundlach, H., et al. (2017). Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. *Science* 357, 93–97. doi:10.1126/science.aan0032

Below, R., Grover-Kopec, E., and Dilley, M. (2007). Documenting drought-related disasters. *J. Environ. Dev.* 16, 328–344. doi:10.1177/1070496507306222

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

Chen, D., Wang, S., Cao, B., Cao, D., Leng, G., Li, H., et al. (2015). Genotypic variation in growth and physiological response to drought stress and Re-watering reveals the critical role of recovery in drought adaptation in maize seedlings. *Front. Plant Sci.* 6, 1241. doi:10.3389/fpls.2015.01241

Colombo, L., Marziani, G., Masiero, S., Wittich, P. E., Schmidt, R. J., Gorla, M. S., et al. (1998). BRANCHED SILKLESSmediates the transition from spikelet to floral meristem duringZea maysear development. *Plant J.* 16, 355–363. doi:10.1046/j.1365-313x.1998.00300.x

Deblieck, M., Fatiukha, A., Grundman, N., Merchuk-Ovnat, L., Saranga, Y., Krugman, T., et al. (2020). GenoTypeMapper: Graphical genotyping on genetic and sequence-based maps. *Plant Methods* 16, 123. doi:10.1186/s13007-020-00665-7

Dempewolf, H., Baute, G., Anderson, J., Kilian, B., Smith, C., and Guarino, L. (2017). Past and future use of wild relatives in crop breeding. *Crop Sci.* 57, 1070–1082. doi:10.2135/cropsci2016.10.0885

Dhanagond, S., Liu, G., Zhao, Y., Chen, D., Grieco, M., Reif, J., et al. (2019). Non-invasive phenotyping reveals genomic regions involved in pre-anthesis drought tolerance and recovery in spring barley. *Front. Plant Sci.* 10, 1307. doi:10.3389/fpls.2019.01307

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

Fahad, S., Bajwa, A. A., Nazir, U., Anjum, S. A., Farooq, A., Zohaib, A., et al. (2017). Crop production under drought and heat stress: Plant responses and management options. *Front. Plant Sci.* 8, 1147. doi:10.3389/fpls.2017.01147

Fang, Yujie, and Xiong, Lizhong (2015). General mechanisms of drought response and their application in drought resistance improvement in plants. *Cell. Mol. Life Sci.* 72 (4), 673–689. doi:10.1007/s00018-014-1767-0

Farooq, M., Hussain, M., and Siddique, K. H. M. (2014). Drought stress in wheat during flowering and grain-filling periods. *Crit. Rev. Plant Sci.* 33, 331–349. doi:10.1080/07352689.2014.875291

FAO (2021). World Food and Agriculture-Statistical Yearbook 2021.

Fatiukha, A., Deblieck, M., Klymiuk, V., Merchuk-Ovnat, L., Peleg, Z., Ordon, F., et al. (2021). Genomic architecture of phenotypic plasticity in response to water stress in tetraploid wheat. *Int. J. Mol. Sci.* 22, 1723. doi:10.3390/ijms22041723

Gao, F., Wen, W., Liu, J., Rasheed, A., Yin, G., Xia, X., et al. (2015). Genome-wide linkage mapping of QTL for yield components, plant height and yield-related physiological traits in the Chinese wheat cross zhou 8425B/Chinese spring. *Front. Plant Sci.* 6, 1099. doi:10.3389/fpls.2015.01099

Gaspar, T., Franck, T., Bisbis, B., Kevers, C., Jouve, L., Hausman, J. F., et al. (2002). *Plant Growth Regul.* 37, 263–285. doi:10.1023/A:1020835304842

Hackett, C. A. (1997). Model diagnostics for fitting QTL models to trait and marker data by interval mapping. *Heredity* 79, 319–328. doi:10.1038/hdy.1997.160

Huang, L., Raats, D., Sela, H., Klymiuk, V., Lidzbarsky, G., Feng, L., et al. (2016). Evolution and adaptation of wild emmer wheat populations to biotic and abiotic stresses. *Annu. Rev. Phytopathol.* 54, 279–301. doi:10.1146/annurev-phyto-080614-120254

Komatsu, M., Chujo, A., Nagato, Y., Shimamoto, K., and Kyozuka, J. (2003). FRIZZY PANICLE is required to prevent the formation of axillary meristems and to establish floral meristem identity in rice spikelets. *Development* 130, 3841–3850. doi:10.1242/dev.00564

Komatsu, M., Maekawa, M., Shimamoto, K., and Kyozuka, J. (2001). The LAX1 and FRIZZY PANICLE 2 genes determine the inflorescence architecture of rice by controlling rachis-branch and spikelet development. *Dev. Biol.* 231, 364–373. doi:10.1006/dbio.2000.9988

Lane, D. R., Coffin, D. P., and Lauenroth, W. K. (2000). Changes in grassland canopy structure across a precipitation gradient. *J. Veg. Sci.* 11, 359–368. doi:10.2307/3236628

Langridge, P., and Fleury, D. (2011). Making the most of 'omics' for crop breeding. *Trends Biotechnol.* 29, 33–40. doi:10.1016/j.tibtech.2010.09.006

Lesk, C., Rowhani, P., and Ramankutty, N. (2016). Influence of extreme weather disasters on global crop production. *Nature* 529, 84–87. doi:10.1038/nature16467

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Available at: https://www.semanticscholar.org/paper/Aligning-sequence-reads%2C-clone-sequences-and-with-Li/74574ee09030e8aadb48fa349eb9b054e2f95ceb.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi:10.1093/bioinformatics/btp352

Li, Y., Li, L., Zhao, M., Guo, L., Guo, X., Zhao, D., et al. (2021). Wheat FRIZZY PANICLE activates VERNALIZATION1-A and HOMEOBOX4-A to regulate spike development in wheat. *Plant Biotechnol. J.* 19, 1141–1154. doi:10.1111/pbi.13535

Merchuk-Ovnat, L., Barak, V., Fahima, T., Ordon, F., Lidzbarsky, G. A., Krugman, T., et al. (2016). Ancestral QTL alleles from wild emmer wheat improve drought resistance and productivity in modern wheat cultivars. *Front. Plant Sci.* 7, 452. doi:10.3389/fpls.2016.00452

Merchuk-Ovnat, L., Fahima, T., Ephrath, J. E., Krugman, T., and Saranga, Y. (2017). Ancestral QTL alleles from wild emmer wheat enhance root development under drought in modern wheat. *Front. Plant Sci.* 8, 703. doi:10.3389/fpls.2017.00703

Neumann, K., Klukas, C., Friedel, S., Rischbeck, P., Chen, D., Entzian, A., et al. (2015). Dissecting spatiotemporal biomass accumulation in barley under different water regimes using high-throughput image analysis. *Plant Cell Environ.* 38 (10), 1980–1996. doi:10.1111/pce.12516

Oyewole, C. I. (2016). *The wheat crop*. Anyigab: Kogi State University.

Pearson, K. (1900). X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Lond. Edinb. Dublin Philosophical Mag. J. Sci.* 50, 157–175. doi:10.1080/14786440009463897

Peleg, Z., Fahima, T., Krugman, T., Abbo, S., Yakir, D., Korol, A. B., et al. (2009). Genomic dissection of drought resistance in durum wheat x wild emmer wheat recombinant inbreed line population. *Plant Cell Environ.* 32, 758–779. doi:10.1111/j.1365-3040.2009.01956.x

Peleg, Z., Saranga, Y., Suprunova, T., Ronin, Y., Röder, M. S., Kilian, A., et al. (2008). High-density genetic map of durum wheat x wild emmer wheat based on SSR and DArT markers. *Theor. Appl. Genet.* 117, 103–115. doi:10.1007/s00122-008-0756-9

Poole, N., Donovan, J., and Erenstein, O. (2021). Viewpoint: Agri-nutrition research: Revisiting the contribution of maize and wheat to human nutrition and health. *Food Policy* 100 (2021), 101976. doi:10.1016/j.foodpol.2020.101976

Shapiro, S. S., and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika* 52, 591. doi:10.2307/2333709

Shi, W., Hao, C., Zhang, Y., Cheng, J., Zhang, Z., Liu, J., et al. (2017). A combined association mapping and linkage analysis of kernel number per spike in common wheat (*Triticum aestivum* L.). *Front. Plant Sci.* 8, 1412. doi:10.3389/fpls.2017.01412

Soleimani, B., Lehnert, H., Keilwagen, J., Plieske, J., Ordon, F., Naseri Rad, S., et al. (2020). Comparison between core set selection methods using different Illumina marker platforms: A case study of assessment of diversity in wheat. *Front. Plant Sci.* 11, 1040. doi:10.3389/fpls.2020.01040

Takenaka, S., and Kawahara, T. (2012). Evolution and dispersal of emmer wheat (Triticum sp.) from novel haplotypes of Ppd-1 (photoperiod response) genes and their surrounding DNA sequences. *Theor. Appl. Genet.* 125, 999–1014. doi:10.1007/s00122-012-1890-y

Tukey, J. W. (1949). Comparing individual means in the analysis of variance. *Biometrics* 5, 99–114. doi:10.2307/3001913

United Nations, Department of Economic and Social Affairs, Population Division (2017). World Population Prospects: The 2017 Revision, Key Findings and Advance Tables. Working Paper No. ESA/P/WP/248.

Wang, S.-S., Chung, C.-L., Chen, K.-Y., and Chen, R.-K. (2020). A novel variation in the FRIZZLE PANICLE (FZP) gene promoter improves grain number and yield in rice. *Genetics* 215, 243–252. doi:10.1534/genetics.119.302862

Wilkinson, P. A., Winfield, M. O., Barker, G. L. A., Tyrrell, S., Bian, X., Allen, A. M., et al. (2016). CerealsDB 3.0: Expansion of resources and data integration. *BMC Bioinforma.* 17, 256. doi:10.1186/s12859-016-1139-x

Wu, Q., Chen, Y., Xie, J., Dong, L., Wang, Z., Lu, P., et al. (2021). A 36 Mb terminal deletion of chromosome 2BL is responsible for a wheat semi-dwarf mutation. *Crop J.* 9, 873–881. doi:10.1016/j.cj.2020.06.015

Xie, W., and Nevo, E. (2008). Wild emmer: Genetic resources, gene mapping and potential for wheat improvement. *Euphytica* 164, 603–614. doi:10.1007/s10681-008-9703-8

Zanke, C. D., Ling, J., Plieske, J., Kollers, S., Ebmeyer, E., Korzun, V., et al. (2014). Whole genome association mapping of plant height in winter wheat (*Triticum aestivum* L.). *PLoS One* 9, e113287. doi:10.1371/journal.pone.0113287

Zhao, C., Liu, B., Piao, S., Wang, X., Lobell, D. B., Huang, Y., et al. (2017). Temperature increase reduces global yields of major crops in four independent estimates. *Proc. Natl. Acad. Sci. U. S. A.* 114, 9326–9331. doi:10.1073/pnas.1701762114

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 11 November 2022
DOI 10.3389/fgene.2022.984720

Check for updates

OPEN ACCESS

EDITED BY
Vijay Gahlaut,
Institute of Himalayan Bioresource Technology (CSIR), India

REVIEWED BY
Amit Kumar Singh,
National Bureau of Plant Genetic Resources (ICAR), India
Rahul Kumar,
Chaudhary Charan Singh University, India

*CORRESPONDENCE
Palaparthi Dharmateja,
dharmateja.palaparthi@gmail.com
Rajbir Yadav,
rajbiryadav@yahoo.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Plant Genomics, a section of the journal Frontiers in Genetics

RECEIVED 02 July 2022
ACCEPTED 17 October 2022
PUBLISHED 11 November 2022

CITATION
Dharmateja P, Yadav R, Kumar M, Babu P, Jain N, Mandal PK, Pandey R, Shrivastava M, Gaikwad KB, Bainsla NK, Tomar V, Sugumar S, Saifi N and Ranjan R (2022), Genome-wide association studies reveal putative QTLs for physiological traits under contrasting phosphorous conditions in wheat (Triticum aestivum L.).
Front. Genet. 13:984720.
doi: 10.3389/fgene.2022.984720

COPYRIGHT
© 2022 Dharmateja, Yadav, Kumar, Babu, Jain, Mandal, Pandey, Shrivastava, Gaikwad, Bainsla, Tomar, Sugumar, Saifi and Ranjan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Genome-wide association studies reveal putative QTLs for physiological traits under contrasting phosphorous conditions in wheat (*Triticum aestivum* L.)

Palaparthi Dharmateja[1*†], Rajbir Yadav[1*†], Manjeet Kumar[1], Prashanth Babu[1], Neelu Jain[1], Pranab Kumar Mandal[2], Rakesh Pandey[3], Manoj Shrivastava[4], Kiran B. Gaikwad[1], Naresh Kumar Bainsla[1], Vipin Tomar[5], S. Sugumar[1], Nasreen Saifi[1] and Rumesh Ranjan[1]

[1]Division of Genetics, ICAR-Indian Agricultural Research Institute, New Delhi, India, [2]ICAR- National Institute for Plant Biotechnology, New Delhi, India, [3]Division of Plant Physiology, ICAR-Indian Agricultural Research Institute, New Delhi, India, [4]Division of Environment Science, ICAR-Indian Agricultural Research Institute, New Delhi, India, [5]Centre for Quantitative Genetics and Genomics (QGG), Aarhus University, Aarhus, Denmark

A Genome-wide association (GWAS) study was conducted for phosphorous (P)-use responsive physiological traits in bread wheat at the seedling stage under contrasting P regimes. A panel of 158 diverse advanced breeding lines and released varieties, and a set of 10,800 filtered single nucleotide polymorphism (SNP) markers were used to study marker-trait associations over the eight shoot traits. Principle component analysis separated the two environments (P regimes) because of the differential response of the traits indicating the essentiality of the separate breeding programmes for each environment. Significant variations for genotypic, environmental, and genotype × environment (GEI) effects were observed for all the traits in the combined analysis of variance with moderately high broad sense heritability traits (0.50–0.73). With the different algorithms of association mapping viz., BLINK, FarmCPU, and MLM, 38 unique QTLs under non-limiting P (NLP) and 45 QTLs for limiting P (LP) conditions for various shoot traits were identified. Some of these QTLs were captured by all three algorithms. Interestingly, a *Q.iari.dt.sdw.1* on chromosome 1D was found to explain the significant variations in three important physiological traits under non-limiting phosphorus (NLP) conditions. We identified the putative candidate genes for QTLs namely *Q.iari.dt.chl.1, Q.iari.dt.sdw.16, Q.iari.dt.sdw.9 and Q.iari.dt.tpc.1* which are potentially involved in the mechanism regulating phosphorus use efficiency through improved P absorption due to improved root architectural traits and better mobilization such as sulfotransferase involved in postembryonic root development, WALLS ARE THIN1 (WAT1), a plant-specific protein that facilitates auxin export; lectin receptor-like kinase

essentially involved in plant development, stress response during germination and lateral root development and F-box component of the SKP-Cullin-F box E3 ubiquitin ligase complex and strigolactone signal perception. Expression profiling of putative genes located in identified genomic regions against the wheat expression atlas revealed their significance based on the expression of these genes for stress response and growth development processes in wheat. Our results thus provide an important insight into understanding the genetic basis for improving PUE under phosphorus stress conditions and can shape the future breeding programme by developing and integrating molecular markers for these difficult-to-score important traits.

KEYWORDS

wheat, GWAS, single nucleotide polymorphism, PUE, non-limiting phosphorus and limiting phosphorus

Introduction

Wheat (*Triticum aestivum* L.), a major cereal crop, meets one-fourth of the protein and one-fifth of the calorie requirement of the total human diet, worldwide. Considering the continuously increasing demand for wheat, its production needs to be increased by at least 50 *percent* by 2050. Additionally, wheat scientists have to look for avenues to increase production under the constraints of declining natural resources and changing climatic conditions (Yadav et al., 2011; Gupta et al., 2016). Globally, the sharp gain in wheat yield was realized with the development and cultivation of lodging tolerant and fertilizer-responsive semi-dwarf wheat varieties during the green revolution era. Since then, yield gain rates are declining in many parts of the world (Yadav et al., 2010). In one of our recent assessments of varieties released over a century-long period (since 1905) for India's north-western plain zone, the wheat yield has grown at the rate of 0.544% (Yadav et al., 2021).

Globally phosphorus (P), being the essential macronutrient, is among the various major yield-deciding factor for increased crop productivity. P, besides being an essential constituent of genetic material, plays a vital role in photosynthesis, energy transfer, improved root development, starch and sugar transformation, and nutrient flow within the plant system (Pasek, 2008). P-limiting conditions have been reported to be causing stunted growth, reduced effective tillering, thin stems, enhanced root:shoot ratio, and substantial yield reduction in rice and many other cereals (Dobermann and Fairhurst, 2000). At early stages, P deficiency affects the emergence and growth of seedlings (Singh et al., 2013) which in turn leads to decreased root volume, total leaf area, and plant dry weight; however, there is also a significant increase in the density of root hairs and root to shoot ratio (Ahmed et al., 2018). Hence, P feeding during the early stages of wheat growth is crucial for better establishment and any supplementation at later stages would not be compensated in plant growth, which would cause a significant decrease in tiller development and head formation (Grant et al., 2001). Early seedling variations correlate well with the high P uptake. A deeper understanding of the biomass and optimized root traits like root length, root width, root tips number, root diameter, root biomass, and shoot biomass can throw some light on dissecting mechanisms underlying PUE. Hence, exploring the variations for seedlings and their initial biomass accumulation traits under limiting and non-limiting P conditions would provide a better scope for improving elite lines for PUE (YANG et al., 2021). In wheat, the realized yield is also significantly compromised under P deficit conditions worldwide (Zhang et al., 2005). Due to its essential role in physiological activity the high energy requirement of modern high-yielding varieties and the limitation of naturally available P in soil, inorganic P fertilizers are much needed in great quantities in contemporary agriculture. It is anticipated that 52.9 MT of P fertilizers will be used in agriculture by 2030 (Brears, 2015). Nevertheless, large-scale injudicious fertilization raises environmental concerns with unused chemical fertilizers seeping into water systems resulting in eutrophication and affecting marine life adversely (Cassman et al., 2003; Chiou and Lin, 2011). Phosphate rock is a limited and non-renewable global resource with a very short expected exhaustion period (50–200 years) at the current pace of P utilization (Herrera-Estrella and López-Arredondo, 2016), it is high time for its judicious application in agriculture.

Among the various management practices for the rational use of P, the development of P use efficient crop cultivars which come with no additional costs (Heuer et al., 2017) is the most economical. Moreover, for wheat, with the largest area and comparatively lower P utilization rate (10.7%) than rice (13.1%) and maize (11%) (Ma et al., 2011), focuses on the fast-track approaches to identify P efficient cultivars is very essential (Dharmateja et al., 2021). Productivity gains with low P-application by growing P-efficient cultivars would pave the way to meet future global food requirements in an eco-friendly, economically feasible, and socially sustainable manner. Moreover, compared to modern wheat varieties which are bred under optimal P conditions, landraces and traditional local germplasm have greater PUE under unfertilized and

P-fixed soil due to their natural selection in the evolutionary process and adaptation to P-deficient conditions (Wissuwa and Ae, 2001). The use of such genetic variability would help to improve modern cultivars for these traits.

To develop P-efficient wheat varieties, knowledge about fundamental traits and their associated genomic regions controlling P uptake and utilisation is of utmost importance. The advances in genomics in wheat over the last decade enabled breeders to identify the genetic markers liked to traits of interest and to facilitate integrating them into the breeding lines for various difficult-to-measure target traits. High throughput genotyping arrays have revolutionized marker-trait-associations studies and have helped to fine-map the target genomic regions. PUE-related traits are largely governed by several minor genes with small cumulative effects. Previously, QTLs for P-deficiency-tolerance have been identified through linkage mapping in biparental wheat populations (Yuan et al., 2017). But, considering large minor genes affecting PUE and its complexity association mapping by exploring more than two possible alleles at various genomic regions in a genetic panel consisting of lines with a diverse pedigree will have more probability of capturing factors explaining the phenotypic variation for P use efficiency than the bi-parental mapping population. Accordingly, Genome-wide association studies (GWAS) have effectively explored allelic variations for P-deficiency tolerance in *Aegilops tauschii* (Liu et al., 2015), Arabidopsis (Bouain, Doumas, and Rouached, 2016) and soybean (Ning et al., 2016).

When measuring a trait response of P among genotypes, precise phenotyping becomes the major limiting factor. Many researchers have struggled to precisely quantify the impact of P on plant growth modifications. Testing genotypes against a particular concentration of measured P is not a feasible task under field conditions. This approach has been limited due to the uncontrollable natural contribution of P through inorganic supply from manures and microbial interaction, and moreover, it is difficult to maintain a uniform supply of P throughout the growing period in the soil. To overcome this limitation and to measure genotypic responses to sole P concentration variations, hydroponic systems are being used in many crops to accurately quantify the P and its uptake by plants. Due to their subterranean nature, these systems have been considered to be the best way to study the root system modifications which would also permit for precise estimation of root and shoot parameters. The simplified hydroponics systems based on aerated nutrient solutions with various levels of P with an option for complete replacement of nutrient solutions at fixed intervals were reported (Byrne et al., 2011; Gong et al., 2011) and successfully used in PUE studies. Hydroponic systems can also prove to be an efficient approach for screening large populations under different nutrient conditions and less amount of space.

In recent decades, several quantitative trait loci (QTLs) for PUE and related traits have been identified and mapped on all 21 chromosomes in wheat under hydroponic culture trials (Guo et al., 2012; Zhang and Wang, 2015). In addition, the phosphorus uptake (PupE), the utilization efficiency (PutE), and the response of morphological traits under different P levels were investigated in hydroponic culture (Yuan et al., 2017b). Categorization of wheat germplasm for PUE was done based on P use efficiency parameters, recorded on genotypes grown in hydroponics with two P regimes (Bilal et al., 2018). PUE-enhanced crop cultivars can have more growth and biomass for the same quantity of P taken up at a given time (Rengel and Marschner, 2005). However, having low heritability for this trait, breeding for PUE is complex and influenced by multiple physiological processes, besides enormous environmental impacts (Heuer et al., 2017). Therefore, dissecting its inheritance patterns into underlying genetic factors at the genomic level can circumvent these environmental influences. With this background, in our current study, we performed a GWAS on a set of high-yielding genotypes with diverse pedigree/parentage to identify marker-trait associations (MTAs) for P use efficiency with a 35 K breeders Affymetrix SNPs array for genotyping. More interestingly, this study has been planned on a panel of indigenously bred, high-yielding advanced breeding lines so that the subsequent gain in PUE can be built upon among these elite lines without compromising yield gain. In addition, the insight gained in understanding P use efficiency among selected genotypes and identified QTNs (Quantitative traits nucleotides) governing PUE will pave the way to develop better genotypes for this challenging-to-score and complex-to-study trait in wheat.

Materials and methods

Plant materials and hydroponic culture

A set of 158 advanced bread wheat lines of spring type with great diversity in their pedigree and lineage were grown under non-limited and limited phosphorus conditions in hydroponics. This set also included released varieties with their wider adaptation to different agro-climatic conditions (Supplementary Table S1). The phenotypic assay under hydroponics was conducted at National Phytotron Facility at the ICAR-Indian Agricultural Research Institute, New Delhi, India. Individual genotypes were grown under a controlled environment with typical growing parameters of 12°C–22°C day/night temperature, 10 h day length of photoperiod, and 70% relative humidity were maintained throughout the growing period. The seeds of individual genotypes were germinated in a separate petri dish with blotting paper containing sufficient moisture for 5–6 days. The uniform-sized five-day-old seedlings were transferred to the hydroponics tank

system consisting of a tray of 18-litre capacity covered with a black ceramic lid. The basal nutrient solution used in the hydroponics experiment consisted of $(NH_4)_2SO_4{\cdot}H_2O$ (1 mmol/L), $Ca(NO_3)_2{\cdot}4H_2O$ (1 mmol/L), KCl (1.8 mmol/L), $MgSO_4{\cdot}7H_2O$ (0.5 mmol/L), $CaCl_2$ (1.5 mmol/L), H_3BO_3 (1 μmol/L), $CuSO_4{\cdot}5H_2O$ (0.5 μmol/L), $ZnSO_4{\cdot}7H_2O$ (1 μmol/L), $MnSO_4{\cdot}H_2O$ (1 μmol/L), FeEDTA (100 μmol/L), and $(NH_4)_6Mo_7O_{24}{\cdot}4H_2O$ (0.1 μmol/L), and two levels of P were maintained using KH_2PO_4 as Non limiting P (0.2 mmol/L) and limiting P (0.02 mmol/L) (Yuan et al., 2017; Bilal et al., 2018). The nutrient solution was continuously aerated, and the pH was maintained between 6 and 6.5 using 1 M KOH and 1 M HCL. The consistent nutrient provision is supported by basal solution supplanted with fresh solution every 4 days.

Advanced phenotyping for phosphorous-responsive traits

The experimental material was tested in the completely randomized design (CRD) design with three replications, and five plants of thirty-day-old seedlings (Zadok's scale: growth stage 29) (Zadoks, Chang, and Konzak, 1974) from each replication under NLP and LP hydroponic conditions were taken out for recording the observations on traits under study. The root and shoots of the individual plant were separated carefully using scissors. LI-COR 3000 (Lincoln, NE, United States) leaf area meter was used to measure the total leaf area (TLA, m^2/plant). The Chl content (μmmol/m^2) was measured with SPAD (Minolta Camera Co., Osaka, Japan) (Minolta 1989). The SPAD readings are measured based on the transmission of red light (650 nm) and infrared light (940 nm), in which red light is absorbed by chlorophyll (Xiong et al., 2015). The Chl data with SPAD was recorded on a fully-opened, well-developed topmost leaf of 30 days old seedlings in each replication on five plants with an average of three readings between 11 a.m. and 12:30 p.m. The variability in SPAD reading is influenced by environmental factors, diurnal variation, and genotypic differences in Chl content under treatment. Therefore, we tried to keep the SPAD reading variability to the minimum level by taking the observations under controlled conditions and with a minimum time difference in data recording. The separated plant parts, i.e., shoot and roots, were dried in a hot air oven at 60°C until stable dry biomass was obtained, then measured for shoot dry weight (SDW, g/plant) and root dry weight (RDW, g/plant). The RSR was calculated as shoot-dry weight (SDW) to root-dry weight (RDW) ratio. The combined shoot and root-dry weight of samples were considered as total dry weight (TDW, g/plant). To estimate the P content, the fine-grind and dried sample of each plant was digested in a diacid mixture (HNO_3: $HCLO_4$) until a clear solution was obtained. The total P (mg) was analyzed using the vanadium molybdate yellow colourimetric method (Ma et al., 2021a). The tissue phosphorus content multiplied by total dry weight is used to calculate the total P uptake by the plant (TPU, mg) (Wang et al., 2017). The P utilization efficiency (PUtE, dry weight (g)/P (mg)) was calculated using the following formula under both NLP and LP conditions (Wang et al., 2017).

PUtE [dry weight (g)/P (mg)] = Total dry weight/total P uptake by plant.

Phenotypic data analysis

The presence of outliers in the data was confirmed by boxplot analysis and with the Z score test. The phenotypic data generated through trials in a completely randomised design were subjected to the analysis of variance (ANOVA) with STAR version 2.1.0 (Statistical Tool for Agricultural Research) an R-based software (Gulles et al., 2014). The variability analysis and adjusted means were calculated as best linear unbiased predictions (BLUPs) considering the replication and genotypes as random effects separately for each P regimes using META-R version 6.0 (Alvarado et al., 2020). The adjusted means of replicates in the trial were obtained by fitting mixed linear models (MLM) using the equation.

$$Y_{ik} = \mu + R_i + G_k + \epsilon_{ik}$$

where Y_{ik} is the trait of interest; μ is the mean effect; R_i is the effect of the ith replicate; G_k is the effect of the kth genotype; ϵ_{ik} is the error associated with the ith replication and the kth genotype, which is assumed to be normally and independently distributed, with mean zero and homoscedastic variance σ^2. The genotypes and replicates were considered random effects to calculate adjusted means, while P regimes as a fixed effect. The broad-sense heritability was estimated using the formula.

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_e^2 / nreps}$$

Where σ_g^2 is the genotypic variance; σ_e^2 is the error variance and $nreps$ is the number of replications. The broad-sense heritability estimated the quality of the breeding program for the traits and the environments. The LSD with type I error, $\alpha = 0.05$ of the level of significance, was calculated using the formula:

$$LSD = t_{(1-0.05, df error)} \times ASED$$

Where t is the cumulative "Student's t-test" distribution; $df\ error$ is the degrees of freedom for the variance of error, and ASED is the average standard error of the differences between pairs of means. And the coefficient of variation (%) is calculated using the formula:

$$CV\,(\%) = 100 \times \frac{ASED}{grand\ mean}$$

The phosphorus deficiency tolerance coefficient (PDTC) was calculated as the ratio of LP over NLP treatment (Li et al., 2015). The correlations were calculated as simple pairwise Pearson's correlations among traits. Further, principal component analysis (PCA) was performed to identify the number of principal components required to explain the variation across the environments and to study the relationship between traits and the effect of treatments using R package version 4.0.1 (R core Team, 2022).

Genome-wide studies to establish marker traits associations

Genomic DNA was extracted by following the CTAB method (Doyle and Doyle, 1987) and the quantity and quality of DNA were estimated by using an agarose gel electrophoresis approach and UV spectrophotometer. The 35 K Axiom® Wheat Breeder's Array (Affymetrix UK Ltd., United Kingdom) was used for genome-wide association analysis. Monomorphic markers, markers with >30% missing data, 5% minor allele frequency, and greater than 20% heterozygosity were removed. A filtered set of 10,800 highly informative SNP markers was used for GWAS. Population structure among the 158 genotypes was determined by using STRUCTURE v. 2.3.4 (Pritchard, Stephens, and Donnelly, 2000). Population structure was generated the over 10,000 length of burnin period and 100,000 MCMC reps with three iterations. The optimum population number (K) value was determined by the *ad-hoc*, delta K method (Evanno et al., 2005). Constellation plot plotted by using the Ward method in JMP v.14 (Lehman et al., 2005). Linkage disequilibrium (LD) analysis was performed on a sub-population basis, using the LD function in TASSEL (Bradbury et al., 2007) and draw the LD plot with R software version 4.2.1 (R core Team 2022). Associations between genotypic and phenotypic data were evaluated in GAPIT version 3.0 (Wang et al., 2021) using MLM, BLINK and FarmCPU algorithms. The Mixed linear model (MLM) is carried out by taking into account both the Q-matrix and the K-matrix (the kinship matrix expressing family relatedness among the genotypes) (Yu and Buckler, 2006; Tadesse et al., 2015). The FarmCPU makes use of stepwise regression [fixed-effect model (FEM)] and the mixed linear model (MLM) (X. Liu et al., 2016). The FEM is used to evaluate genetic markers, while the REM is used to control false positives by including associations or pseudo-quantitative trait nucleotides as covariates in the model. The Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway (BLINK) is very computationally efficient and reliable, and algorithum is based on the fixed effect model involving the Bayesian information criteria and linkage disequilibrium information (Huang et al., 2019).

The kinship matrix was calculated from the 10,800 markers, and QQ and Manhattan plots were generated to evaluate the results. The adjusted *p*-value threshold of significance was corrected for multiple comparisons according to the false-discovery rate (FDR) with cut-off ≤ 0.05 with Benjamini and Hochberg method (https://tools.carbocation.com/FDR) (Benjamini and Hochberg, 1995). After the identification of MTAs, an *in silico* search of the putative candidate genes with their annotated functions was conducted in the Ensembl Plants database (http://plants.ensembl.org/index.html) of the bread wheat genome (*Triticum aestivum* L.). The Chinese spring wheat cultivar was utilized in making of genome assembly IWGSC-refseq version 1.0. The data from CerealDB was utilized to locate the genes in the wheat version hosted on Ensembl. The gene expression atlas was used to analyse the expression of detected putative candidate genes (http://www.wheat-expression.com/) (Borrill et al., 2016).

Results

Phenotypic variation and heritability for traits under non-limiting phosphorus and limiting phosphorus conditions

Phenotypic data on P-responsive traits among the 158 wheat genotypes were recorded under NLP and LP conditions. The combined analysis of variance indicated significant variations due to the genotype (G), different P-treatment (T) regimes (NLP and LP), and genotype-P treatment interaction effect (G × T) for all the traits. The broad-sense heritability expressed as the proportion of total phenotypic variance for each trait under this study was moderate, ranging from 0.50 (PUtE) to 0.73 (TDW), indicating the involvement of both genetic and environmental variations in governing these traits related to P-response (Table 1). The boxplots of various traits under study have been presented in (Figure 1), with mean values as "*". The traits namely TLA, SDW, TDW, TPU, and TPC were higher in NLP than LP, while RSR, Chl content, and PUtE were higher in LP. The spread of variability for all traits under study was higher in NLP except for PUtE, which was limited. The overall coefficient of variation was ranging from 3.10% to 13.70% across the traits tested. Limiting the phosphorous supply reduced TLA, SDW, TDW, TPC, and TPU. TLA ranged from 18.94 to 97.55 m^2/plant in NLP and 13.32–55.65 m^2/plant in LP while SDW ranged from 0.114 to 0.451 g/plant for NLP and 0.067–0.214 g/plant in LP. TDW ranged from 0.139 to 0.545 g/plant under NLP condition against 0.111–0.362 g/plant in LP. A higher value was observed for TPC and TPU in a P-rich environment with a range of 3.07–9.33 against 0.499–2.54 for LP. TPU varied from 0.541 to 4.075 under NLP and 0.092–0.647 in LP. In contrast, under LP, increased values were observed for Chl, RSR, and PUtE. Chl varied from 25.14 to 40.66 μm mol/m^2 for NLP and 23.42–48.06 μm mol/m^2 for LP, with a mean value higher in LP than NLP (Table 2). Similarly, a higher ratio was observed for RSR in LP across all

TABLE 1 Analysis of variance and heritability for the traits under non-limiting and limiting phosphorus.

	Source of variation	TLA	Chl	SDW	RSR	TDW	TPC	TPU	PUtE
MSS	Phosphorus (P)	112,084.50**	6138.24**	2.47**	21.04**	1.93**	5089.56**	490.76**	99.68**
	Genotype (G)	515.10**	57.78**	0.0132**	0.0520**	0.026**	2.892**	0.9996**	0.1623**
	G*P	414.11**	28.44**	0.0074**	0.0328**	0.0093**	2.60**	0.8431**	0.1628**
	Error	26.89	1.42	0	0.0005	0.0001	0.0096	0.0019	0.0014
V_G		122.05	14.09	0.0033	0.0129	0.0064	0.7206	0.2494	4.02E-02
V_G P_H		193.61	13.51	0.0037	0.0162	0.0046	1.29	0.4206	8.07E-02
V_E		26.89	1.42	0	0.0005	0.0001	0.0096	0.0019	1.40E-03
V_P		225.58	21.20	0.0052	0.0211	8.8E-03	1.37	0.4602	0.0809
Grand mean		37.85	35.86	0.1751	0.3749	0.233	3.71	0.9807	0.4982
CV (%)		13.70	3.32	3.53	5.92	3.1	2.64	4.43	7.6
LSD		7.27	3.07	0.0458	0.0887	0.0622	0.4138	0.2929	4.23E-06
Heritability		0.54	0.66	0.64	0.61	0.73	0.53	0.54	0.50

TLA, total leaf area; Chl, chlorophyll content; SDW, shoot dry weight; RSR, root:shoot ratio; TDW, total dry weight; TPC, total phosphorus content; TPU, total phosphorus uptake; PUtE, phosphorus utilization efficiency; V_G, genotypic variance; V_E, error variance; V_P, total phenotypic variance; CV, coefficient of variation; LSD, Least Significant Difference.
*$p < 0.05$; **$p < 0.01$ and ***$p < 0.001$ include significance in tables.

genotypes tested, with a range of 0.287–1.03 in LP. The PUtE was found higher under limiting phosphorus with a range of 0.396–2.02 against 0.107–0.326 under NLP. The phosphorus deficiency tolerance coefficient (PDTC) was calculated as the ratio of different trait mean values in LP over NLP for a better and quicker understanding of trait response under two conditions. The traits like Chl (1.15), RSR (2.33), and PUtE (4.73) with >1 PDTC value indicated their better expression under LP. The remaining traits TLA (0.556), SDW (0.550), TDW (0.677), TPC (0.231), TPU (0.154) expressed better with availability of sufficient phosphorus as their PDTC value was < 1.

Trait correlations and principal component analysis

Pearson's pairwise correlation coefficient was calculated among various traits under both environments. Chl was not associated with any other traits, neither in NLP nor in LP. The most significant ($p < 0.001$) correlations were observed between TLA and SDW under both conditions (Figures 2A,B). Interestingly, RSR was positively correlated with TLA under NLP and negatively under LP conditions. Very strong positive correlations were observed for TLA with SDW, TDW and TPU; SDW with TDW and TPU and TPU with TDW and TPC under both NLP and LP conditions indicating their strong influence on each other. The TDW showed a significant positive correlation with TPC in the NLP condition, and a non-significant correlation was observed in the LP condition. TPC had a significant positive correlation with TPU, and it was interesting to note that both of these showed a strong negative association with PUtE under both environments. In addition, the PUtE was also negatively associated with most of the traits in NLP and LP conditions except for a significant positive association with RSR in LP conditions. PCA-based grouping of traits over LP and NLP conditions indicated that the first two principal components (PC1 and PC2) had explained 84.5% (68.9% and 15.6%) of the total variation (Figure 3). The existence of a high G x T effect indicates that these treatments significantly affected the traits studied. PCA analysis showed that TLA was highly dependent on SDW, TDW, TPC, and TPU and least dependent on Chl, RSR, and PUtE. Similarly, Chl was dependent on RSR and, PUtE, with the least dependence on SDW, TDW, TPC, and TLA. The results clearly depicted in the figure that the trait expression varies with the availability of P. The genotypes under NLP condition exhibit better expression for TLA, SDW, TDW, TPC, and TPU, and in the LP condition exhibit higher expression for Chl, RSR, and PUtE. Along with the treatment, genotypes are separated and all eight traits fall into two distinct groups. The correlation between these traits is represented by the angle between their vectors. The traits TLA, SDW, TDW, TPC, TPU and Chl, RSR, and PUtE are highly correlated because the angle between them is very small (acute angle i.e., < 90°). But there is a significant crossover between the traits of these distinct groups because the presence of a wider angle (i.e., > 90°).

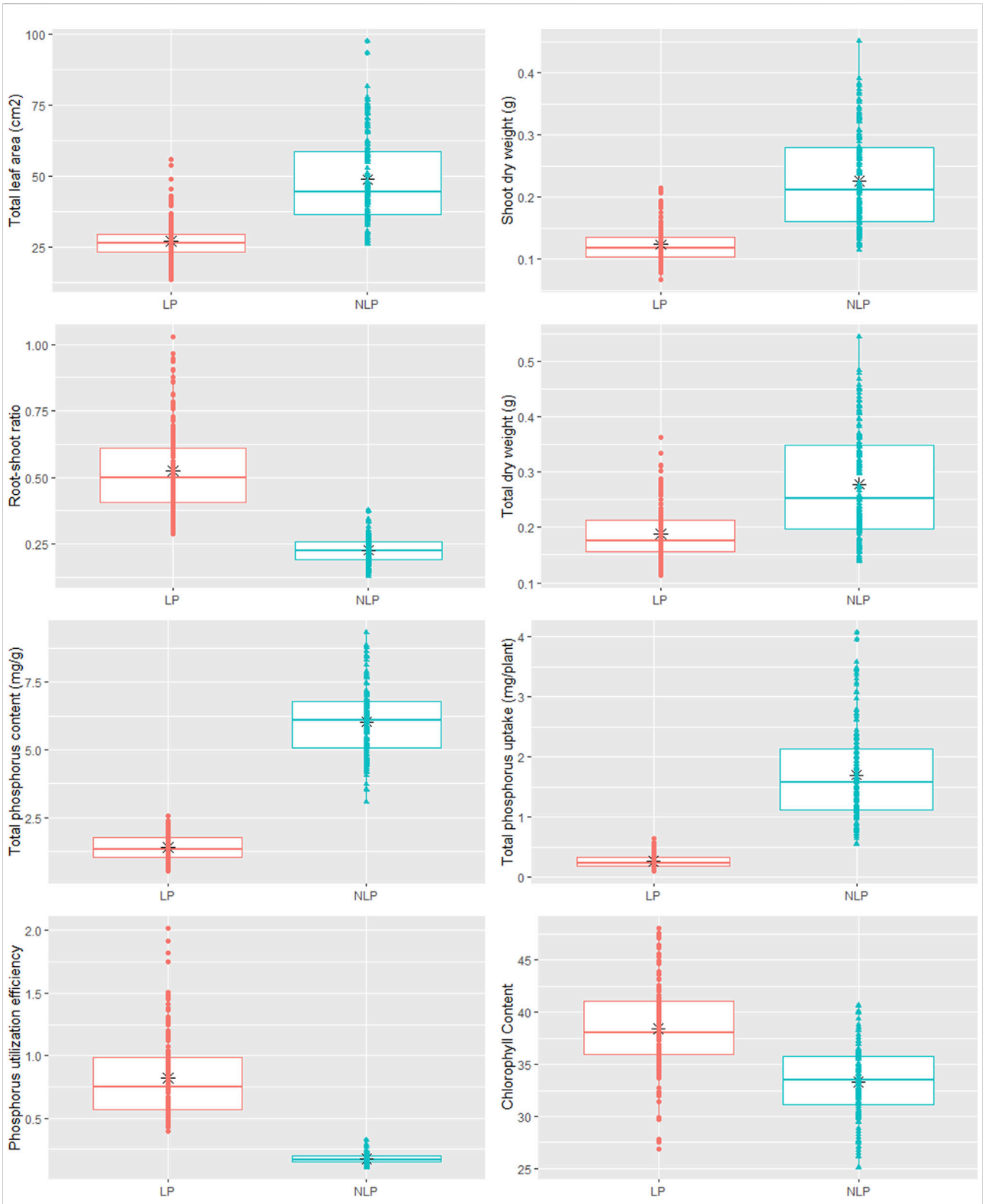


FIGURE 1
Box plot showing phenotypic variation in wheat genotypes in non-limiting and limiting phosphorus conditions.

TABLE 2 Measures of variability and phosphorus deficiency tolerance coefficient (PDTC) for the traits under study.

Trait	NLP		LP		Mean		PDTC
	MIN	MAX	MIN	MAX	NLP	LP	
TLA	18.94	97.55	13.32	55.65	48.48	26.95	0.556
Chl	25.14	40.66	23.42	48.06	33.31	38.37	1.15
SDW	0.114	0.451	0.067	0.214	0.226	0.124	0.550
RSR	0.122	0.377	0.287	1.03	0.225	0.523	2.33
TDW	0.139	0.545	0.111	0.362	0.278	0.188	0.677
TPC	3.07	9.33	0.499	2.54	6.02	1.39	0.231
TPU	0.541	4.075	0.092	0.647	1.70	0.261	0.154
PUtE	0.107	0.326	0.396	2.02	0.174	0.822	4.73

TLA, total leaf area; Chl, chlorophyll content; SDW, shoot dry weight; RSR, root:shoot ratio; TDW, total dry weight; TPC, total phosphorus content; TPU, total phosphorus uptake; PUtE, phosphorus utilization efficiency; MIN, minimum value; MAX, maximum value; PDTC, phosphorus deficiency tolerance coefficient.

Genome-wide association mapping of phosphorous-responsive physiological traits

The association analysis was performed using three algorithms (Blink, FarmCPU, and MLM) for the physiological traits studied under NLP and LP conditions. A total of 10,800 SNP's derived from the SNP chip array for 158 genotypes were used for mapping and their distribution across the chromosomes was depicted (Supplementary Table S2). Among association mapping panel population structure was calculated by using STRUCTUREv2.3.4 software (Figure 4A). The population structure showed a sharp peak at K = 2 when the clusters were plotted against ΔK indicating two subgroups in the population and the dendrogram, constellation plot revealed two major groups, thus results were further confirmed (Figures 4B,C). A set of 10,800 high-quality SNP markers were distributed across the genome with the B genome (4041) having the highest number of markers, followed by the A genome (3409) and the D genome (3350) respectively. According to chromosome-wise distribution, chromosome 2B (749) had the most markers mapped, followed by chromosome 2D (707). The least number of markers were found on chromosomes 4D (180) and 6B (277). The LD was estimated by calculating the squared correlation coefficient (*r2*) for all the 10,800 markers. The genome-wide LD decay with physical distance, the LD decay to its half at 9.04 Mb for the whole genome (Figure 5). The –log *p*-value = 3.5 is considered as threshold to call the MTAs as significant associations and corrected according to the false-discovery rate (FDR cut-off ≤ 0.05). A total of 83 QTLs were detected, among them 45 in NLP and 38 in LP conditions identified for the eight traits under study. With different algorithms under NLP conditions, 19 (BLINK), 32 (Farm CPU), 9 (MLM) QTLs were identified, and similarly, under LP conditions, 26 (BLINK), 34 (FARM CPU), and 21 (MLM) QTLs were detected (Supplementary Table S3) (Figures 6A,B; Supplementary Figures S1A–D). Fourteen QTLs (4 in NLP and 10 in LP) were detected in all the algorithms (Supplementary Table S4). Four QTLs were detected for TLA, located one each on chromosome 7B, 1D, 6A, and 1B across both the conditions. In NLP condition *Q.iari.dt.tla.1* and *Q.iari.dt.tla.2* were detected on chromosome 7B and 1D with their associated SNPs AX94470386 and AX94765690 with –log10 *p*-value with 3.70 and 3.5 respectively. Two QTLs namely *Q.iari.dt.tla.3* and *Q.iari.dt.tla.4* with their associated SNPs as AX94770913 and AX95190390 with –log10 *p*-value ranging from 4.2 to 4.6 and 4.0 to 4.3, respectively, were detected in LP conditions on chromosome 6A and 1B. For Chl content, three QTLs, namely *Q.iari.dt.chl.1*, *Q.iari.dt.chl.2*, and *Q.iari.dt.chl.3* associated with SNPs AX94832883, AX94676652, and AX95105278 were detected in NLP conditions on chromosomes 7D, 2A, and 2B. Under LP condition, *Q.iari.dt.chl.4, Q.iari.dt.chl.5, Q.iari.dt.chl.6, Q.iari.dt.chl.7,* and *Q.iari.dt.chl.8* were detected on chromosomes 6B, 6A, 6D, 2D, and 6D for Chl content with their associated SNPs AX94597699, AX95241386, AX94702861, AX94622481, and AX95230097 with –log10 *p*-value ranging from 3.5 to 4.8. Seven QTLs under NLP and thirteen QTLs under LP were detected for SDW with a range of –log10 *p*-value from 3.74 to 8.13. For RSR, thirteen QTLs in NLP and seven QTLs in LP conditions were detected. SDW is very strongly associated with TDW, seven QTLs in the NLP condition and twelve QTLs in the LP condition were detected to explain the variation for TDW. For TPC five QTLs, two each on chromosomes, 3A and 3 D and one on 7A were identified under both the treatments. Under NLP conditions, three QTLs and in LP four QTLs were detected for the trait TPU. However, only two QTLs were found to explain the variation for PUtE in both NLP and LP conditions.

Interestingly, a QTL *Q.iari.dt.sdw.1* associated with SNP AX94514240 located on chromosome 1D was found to explain the significant variation for all three traits, namely SDW, TDW, and TPU under NLP conditions (Table 3). Though SDW is part of TDW, only four common QTLs viz. *Q.iari.dt.sdw.2* associated with SNP AX94646448 on 1B; *Q.iari.dt.sdw.8* and *Q.iari.dt.sdw.14* on 7B associated with SNPs AX94638774 and AX94626370, under NLP condition, and a QTL *Q.iari.dt.sdw.20* associated with SNP AX94456805 detected on chromosome 2D under LP condition explained the variation for both traits. The traits like SDW and TPU shared a common QTL named *Q.iari.dt.sdw.3* present on chromosome 7B and associated with SNP AX94396598 was detected to explain the variation for traits SDW and TPU, simultaneously under NLP condition only. Another QTL named *Q.iari.dt.rsr.1* on chromosome 6A

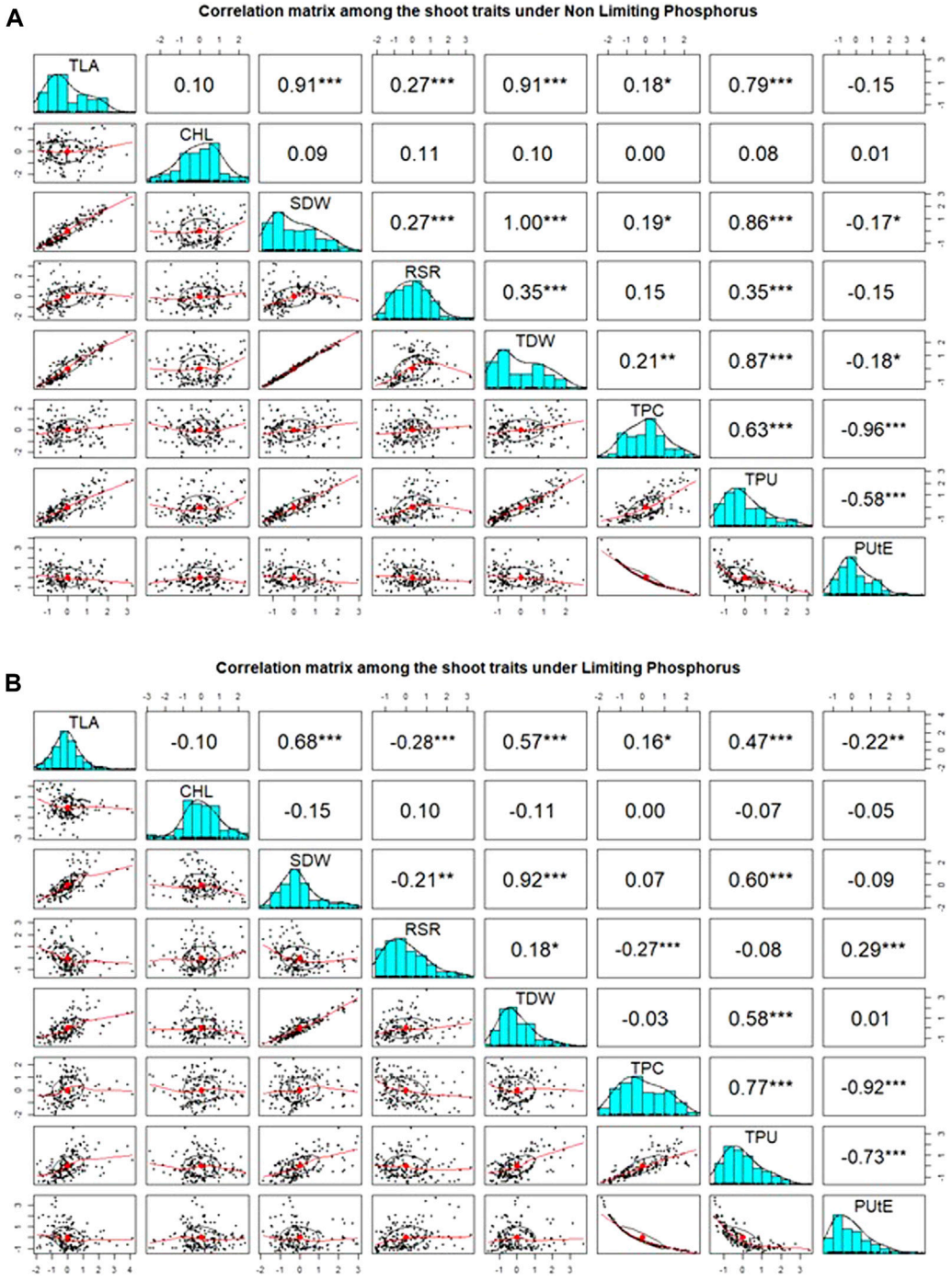


FIGURE 2
Association between measured traits and distribution among wheat genotypes under non-limiting (NLP) **(A)** and limiting (LP) phosphorus **(B)** conditions. TLA, total leaf area; Chl, chlorophyll content; SDW, shoot dry weight; RSR, root:shoot ratio; TDW, total dry weight; TPC, total phosphorus content; TPU, total phosphorus uptake; PUtE, phosphorus utilization efficiency.

and associated with SNP AX94475513 was associated with two traits, RSR and TDW under both conditions. In the NLP condition, TPC shared common QTLs named *Q.iari.dt.tpc.2* and *Q.iari.dt.tpc.3* which were associated with SNPs AX94584110 and AX94397869 on the chromosome 3A and 3D with the other traits PUtE and TPU, respectively.

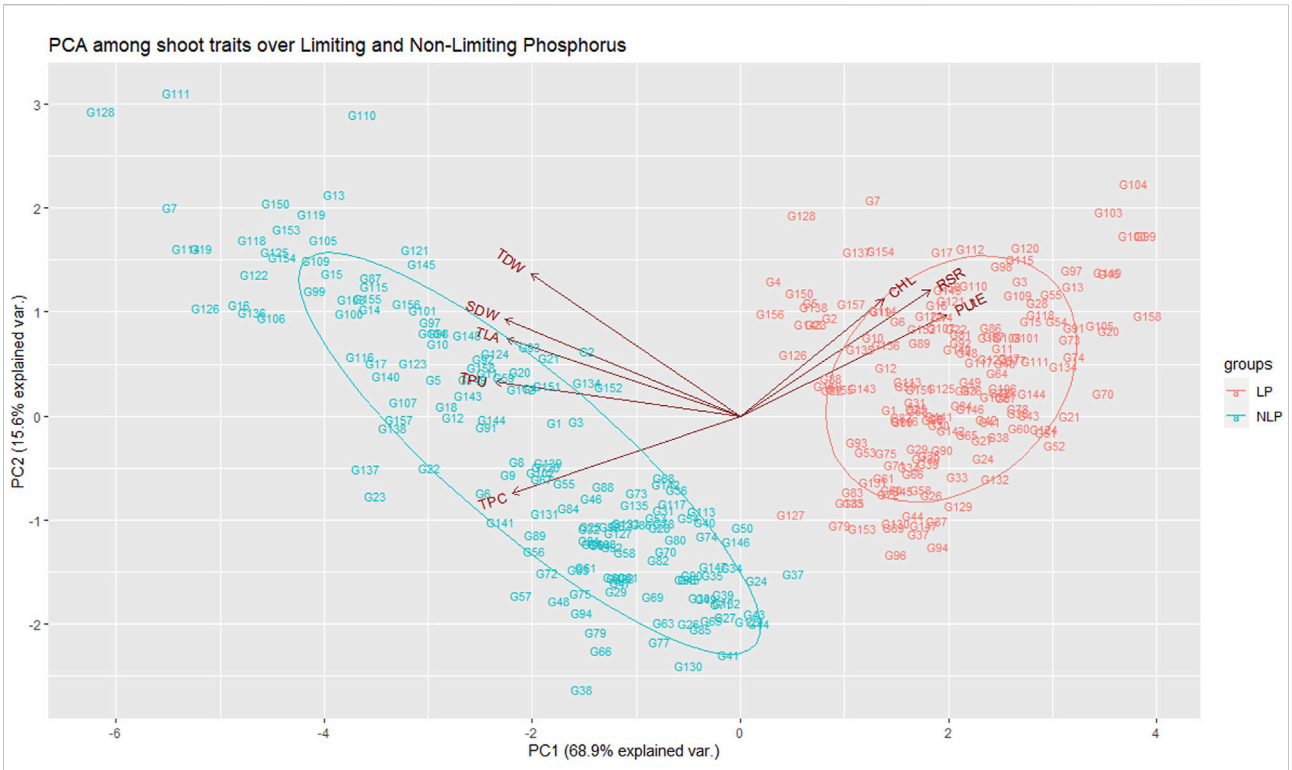


FIGURE 3
Principal component analysis among the traits over non-limiting and limiting conditions. TLA, total leaf area; Chl, chlorophyll content; SDW, shoot dry weight; RSR, root:shoot ratio; TDW, total dry weight; TPC, total phosphorus content; TPU, total phosphorus uptake; PUtE, phosphorus utilization efficiency.

The annotation of putative candidate genes functions

Based on the physical locations of associated SNPs and their linked QTLs, an attempt was made to identify the candidate genes harbouring the associated SNPs (Table 4). The SNP AX94470386 linked to QTL *Q.iari.dt.tla.1* of TLA was found in the gene TraesCS7B02G382300 coding for ATPase-associated with cation transmembrane transporter activity. Similarly, the SNP AX94765690 is associated with QTL *Q.iari.dt.tla.2* of TLA was found within the gene TraesCS1D02G075500 coding for basic region/leucine zipper protein; a positive regulator of transcription. The SNP AX94770913 linked to QTL *Q.iari.dt.tla.3* of TLA was also found in TraesCS6A02G406500 gene transcribing ribosomal protein L9 in bacteria/chloroplast. The SNP AX94832883 is linked to QTL Q.iari.dt.chl.1 and was found in gene TraesCS7D02G040300 coding for the multi-protein family of sulfotransferase involved in postembryonic root development. In the same way, the QTLs namely. *Q.iari.dt.chl.4*, *Q.iari.dt.chl.5*, *Q.iari.dt.chl.6,* and *Q.iari.dt.chl.8* linked to SNPs AX94597699, AX95241386, AX94702861, and AX95230097 were harbouring the genes TraesCS6B02G056100, TraesCS6A02G037800, TraesCS6D02G046800, and TraesCS6D02G047400 coding for RNA-binding domain S1, ribosomal small subunit biogenesis (cleavage involved in rRNA processing), F-box-like domain superfamily and chloroplast rRNA processing, respectively.

The SNP AX94514240 linked to the QTL *Q.iari.dt.sdw.1* was found in gene TraesCS1D02G029900 coding for leucine-rich repeat domain superfamily, NB ARC, P-loop containing nucleoside triphosphate hydrolase and in our study this gene has been found to play role in the expression in several traits like SDW, TDW, and TPU. The QTLs namely *Q.iari.dt.sdw.4*, *Q.iari.dt.sdw.5*, *Q.iari.dt.sdw.6,* and *Q.iari.dt.sdw.9* linked to SNPs AX94503640, AX94812403, AX94635019, and AX95081347 were having the genes TraesCS6D02G193000, TraesCS5A02G023600, TraesCS4A02G307900 and TraesCS2D02G126300 coding for elongation factor Tu and translational elongation, glycosyltransferase activity, polysaccharide catabolic process and bulb-type lectin domain superfamily (LecRLKs play important roles in plant development and stress responses, respectively). The protein coded by these genes plays an important role in seed germination and lateral root development. Other SNPs AX94544797, AX94939596, and AX95113278 and their associated QTLs *Q.iari.dt.sdw.12*, *Q.iari.dt.sdw.15,* and *Q.iari.dt.sdw.16* are localised in genes TraesCS1D02G218300, TraesCS1D02G058900, and

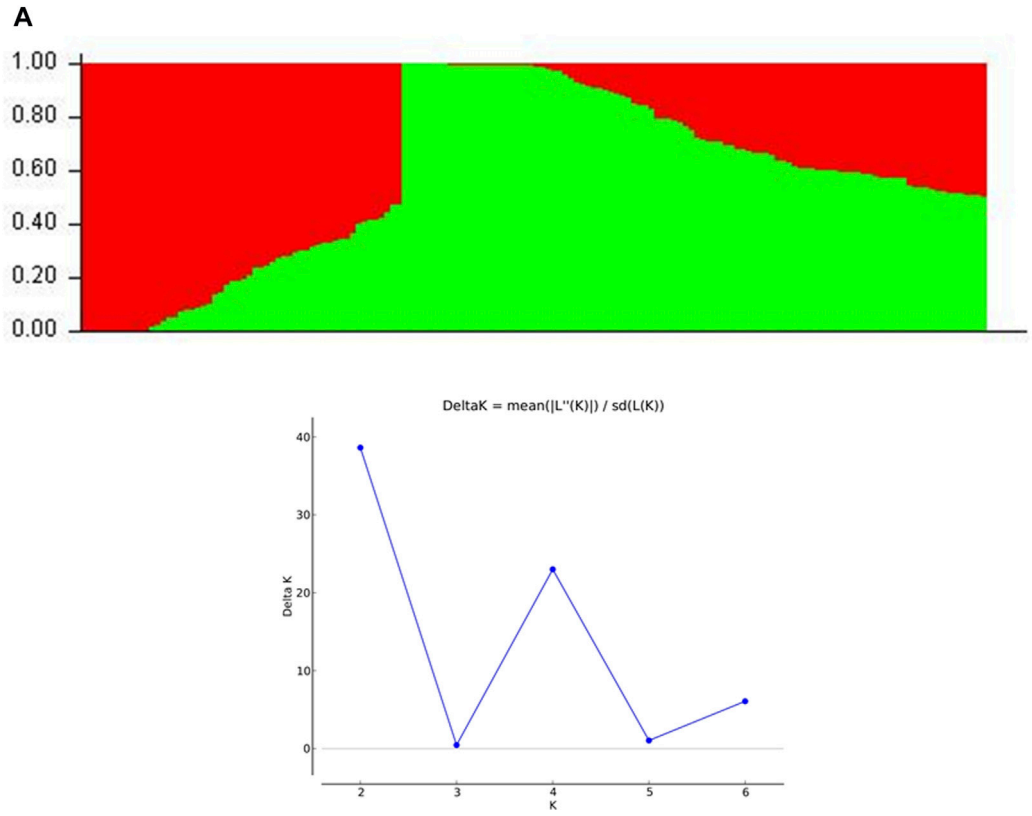


FIGURE 4
(A) Population genetic structure plot in association panel of 158 wheat genotypes (optimal population number K = 2 with two different colours) and Delta K plot depicting peak at K = 2. **(B)** Dendrogram **(C)** Constellation plot (showing three groups) using the Ward method in JMP v.14.Figure.

TraesCS5D02G496600 coding for Phospho-2-dihydro-3-dioxyheptonate aldolase, protein phosphorylation and WAT1-related protein (facilitates auxin export), respectively.

Incidentally, the QTLs namely *Q.iari.dt.sdw.14* and *Q.iari.dt.sdw.20* linked to SNPs AX94626370 and AX94456805 and associated with SDW and TDW carried the genes TraesCS7B02G149200 and TraesCS2D02G584900 coding for Heat shock protein 90, cellular response to heat, protein stabilization, and Tubby-like F-box protein. The SNP AX94475513 linked with *Q.iari.dt.rsr.1* and associated with RSR and TDW has localized in gene TraesCS6A02G095100 coding for F-box-like domain superfamily, leucine-rich repeat domain superfamily, and leucine-rich repeat 2. The four QTLs associated with RSR viz., *Q.iari.dt.rsr.2*, *Q.iari.dt.rsr.6*, *Q.iari.dt.rsr.12,* and *Q.iari.dt.rsr.13* are located in genes coding for protein dephosphorylation, acetylglucosaminyltransferase activity, aldehyde dehydrogenase and calcium-dependent protein binding. Similarly, six QTLs *Q.iari.dt.tdw.4*, *Q.iari.dt.tdw.6*, *Q.iari.dt.tdw.7*, *Q.iari.dt.tdw.9*, *Q.iari.dt.tdw.10,* and *Q.iari.dt.tdw.14* are localized in genes TraesCS4B02G000900, TraesCS7A02G383000, TraesCS7D02G243800, TraesCS5B02G271700, TraesCS2A02G556400, and TraesCS7B02G149200 coding for ATG8-interacting protein, protein phosphorylation, negative regulation of mRNA polyadenylation, protein Iojap, chloroplastic, methyltransferase activity, and heat shock protein 90, respectively.

The QTLs namely *Q.iari.dt.tpc.1* and *Q.iari.dt.tpc.4* linked to SNPs AX94905933 and AX94978370 were harboring the genes TraesCS7A02G110500 and TraesCS3A02G298600 coding for the F-box component of the skp-cullin-f box (SCF) E3 ubiquitin ligase complex (Strigolactone (SL) signal perception, the unidirectional movement of auxin in the stem from tip to base basipetal) and integral component of membrane. *Q.iari.dt.tpc.2* QTL associated with traits TPC and PUtE was having the gene TraesCS3D02G267000 coding for glycerol lipid biosynthetic process, diacylglycerol O-acyltransferase activity. The *Q.iari.dt.tpu.3* was having the gene TraesCS3A02G288900 coding for Zinc ion binding.

Putative candidate genes against wheat gene expression atlas

Using the publicly available global gene expression atlas of wheat, the identified putative genes for P-responsive traits in our

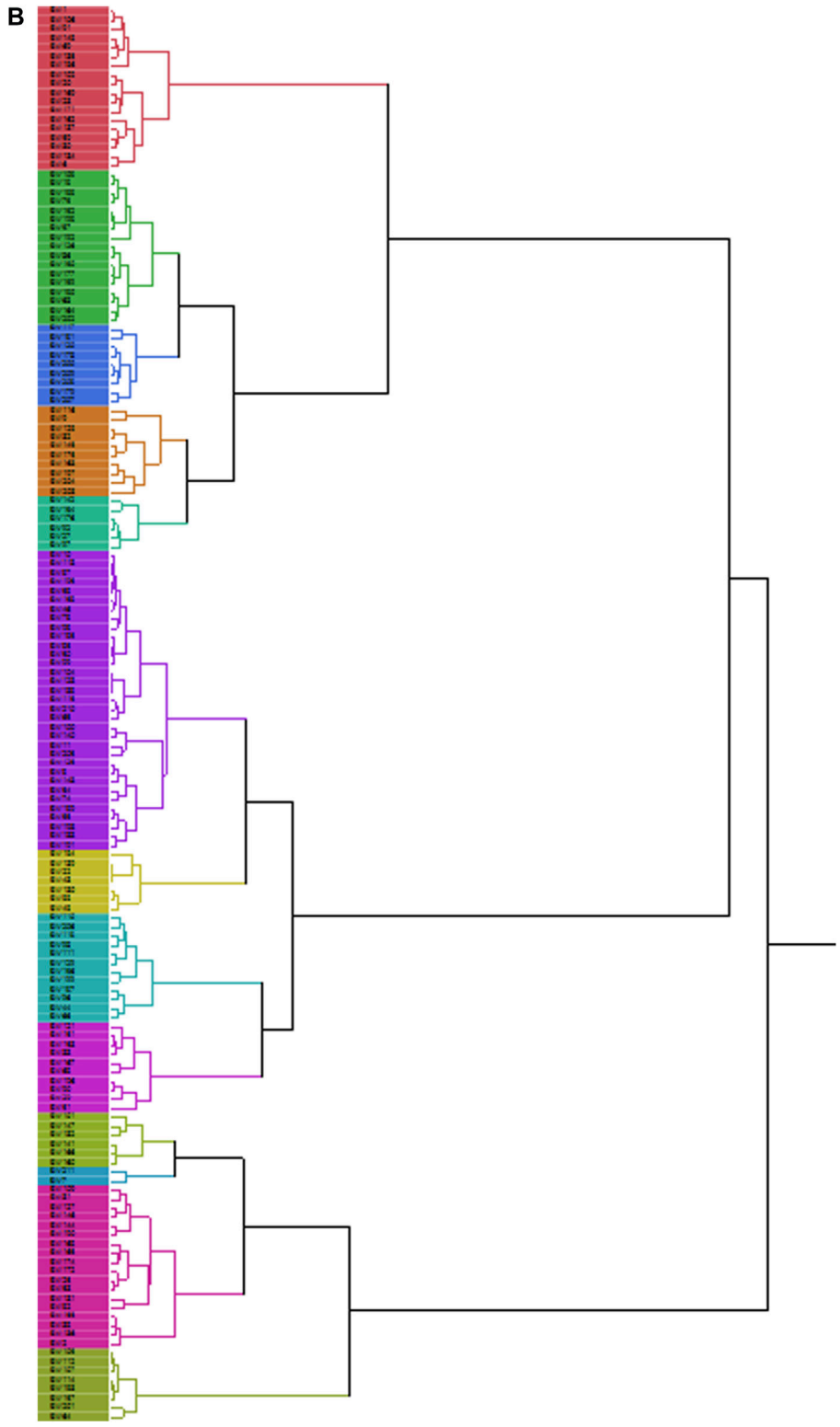


FIGURE 4
(Continued)

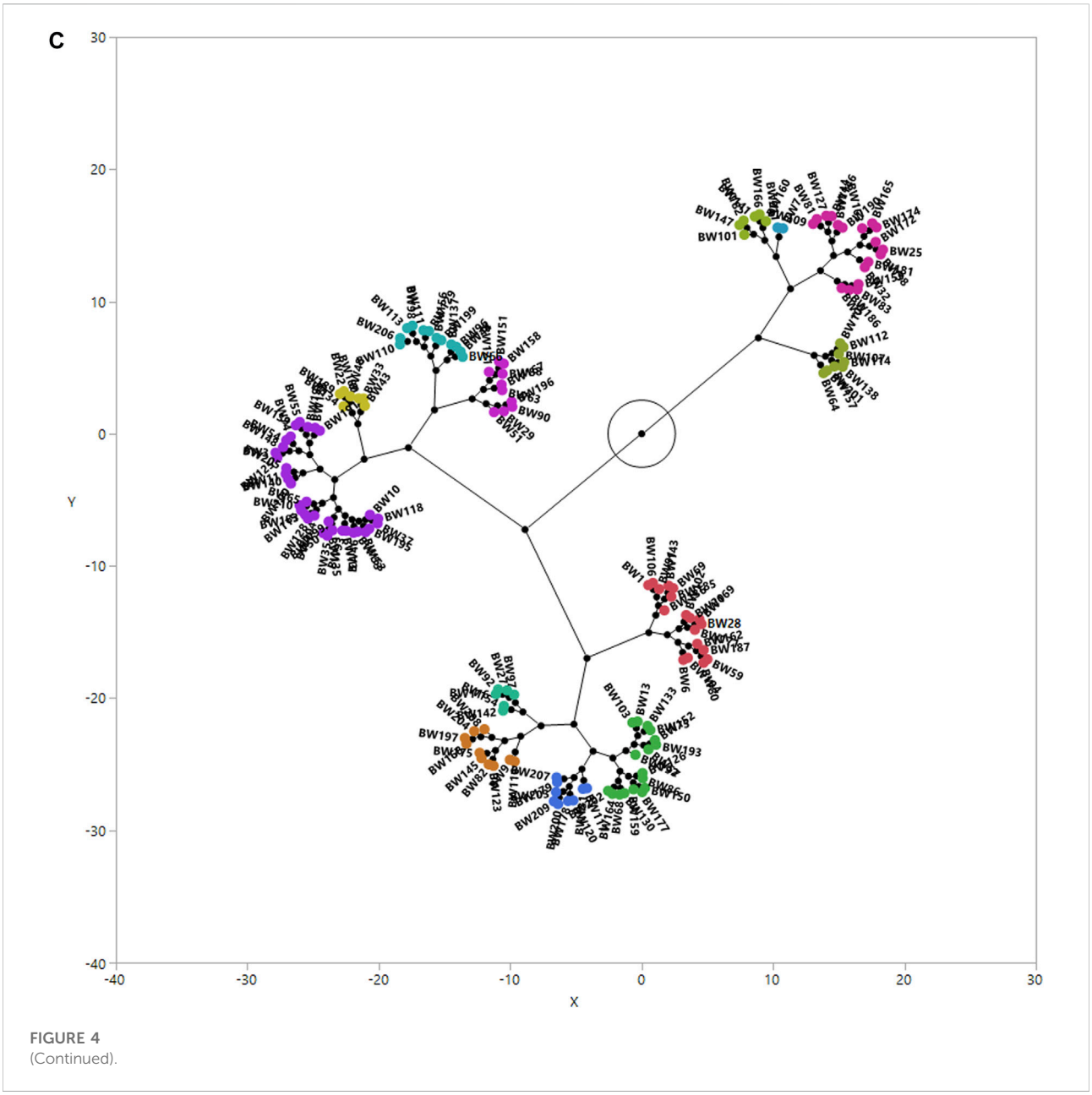


FIGURE 4
(Continued).

study were analysed against the gene expression atlas of wheat targeting the leaf and root traits under normal and phosphate deprivation conditions, to know the reliable expression of these genes (Borrill et al., 2016; Ramírez-González et al., 2018). Most of the identified genes were having the moderate to high consecutive expression in leaves and roots over normal and phosphate-deprivation conditions (Figure 7). These genes are TraesCS7B02G382300 and TraesCS1D02G075500 for TLA-NLP; TraesCS6A02G037800 for Chl-LP; TraesCS1D02G058900 and TraesCS2D02G584900 for SDW/TDW-LP; TraesCS2D02G190700, TraesCS7A02G374700, TraesCS4D02G094500 and TraesCS2D02G435000 for RSR-LP; TraesCS4B02G000900; TraesCS7A02G383000; TraesCS7D02G243800 and TraesCS7B02G149200 for TDW-LP; TraesCS7A02G110500 for TPC-NLP; TraesCS3D02G267000 for TPC/PUtE-NLP and TraesCS3A02G288900 for TPU-LP, which are consecutively expressed genes, except the non-expressed genes like TraesCS7D02G040300, TraesCS7A02G383000, TraesCS2A02G556400, and TraesCS3A02G298600. The other genes like TraesCS6A02G406500 for TLA-LP; TraesCS6B02G056100 and TraesCS6D02G047400 for Chl-LP, TraesCS6D02G193000 for SDW-LP; TraesCS1D02G218300 for SDW-LP and TraesCS5B02G271700 for TDW-LP were

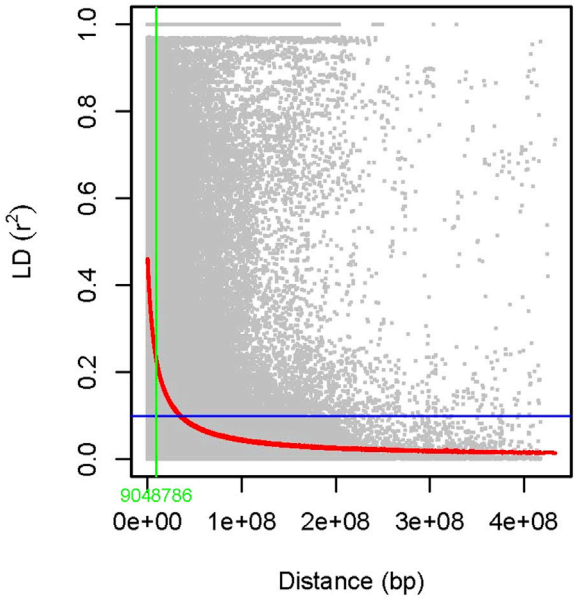


FIGURE 5
Scatter plot showing linkage disequilibrium (LD) decay estimated by plotting (r^2) against genetic distance (bp).

expressed in leaves and shoot of the wheat plants under normal and phosphorus deprivation conditions which coincides with our finding. The genes, TraesCS5A02G023600 for SDW-NLP; TraesCS2D02G126300 for SDW-LP; and TraesCS5D02G496600 for SDW/TDW-LP were expressed in root tissues under normal and phosphorus deprivation conditions.

Discussion

To sustain wheat production globally through breeding intervention is the target of most breeding programs, primarily because wheat is the most important source of food and energy. The challenges to meeting this target are compounded by quickly depleting natural resources due to unsustainable management practices, declining soil health, and changing climatic conditions (Yadav et al., 2010). Intensive cropping with no concern for resource use efficiency—such as is presently the case with phosphorus fertilizers—cannot be sustained for a very long period (Syers, Johnston, and Curtin, 2008; Lott et al., 2011). Under intensive cropping systems, increasing phosphorus efficiency has long been a target to sustain food production (Shenoy and Kalagudi, 2005; Schröder et al., 2011). The residual soil phosphorus from past intense fertilization though contributes considerably to future crop output but with a large lag period as most of the applied phosphorus stays in the soil in the absence of larger uptake and efficient utilization (Sattari et al., 2012). As a result, improved plant capacity to utilize phosphorus effectively will be highly beneficial to crop output. The declining availability of rock phosphate as a source of phosphorus fertilizer and growing awareness about the negative consequences on the environment has piqued interest in improving plant phosphorus uptake and use efficiency (van de Wiel et al., 2016).

The results showed an increase in Chl, RSR, and PUtE and a decrease in TLA, SDW, TDW, TPC, and TPU under P-limited conditions. Improved phosphorus scavenging and uptake (phosphorus acquisition efficiency, PAE) achieved through better RSR with more economical and better utilization in the plant (phosphorus utilization efficiency, PUtE) as indicated in the present study can both improve phosphorus use efficiency (Wang, Shen, and Liao, 2010; Rose and Wissuwa, 2012; Veneklaas et al., 2012; van de Wiel et al., 2016). Correlation studies help to detect the relation between the traits, with respect to a specific treatment. Leaf area plays the most important role in carbon assimilation and therefore, was positively correlated with SDW, TDW, TPC, and TPU. The substantial contribution of TDW, phosphorus concentration, and total phosphorus uptake towards PUE was reported in rice (Wissuwa and Ae, 2001) and wheat (Valizadeh, Rengel, and Rate, 2002). However, better cell expansion to achieve optimum leaf area requires sufficient P in various plant parts, assimilation of which in fact depends upon better root proliferation and higher RSR. RSR exhibited considerably better expression in LP condition. In crop plants, non-availability of P predominantly stimulates root growth as opposed to shoot growth (Lambers et al., 2011). Chl was observed to have weak-to-no association with most of the other traits, but its expression manifests in LP conditions. Plants turn dark green in color when P deprivation is more severe (Hoppo, Elliott, and Reuter, 1999). Respiration and photosynthesis can be slowed down under phosphorus-deprived conditions (Glass, Beaton, and Bomke, 1980), but if respiration is slowed down more than photosynthesis, carbohydrates will be deposited, resulting in dark green leaves.

Principle component analysis clearly shows the importance of chlorophyll content, RSR, and PUtE under P-limited conditions whereas TPU, TPC, TLA, and TDW under nonlimiting P achieve higher yield. Plants have evolved highly specialized adaptive mechanisms through morphological, physiological, and molecular modifications such as an increased root/shoot ratio, an increase in the number of root hairs, association with arbuscular mycorrhizal fungi (AMF), synthesis and release of phosphatases and organic acids, and enhanced expression of phosphatases, to optimize access to soil phosphorus under limited phosphorus (Péret et al., 2011). The length of the vector (distance from the origin) explains the variation contributed by each trait. Along with mean values and correlation, PCA explains the traits TPU, TPC, TLA, SDW, and TDW are highly correlated and explains the large amount of variation contributed by these traits. The significant

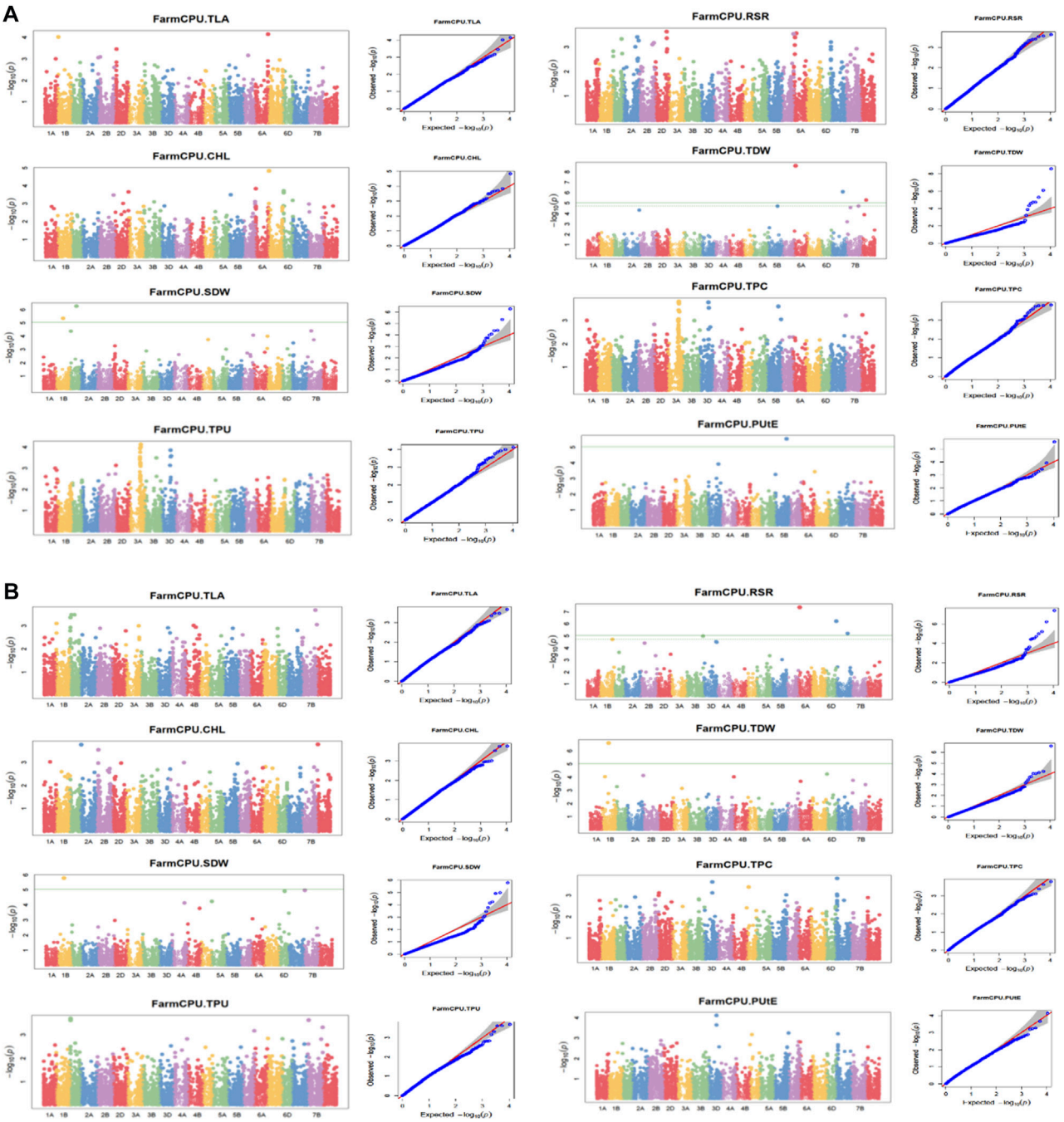


FIGURE 6
(A) (NLP) and **(B)** (LP): Manhattan plots and Q-Q plots (using FARM CPU) for Phosphorus use efficiency traits. TLA, total leaf area; Chl, chlorophyll content; SDW, shoot dry weight; RSR, root: shoot ratio; TDW, total dry weight; TPC, total phosphorus content; TPU, total phosphorus uptake; PUtE, phosphorus utilization efficiency.

crossover interaction between traits across environments (i.e > 90°) helps to identify the traits to be useful for the selection/improvement of genotypes for a specific environment (Yan and Tinker, 2006). In the present study HDCSW18, a wheat variety of very high yield potential and specifically bred for conservation agriculture conditions, exhibits very high phosphorus limitation tolerance largely because of its strong root traits. This genotype because of its strong RSA traits (Dharmateja et al., 2021) has the inherent ability to explore even the deeper layers of soil. Our study clearly shows that better P uptake though, largely depends

TABLE 3 List common QTLs detected across the traits and treatments.

S. No.	Traits	Treatment	QTLs	Chr	Traes ID	Function	References
1	SDW, TDW, TPU	NLP	*Q.iari.dt.sdw.1*	1D	TraesCS1D02G029900	Leucine-rich repeat domain superfamily, NB ARC, P-loop containing nucleoside tri phosphate hydrolase	Flor (1956), Hammond-Kosack and Jones (1997)
2	SDW, TDW	NLP	*Q.iari.dt.sdw.2*	1B	TraesCS1B02G167700	Protien kinase like-domain super family	Sondergaard et al. (2004)
3	SDW, TPU	NLP	*Q.iari.dt.sdw.3*	7B	—	—	—
4	SDW, TDW	LP	*Q.iari.dt.sdw.8*	7B	—	—	—
5	SDW, TDW	LP	*Q.iari.dt.sdw.14*	7B	TraesCS7B02G149200	Heat shock protein 90, cellular response to heat, protein stabilization	Kumar et al. (2020); G. Wang et al. (2011)
6	SDW, TDW	LP	*Q.iari.dt.sdw.20*	2D	TraesCS2D02G584900	Tubby-like F-box protein	Hong et al. (2016)
7	RSR, TDW	NLP and LP	*Q.iari.dt.rsr.1*	6A	TraesCS6A02G095100	F-box-like domain super family, FBD domain, Leucine-rich repeat domain superfamily, Leucine-rich repeat 2	Lechner et al. (2006)
8	TPC, PUtE	NLP	*Q.iari.dt.tpc.2*	3D	TraesCS3D02G267000	Glycerolipid biosynthetic process, diacylglycerol O-acyltransferase activity	Hernández, M. L., 2012
9	TPC, TPU	NLP	*Q.iari.dt.tpc.3*	3A	—	—	—

upon root traits, and is essential for better ground coverage and C assimilation but can improve PUE up to a limited extent because of the limitation imposed by PUtE. It also vouches for a separate breeding programme for both sets of conditions.

For mapping, a panel of 158 wheat genotypes was used, including advanced breeding lines, obsolete varieties, and recently released varieties. The population structure revealed by STRUCTURE analysis infers two major sub-populations (Figure 4A). The presence of two major sub-populations, was further confirmed with the dendrogram and constellation plot. The genotypes in a group shared alleles descended from common parents leading to genetic relatedness among the genotypes. The genotypes are mostly grouped based on pedigree lineage, and evolutionary and geographical origin (Gorafi et al., 2018; Tomar et al., 2021). In view of genetic relatedness, we have adopted the FarmCPU, BLINK, and MLM with population structure and kinship relatedness matrix in association analysis to avoid spurious false positives. The genomic regions responsible for better trait expression both under abundant and deficit P conditions were identified through GWAS. In total 38 QTLs under NPL and 45 QTLs under LP were associated with various traits in the present study. Four QTLs for TLA, eight for Chl, twenty for SDW, seventeen for RSR, eighteen for TDW, five for TPC, seven for TPU, and four QTLs for PUtE were found to be associated in both conditions. In developing countries like India, the major focus of the breeders throughout the 20th century was to achieve higher yields with almost nil or negligible effort on the development of nutrient-efficient genotypes. However, with the faster depletion of natural resources for P and stronger dependence on imports, it becomes inevitable to focus on the development of phosphorus efficient genotypes in wheat. In contrast, breeding phosphorus-efficient wheat genotypes has received a lot of attention (Davies et al., 2002; Wang et al., 2005) in the developed world where genotypic differences in phosphorus acquisition efficiency and phosphorus utilization efficiency for wheat have been frequently reported (Batten, 1993; Manske and Vlek, 2002; Wang et al., 2005; Gunes et al., 2006). The shoot traits are mostly associated and their biological process are also presumed to coordinate with their expression. Strong pleiotropic gene action or tight linkage between the genes results in strong correlation between the traits. Five QTLs in NLP, three QTLs in LP, and one in both treatments were detected for multiple traits. The loci affecting multiple traits should be a potential marker for marker-assisted selection for varietal improvement (Chebib and Guillaume, 2021).

Putative candidate gene functions of identified QTLs

Studying the annotated genomic region in wheat enabled us to identify the genes within the associated SNPs/identified QTLs. The identified putative candidate gene for *Q.iari.dt.tla.1* on 7B for TLA is reported to be responsible for ATPase-coupled cation transmembrane transporter activity. H + -ATPase had a role in nutrient uptake in the root and translocation of these nutrients to the shoots (Sondergaard, Schulz, and Palmgren, 2004). Similarly, *Q.iari.dt.tla.2* for TLA was linked with the gene responsible for basic leucine zipper protein, which is reported to have a role in the positive regulation of transcription. Plants regulate various physiological processes through a regulatory network of transcription factors. Under nutrient starvation, the conserved

TABLE 4 List of putative candidate genes and their functions.

Trait	Treatment	QTL	SNP ID	Chr	Traes ID	Position (bp)	Function	References
TLA	NLP	*Q.iari.dt.tla.1*	AX94470386	7B	TraesCS7B02G382300	648,103,932–648,109,965	ATPase-coupled cation transmembrane transporter activity	Sondergaard et al. (2004)
TLA	NLP	*Q.iari.dt.tla.2*	AX94765690	1D	TraesCS1D02G075500	57,476,559–57,477,974	Basic region/leucine zipper protein, positive regulation of transcription	Lastdrager et al. (2014), Zhang et al. (2020)
TLA	LP	*Q.iari.dt.tla.3*	AX94770913	6A	TraesCS6A02G406500	611,805,720–611,808,712	Ribosomal protein L9	Qi et al. (2019)
CHL	NLP	*Q.iari.dt.chl.1*	AX94832883	7D	TraesCS7D02G040300	20,397,341–20,399,320	Sulfotransferase involved in postembryonic root development	Zhou et al. (2010)
CHL	LP	*Q.iari.dt.chl.4*	AX94597699	6B	TraesCS6B02G056100	36,204,996–36,210,542	RNA-binding domain, S1	Young and Karbstein (2011)
CHL	LP	*Q.iari.dt.chl.5*	AX95241386	6A	TraesCS6A02G037800	18,704,699–18,708,252	Ribosomal small subunit biogenesis	Han et al. (2015)
CHL	LP	*Q.iari.dt.chl.6*	AX94702861	6D	TraesCS6D02G046800	21,050,249–21,055,169	F-box-like domain superfamily	Pérez-Torres et al. (2008)
CHL	LP	*Q.iari.dt.chl.8*	AX95230097	6D	TraesCS6D02G047400	21,980,166–21,985,295	chloroplast rRNA processing	Han et al. (2015)
SDW	NLP	*Q.iari.dt.sdw.4*	AX94503640	6D	TraesCS6D02G193000	267,200,426–267,202,351	Elongation factor Tu, translational elongation	Harvey *et al.* (2019)
SDW	NLP	*Q.iari.dt.sdw.5*	AX94812403	5A	TraesCS5A02G023600	18,795,384–18,797,642	Glycosyltransferase activity	Shi et al. (2020)
SDW	NLP	*Q.iari.dt.sdw.6*	AX94635019	4A	TraesCS4A02G307900	601,364,145–601,368,127	polysaccharide catabolic process	Yuan *et al.* (2019)
SDW	LP	*Q.iari.dt.sdw.9*	AX95081347	2D	TraesCS2D02G126300	73,425,329–73,426,739	Bulb-type lectin domain superfamily involved in plant development, stress response during germination and lateral root development	Vaid et al. (2013), Cheng et al. (2013), Deb et al. (2014)
SDW	LP	*Q.iari.dt.sdw.12*	AX94544797	1D	TraesCS1D02G218300	305,111,236–305,114,814	Phospho-2-dehydro-3-deoxyheptonate aldolase	Entus et al. (2002)
SDW	LP	*Q.iari.dt.sdw.15*	AX94939596	1D	TraesCS1D02G058900	38,780,587–38,789,555	Protein phosphorylation	Li *et al.* (2020)
SDW	LP	*Q.iari.dt.sdw.16*	AX95113278	5D	TraesCS5D02G496600	527,153,578–527,156,833	WAT1-related protein	Ranocha et al. (2013)
RSR	NLP	*Q.iari.dt.rsr.2*	AX94601118	2D	TraesCS2D02G190700	134,790,880–134,792,691	Protein dephosphorylation	Ma et al. (2021a)
RSR	NLP	*Q.iari.dt.rsr.6*	AX95190609	7A	TraesCS7A02G374700	547,568,878–547,574,155	Acetylglucosaminyltransferase activity	Yoo et al. (2021)
RSR	LP	*Q.iari.dt.rsr.12*	AX94944176	4D	TraesCS4D02G094500	69,846,691–69,850,833	Aldehyde dehydrogenase	Tola et al. (2020)
RSR	LP	*Q.iari.dt.rsr.13*	AX94460476	2D	TraesCS2D02G435000	545,937,573–545,939,217	Calcium-dependent protein binding	Reddy and Reddy, (2004)
TDW	NLP	*Q.iari.dt.tdw.4*	AX94572741	4B	TraesCS4B02G000900	621,192–623,232	ATG8-interacting protein	Michaeli *et al.* (2014)
TDW	LP	*Q.iari.dt.tdw.6*	AX94621027	7A	TraesCS7A02G383000	558,079,939–558,082,558	protein phosphorylation	Li et al. (2020)
TDW	LP	*Q.iari.dt.tdw.7*	AX94734828	7D	TraesCS7D02G243800	210,392,267–210,399,523	negative regulation of mRNA polyadenylation	Hunt (2012)
TDW	LP	*Q.iari.dt.tdw.9*	AX94426211	5B	TraesCS5B02G271700	457,128,028–457,130,661	Protein Iojap, chloroplastic	Carey (2016)
TDW	LP	*Q.iari.dt.tdw.10*	AX94884567	2A	TraesCS2A02G556400	760,619,027–760,620,516	Methyltransferase activity	Mishra et al. (2019)
TPC	NLP	*Q.iari.dt.tpc.1*	AX94905933	7A	TraesCS7A02G110500	67,671,323–67,674,222	F-box component of the SKP-Cullin-F box (SCF) E3 ubiquitin ligase complex, Strigolactone signal perception.	R. Liu et al. (2021)

(Continued on following page)

TABLE 4 (*Continued*) List of putative candidate genes and their functions.

Trait	Treatment	QTL	SNP ID	Chr	Traes ID	Position (bp)	Function	References
TPC	LP	*Q.iari.dt.tpc.4*	AX94978370	3A	TraesCS3A02G298600	532,846,752–532,851,697	Integral component of membrane	Cvrcková (2000)
TPU	LP	*Q.iari.dt.tpu.3*	AX94713349	3A	TraesCS3A02G288900	517,074,705–517,080,156	Zinc Ion Binding	Cabot *et al.* (2019)

*Excluding common QTLs.

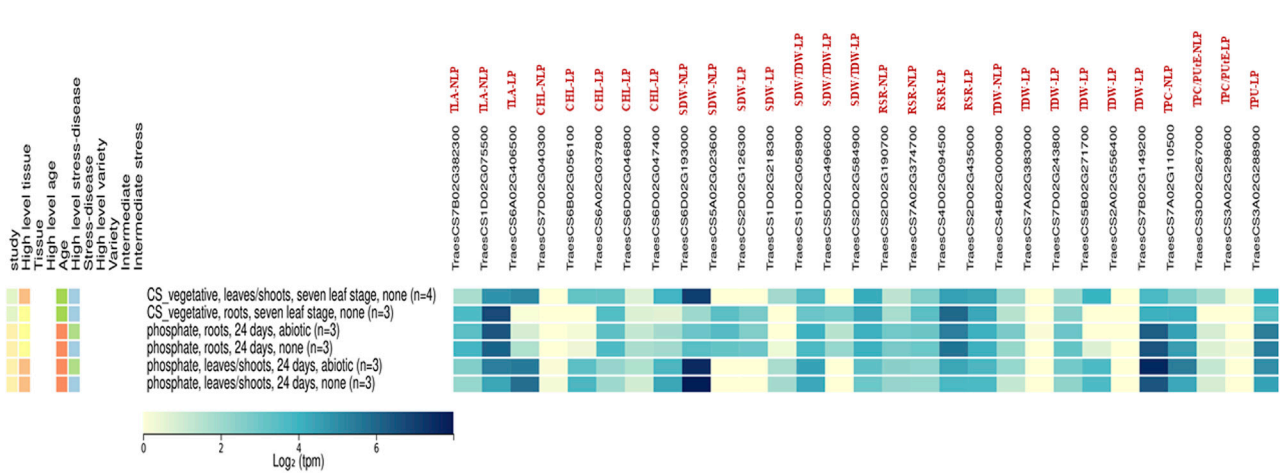


FIGURE 7
The expression analysis for identified Putative Candidate genes in Non-limiting and Limiting conditions by using the gene expression atlas.

sucrose-non-fermenting-1-related protein kinase-1 (SnRK1) mediates the phosphorylation of S1-bZIPs (basic region/leucine zipper) to regulate plant growth and development (Lastdrager et al., 2014). SnRK1 also play a major role in low energy syndrome response under stress condition (Zhang et al., 2020). Mitochondrial ribosomal protein L9, crucial for cells' survival and kernel development regulation (Qi et al., 2019), was linked with *Q.iari.dt.tla.3*.

Q.iari.dt.chl.1 on chromosome 7D is linked to the gene responsible for sulfotransferase. In *Arabidopsis thaliana*, tyrosyl protein sulfotransferase (TPST) was found to act in the auxin/plethora pathway to maintain the stem cell niche of the roots (Zhou et al., 2010). *Q.iari.dt.chl.4* is associated with the RNA-binding domain, S1. As distinct RNA-binding domains (RBDs) are very limited in number, they often combine with multiple RNA-binding motifs for higher affinity and target selectivity. S1 domains binding RNA specifically and non-specifically with high affinity indicate its importance in Rrp5 (ribosomal assembly factor), pre-rRNA complex (Young and Karbstein, 2011). *Q.iari.dt.chl.6* was an ensemble with a gene F-box-like domain superfamily. F-box containing highly conserved motif and their association with cellular degradation with other interacting domains have a robust adaptive role under biotic and abiotic stress conditions, including low-P stress in crops (Pérez-Torres et al., 2008). *Q.iari.dt.chl.8* on chromosome 6D is associated with the gene responsible for chloroplast rRNA processing. The rRNA processing is critical for chloroplast biogenesis and photosynthetic activity resulting in the normal growth of Arabidopsis (Han et al., 2015).

Q.iari.dt.sdw.1 is present on chromosome 1D and associated with nucleotide-binding sites -leucine-rich repeat domain responsible for plant proteins' key role in host-pathogen interaction (Flor, 1956; Hammond-Kosack and Jones, 1997). Similarly, *Q.iari.dt.sdw.5* associated with gene glycosyltransferase OsUGT90A1 on chromosome 5A helps in protecting the plasma membrane during stress (Shi et al., 2020). *Q.iari.dt.sdw.9* present on chromosome 2D is linked to the bulb-type lectin domain superfamily. Lectin receptor-like kinase (LecRLKs) is reported to play essential roles in plant development and stress responses (Vaid, Macovei, and Tuteja, 2013) besides their involvement

in germination processes (Cheng et al., 2013) and lateral root development (Deb et al., 2014). *Q.iari.dt.sdw.14* is associated with heat shock protein 90 which plays an important role in plant adaptation under different environmental conditions (Wang et al., 2011; Kumar et al., 2020). *Q.iari.dt.sdw.15* is associated with gene-responsible protein phosphorylation, which is an integral part of abiotic stress-responsive pathways including phytohormones and ion homeostasis. Auxin hormone plays a very important role in plant adaptation response and the QTL *Q.iari.dt.sdw.16* is associated with a candidate gene responsible for WALLS ARE THIN1 (WAT1), a plant-specific protein that facilitates auxin export from vacuoles (Ranocha et al., 2013). Tubby-like proteins (TLPs) are present in all eukaryotic species (Liu, 2008), including wheat (Hong, Kim, and Seo, 2016). *Q.iari.dt.sdw.20* on chromosome 2D was found associated with a Tubby-like F-box protein-producing gene reported to have a role in adaptation response though with some indistinct mechanism.

QTL *Q.iari.dt.rsr.1* was found to be putatively associated with the gene F-box-like domain superfamily. F-box genes in plants control many important processes including embryogenesis, hormonal responses, seedling development, floral organogenesis, senescence, and pathogen resistance (Lechner et al., 2006). *Q.iari.dt.rsr.2* on chr2D was found putatively associated with gene coding protein dephosphorylation. Plants regulate protein through phosphorylation and dephosphorylation during their response to biotic and abiotic stresses (Ma et al., 2021b). Gene regulating acetylglucosaminyltransferase was found putatively associated with *Q.iari.dt.rsr.6*, the role in various developmental processes under stress conditions has been established in Arabidopsis (Yoo et al., 2021). *Q.iari.dt.rsr.12* was associated with gene-regulating Aldehyde dehydrogenase production, which oxidizes excessive endogenous and exogenous aliphatic and aromatic aldehyde molecules into corresponding carboxylic acids (Tola et al., 2020). *Q.iari.dt.rsr.13* on chromosome 2D is putatively associated with the calcium-dependent protein binding gene. Calcium ions are a messenger for physiological responses to various developmental signals (Reddy and Reddy, 2004).

During the energy deprivation-induced by limiting nutrient supply or other biotic and abiotic stresses, including toxicity, plants have evolved the autophagy process to counter the negative outcome. *Q.iari.dt.tdw.4* links were found putatively associated with ATG8-interacting protein. Similarly, phosphorylation is an important means through which plants regulate post-translational gene expression. *Q.iari.dt.tdw.6* was found to be associated with genes responsible for phosphorylation. *Q.iari.dt.tdw.7* was found to be associated with gene-regulating negative mRNA polyadenylation. Plants also regulate gene expression quantitatively and qualitatively through mRNA polyadenylation (Hunt, 2012). *Q.iari.dt.tdw.9* is an ensemble with a gene responsible for IOJAP protein localized in the chloroplast in Arabidopsis and found to have a role in adaptation to cold stress. *Q.iari.dt.tdw.10* is associated with gene methyltransferase conferring tolerance to salinity stress in Arabidopsis (Mishra et al., 2019). *Q.iari.dt.tpc.1* is related to the gene F-box component of the SKP-Cullin-F box (SCF) E3 ubiquitin ligase complex. The complex is involved in various growth, flower development, and other physiological processes in wheat (Liu et al., 2021).

As earlier discussed, some of the QTLs were associated with more than one trait and which harbors the putative candidate genes. The *Q.iari.dt.sdw.1* associated with SDW, TDW, and TPU under NLP, and harboring the TraesCS1D02G029900 which code for Leucine-rich repeat domain superfamily, NB ARC, P-loop containing nucleoside triphosphate hydrolase responsible for plant proteins' key role in host-pathogen interaction (Flor, 1956; Hammond-Kosack and Jones, 1997) (Table 3). Another two more QTLs viz. *Q.iari.dt.sdw.14* and *Q.iari.dt.sdw.20* were associated with SDW and TDW, and harbouring the TraesCS7B02G149200 and TraesCS2D02G584900, respectively. The putative candidate genes, TraesCS7B02G149200 codes for heat shock protein 90 which plays an important role in plant adaptation under different environmental conditions (Wang et al., 2011; Kumar et al., 2020), while, TraesCS2D02G584900 codes for Tubby-like F-box protein involved in various physiological activities (Hong, Kim, and Seo, 2016). Likewise, QTL named *Q.iari.dt.rsr.1* was associated with RSR and TDW under both conditions, and harbouring the TraesCS6A02G095100 responsible for F-box-like domain super family, FBD domain, Leucine-rich repeat domain superfamily influences the embryogenesis, hormonal responses, seedling development, floral organogenesis, senescence, and pathogen resistance (Lechner et al., 2006). The QTL, *Q.iari.dt.tpc.2* associated with PUtE and TPU, was harbouring the TraesCS3D02G267000 which codes for Glycerolipid biosynthetic process, diacylglycerol O-acyltransferase activity responsible for lipid metabolism (Sanjaya et al., 2013).

In the Wheat Gene Expression Atlas data, several transcripts of different wheat tissues with various ID contents exhibit differential gene expression (Borrill, Ramirez-Gonzalez, and Uauy, 2016). The several genes expressed in leaves, roots, and shoots of the wheat plants under normal and phosphorus-deprived conditions which indicate the potential tissue-specific roles that these genes play in phosphorus stress. Many QTLs explaining the significant variation for many

root traits relevant for PUE in the present study are co-localized in the chromosome regions harbouring essential pertinent genes for stress response and growth development processes in crop plants. Isolation, cloning, and verifying their role in PUE may pave the way for developing stable molecular markers in the crop improvement programme. The expression pattern of identified putative genes against the gene expression atlas indicates that the identified genes in this study have relevance, which can be converted into PCR-based primers for marker-assisted selection.

Conclusion

Identifying regions on the chromosome in the form of QTLs to explain the phenotypic variation for the various breeding traits is an important tool to improve breeding programme efficiency. The identified genomic loci (83 loci across the models and treatments) in the present study explain the significant variation in phosphorus uptake and utilization or its associated root and shoot traits. Their connotation with putative functions or proteins can lead to the validation of gene(s) underlying these loci. The discovered common QTLs controlling several phenotypes may serve as candidate markers for marker-assisted breeding. However, the functional markers need to be validated in a separate independent panel with different genetic backgrounds of wheat genotypes. The proteins encoded by the identified genes are involved in many developmental processes, particularly stress responses. Therefore, there is a need to investigate the functions, activation, deactivation, or changes in the expression rate of these genes in the different developmental stages, to know how these genes enhance the efficiency of a few genotypes under P deficient or P surplus conditions.

Further research into the genic regions of associated SNPs at the transcriptional level is needed to determine the trustworthy source of efficiency imparted in the genotypes studied, which will aid in the identification of distinct developmental pathways. Given the rising cost and relevance of phosphorus as an agricultural input, Crop improvement, PUE is an intrinsically worthwhile objective. However, a PUE-focused breeding programme will compete with other breeding goals like disease resistance and climate change adaptability. The discovery of QTL allows for the creation of trait-relevant markers for marker-assisted or genomic selection methods. Overall, the vast diversity of the genetic resources used in this study will help develop new cultivars of wheat with higher PUE by genomic-assisted breeding.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: https://doi.org/10.6084/m9.figshare.20298006.v1.

Author contributions

RY and PD conceptualized the idea and arranged the resources. All the authors contributed to conducting the experiments and article edits. All the authors contributed to the article and approved the submitted version.

Funding

This research was partly funded by ICAR-World bank-funded NAHEP-CAAST project and the Division of Genetics, ICAR-Indian Agricultural Research.

Acknowledgments

PD acknowledges the National Fellowship for Other Backward Classes (NFOBC) and ICAR-Indian Agricultural Research Institute (IARI), New Delhi for scholarships to complete this work as part of a Ph.D. thesis.

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/fgene.2022.984720/full#supplementary-material

References

Ahmed, M., Khan, S., Irfan, M., Aslam, M. A., Shabbir, G., and Ahmad, S. (2018). *Effect of phosphorus on root signaling of wheat under different water regimes*. Global Wheat Production.

Alvarado, G., Rodriguez, F. M., Pacheco, A., Burgueno, J., Crossa, J., Vargas, M., et al. (2020). META-R: A software to analyze data from multi-environment plant breeding trials. *Crop J.* 8 (5), 745–756. doi:10.1016/j.cj.2020.03.010

Batten, G. D. (1993). *'A review of phosphorus efficiency in wheat'*, Genetic Aspects of Plant Mineral Nutrition, 215–220.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Methodol.* 57 (1), 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x

Bilal, H. M., Aziz, T., Maqsood, M. A., Farooq, M., and Yan, G. (2018). Categorization of wheat genotypes for phosphorus efficiency. *PLoS ONE* 13 (10), e02054711. doi:10.1371/journal.pone.0205471

Borrill, P., Ramirez-Gonzalez, R., and Uauy, C. (2016). expVIP: a customizable RNA-seq data analysis and visualization platform. *Plant Physiol.* 170 (4), 2172–2186. doi:10.1104/pp.15.01667

Bouain, N., Doumas, P., and Rouached, H. (2016). Recent advances in understanding the molecular mechanisms regulating the root system response to phosphate deficiency in Arabidopsis. *Curr. Genomics* 17 (4), 308–304. doi:10.2174/1389202917666160331201812

Brears, R. C. (2015). The circular economy and the water-food nexus. Futur. *Food J. Food, Agric. Soc.* 3, 53–59.

Byrne, S. L., Foito, A., Hedley, P. E., Morris, J. A., Stewart, D., and Barth, S. (2011). Early response mechanisms of perennial ryegrass (Lolium perenne) to phosphorus deficiency. *Ann. Bot.* 107 (2), 243–254. doi:10.1093/aob/mcq234

Cabot, C., Martos, S., Llugany, M., Gallego, B., Tolra, R., and Poschenrieder, C. (2019). A role for zinc in plant defense against pathogens and herbivores. *Front. Plant Sci.* 10, 1171. doi:10.3389/fpls.2019.01171

Carey, A. L. (2016). *Chloroplastic protein IOJAP is important for cold-acclimation in* Arabidopsis thaliana. Knoxville: University of Tennessee.

Cassman, K. G., Dobermann, A., Walters, D. T., and Yang, H. (2003). Meeting cereal demand while protecting natural resources and improving environmental quality. *Annu. Rev. Environ. Resour.* 28 (1), 315–358. doi:10.1146/annurev.energy.28.040202.122858

Chebib, J., and Guillaume, F. (2021). Pleiotropy or linkage? Their relative contributions to the genetic correlation of quantitative traits and detection by multitrait GWA studies. *Genetics* 219 (4), iyab159. doi:10.1093/genetics/iyab159

Cheng, X., Wu, Y., Guo, J., Du, B., Chen, R., Zhu, L., et al. (2013). A rice lectin receptor-like kinase that is involved in innate immune responses also contributes to seed germination. *Plant J.* 76 (4), 687–698. doi:10.1111/tpj.12328

Chiou, T.-J., and Lin, S.-I. (2011). Signaling network in sensing phosphate availability in plants. *Annu. Rev. Plant Biol.* 62, 185–206. doi:10.1146/annurev-arplant-042110-103849

Cvrcková, F. (2000). Are plant formins integral membrane proteins? *Genome Biol.* 1 (1), RESEARCH001–7. doi:10.1186/gb-2000-1-1-research001

Davies, T. G. E., Ying, J., Xu, Q., Li, Z. S., Li, J., and Gordon-Weeks, R. (2002). Expression analysis of putative high-affinity phosphate transporters in Chinese winter wheats. *Plant, Cell & Environ.* 25 (10), 1325–1339. doi:10.1046/j.1365-3040.2002.00913.x

Deb, S., Sankaranarayanan, S., Wewala, G., Widdup, E., and Samuel, M. A. (2014). The S-domain receptor kinase arabidopsis receptor kinase2 and the U box/armadillo repeat-containing E3 ubiquitin ligase9 module mediates lateral root development under phosphate starvation in Arabidopsis. *Plant Physiol.* 165 (4), 1647–1656. doi:10.1104/pp.114.244376

Dharmateja, P., Kumar, M., Pandey, R., Mandal, P. K., Babu, P., Bainsla, N. K., et al. (2021). Deciphering the change in root system architectural traits under limiting and non-limiting phosphorus in Indian bread wheat germplasm. *PloS one* 16 (10), e0255840. doi:10.1371/journal.pone.0255840

Dobermann, A., and Fairhurst, T. H. (2000). *Nutrient disorders and nutrient management*. Singapore: Potash and Phosphate Institute, Potash and Phosphate Institute of Canada and International Rice Research Institute.

Doyle, J. J., and Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bulletin* 19 (1), 11–15.

Entus, R., Poling, M., and Herrmann, K. M. (2002). Redox regulation of Arabidopsis 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase. *Plant Physiol.* 129 (4), 1866–1871. doi:10.1104/pp.002626

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* 14 (8), 2611–2620. doi:10.1111/j.1365-294X.2005.02553.x

Flor, H. H. (1956). The complementary genic systems in flax and flax rust. *Adv. Genet.* 8, 29–54. doi:10.1016/s0065-2660(08)60498-8

Glass, A. D. M., Beaton, J. D., and Bomke, A. (1980). "Role of P in plant nutrition," in *Proceedings of the western Canada phosphate symposium*, 357–368.

Gong, Y., Guo, Z., He, L., and Li, J. (2011). Identification of maize genotypes with high tolerance or sensitivity to phosphorus deficiency. *J. plant Nutr.* 34 (9), 1290–1302. doi:10.1080/01904167.2011.580816

Gorafi, Y. S. A., Ishii, T., Kim, J. S., Elbashir, A. A. E., and Tsujimoto, H. (2018). Genetic variation and association mapping of grain iron and zinc contents in synthetic hexaploid wheat germplasm. *Plant Genet. Resour.* 16 (1), 9–17. doi:10.1017/s1479262116000265

Grant, C. A., Flaten, D. N., Tomasiewicz, D. J., and Sheppard, S. C. (2001). The importance of early season phosphorus nutrition. *Can. J. Plant Sci.* 81 (2), 211–224. doi:10.4141/P00-093

Gulles, A. A., Bartolome, V., Morantte, R., Nora, L. A., Relente, C., Talay, D., et al. (2014). *Randomization and analysis of data using STAR [statistical tool for agricultural research]*. Philippines: Philippine Journal of Crop Science.

Gunes, A., Inal, A., Alpaslan, M., and Cakmak, I. (2006). Genotypic variation in phosphorus efficiency between wheat cultivars grown under greenhouse and field conditions. *Soil Sci. Plant Nutr.* 52 (4), 470–478. doi:10.1111/j.1747-0765.2006.00068.x

Guo, Y., Kong, F. m., Xu, Y. f., Zhao, Y., Liang, X., Wang, Y. y., et al. (2012). QTL mapping for seedling traits in wheat grown under varying concentrations of N, P and K nutrients. *Theor. Appl. Genet.* 124 (5), 851–865. doi:10.1007/s00122-011-1749-7

Gupta, S., Yadav, R., Gaikwad, K., Kushwah, A., Singh, A. M., and Bainsla, N. K. (2016). Genetic improvement trend analysis for end-use quality characteristics among wheat cultivars of North-Western India. *Indian J. Genet. Plant Breed.* 76, 137–143. doi:10.5958/0975-6906.2016.00030.4

Hammond-Kosack, K. E., and Jones, J. D. G. (1997). Plant disease resistance genes. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 48 (1), 575–607. doi:10.1146/annurev.arplant.48.1.575

Han, J. H., Lee, K., Lee, K. H., Jung, S., Jeon, Y., Pai, H. S., et al. (2015). A nuclear-encoded chloroplast-targeted S1 RNA-binding domain protein affects chloroplast rRNA processing and is crucial for the normal growth of *Arabidopsis thaliana*. *Plant J.* 83 (2), 277–289. doi:10.1111/tpj.12889

Harvey, K. L., Jarocki, V. M., Charles, I. G., and Djordjevic, S. P. (2019). The diverse functional roles of elongation factor Tu (EF-Tu) in microbial pathogenesis. *Front. Microbiol.* 10, 2351. doi:10.3389/fmicb.2019.02351

Herrera-Estrella, L., and López-Arredondo, D. (2016). Phosphorus: The underrated element for feeding the world. *Trends Plant Sci.* 21 (6), 461–463. doi:10.1016/j.tplants.2016.04.010

Heuer, S., Gaxiola, R., Schilling, R., Herrera-Estrella, L., Lopez-Arredondo, D., Wissuwa, M., et al. (2017). Improving phosphorus use efficiency: A complex trait with emerging opportunities. *Plant J.* 90 (5), 868–885. doi:10.1111/tpj.13423

Hong, M. J., Kim, D. Y., and Seo, Y. W. (2016). Interactions between wheat tubby-like and SKP1-like proteins *Genes Genet. Syst.* 90 (5), 293–304. doi:10.1266/ggs.14-00084

Hoppo, S. D., Elliott, D. E., and Reuter, D. J. (1999). Plant tests for diagnosing phosphorus deficiency in barley (Hordeum vulgare L.) *Aust. J. Exp. Agric.* 39 (7), 857–872. doi:10.1071/ea99029

Huang, M., Liu, X., Zhou, Y., Summers, R. M., and Zhang, Z. (2019). Blink: A package for the next level of genome-wide association studies with both individuals and markers in the millions. *Gigascience* 8 (2), giy154. doi:10.1093/gigascience/giy154

Hunt, A. G. (2012). RNA regulatory elements and polyadenylation in plants. *Front. Plant Sci.* 2, 109. doi:10.3389/fpls.2011.00109

Kumar, Ashish, Sharma, S., Chunduri, V., Kaur, A., Kaur, S., Malhotra, N., et al. (2020). Genome-wide identification and characterization of Heat Shock Protein Family reveals role in development and stress conditions in *Triticum aestivum* L. *Sci. Rep.* 10, 7858. doi:10.1038/s41598-020-64746-2

Lambers, H., Finnegan, P. M., Laliberte, E., Pearse, S. J., Ryan, M. H., Shane, M. W., et al. (2011). Phosphorus nutrition of proteaceae in severely phosphorus-impoverished soils: Are there lessons to be learned for future crops? *Plant Physiol.* 156 (3), 1058–1066. doi:10.1104/pp.111.174318

Lastdrager, J., Hanson, J., and Smeekens, S. (2014). Sugar signals and the control of plant growth and development. *J. Exp. Bot.* 65 (3), 799–807. doi:10.1093/jxb/ert474

Lechner, E., Achard, P., Vansiri, A., Potuschak, T., and Genschik, P. (2006). F-box proteins everywhere. *Curr. Opin. Plant Biol.* 9 (6), 631–638. doi:10.1016/j.pbi.2006.09.003

Lehman, A., Creighton, L., Sall, J., Vang, E., Blackwelder, M., Zangi, A. D., et al. (2005). *JMP IN 5 software*. Nc: SAS Inc., Ltd.,JMP.

Li, R., Zeng, Y., Xu, J., Wang, Q., Wu, F., Cao, M., et al. (2015). Genetic variation for maize root architecture in response to drought stress at the seedling stage. *Breed. Sci.* 65 (4), 298–307. doi:10.1270/jsbbs.65.298

Li, X., Sanagi, M., Lu, Y., Nomura, Y., Stolze, S. C., Yasuda, S., et al. (2020). Protein phosphorylation dynamics under carbon/nitrogen-nutrient stress and identification of a cell death-related receptor-like kinase in Arabidopsis. *Front. Plant Sci.* 11, 377. doi:10.3389/fpls.2020.00377

Liu, Q. (2008). Identification of rice TUBBY-like genes and their evolution. *FEBS J.* 275 (1), 163–171. doi:10.1111/j.1742-4658.2007.06186.x

Liu, R., Xia, R., Xie, Q., and Wu, Y. (2021). Endoplasmic reticulum-related E3 ubiquitin ligases: Key regulators of plant growth and stress responses. *Plant Commun.* 2 (3), 100186. doi:10.1016/j.xplc.2021.100186

Liu, Y., Wang, L., Deng, M., Li, Z., Lu, Y., Wang, J., et al. (2015). Genome-wide association studies of phosphorus-deficiency-tolerance traits in Aegilops tauschii. *Theor. Appl. Genet.* 128 (11), 2203–2212. doi:10.1007/s00122-015-2578-x

Lott, J. N. A., Kolasa, J., Batten, G. D., and Campbell, L. C. (2011). The critical role of phosphorus in world production of cereal grains and legume seeds. *Food Secur.* 3 (4), 451–462. doi:10.1007/s12571-011-0144-1

Ma, C., Xie, P., Zhang, K., Yang, J., Li, X., Liu, F., et al. (2021b). Contribution of the flag leaf to lead absorption in wheat grain at the grain-filling stage. *Ecotoxicol. Environ. Saf.* 225, 112722. doi:10.1016/j.ecoenv.2021.112722

Ma, W., Ma, L., Li, J., Wang, F., Sisak, I., and Zhang, F. (2011). Phosphorus flows and use efficiencies in production and consumption of wheat, rice, and maize in China. *Chemosphere* 84 (6), 814–821. doi:10.1016/j.chemosphere.2011.04.055

Ma, Z., Wang, J., Li, C., Ren, P., Yao, L., Li, B., et al. (2021a). Global profiling of phosphorylation reveals the barley roots response to phosphorus starvation and resupply. *Front. Plant Sci.* 12, 676432. doi:10.3389/fpls.2021.676432

Manske, G. G. B., and Vlek, P. L. G. (2002). Root architecture–wheat as a model plant. *Plant roots hidden half* 3, 249–259.

Michaeli, S., Honig, A., Levanony, H., Peled-Zehavi, H., and Galili, G. (2014). Arabidopsis ATG8-INTERACTING PROTEIN1 is involved in autophagy-dependent vesicular trafficking of plastid proteins to the vacuole. *Plant Cell* 26 (10), 4084–4101. doi:10.1105/tpc.114.129999

Mishra, P., Jain, A., Takabe, T., Tanaka, Y., Negi, M., Singh, N., et al. (2019). Heterologous expression of serine hydroxymethyltransferase-3 from rice confers tolerance to salinity stress in *E. coli* and Arabidopsis. *Front. Plant Sci.* 10, 217. doi:10.3389/fpls.2019.00217

Ning, L., Kan, G., Du, W., Guo, S., Wang, Q., Zhang, G., et al. (2016). Association analysis for detecting significant single nucleotide polymorphisms for phosphorus-deficiency tolerance at the seedling stage in soybean [Glycine max (L) Merr.]. *Breed. Sci.* 66 (2), 191–203. doi:10.1270/jsbbs.66.191

Pasek, M. A. (2008). Rethinking early Earth phosphorus geochemistry. *Proc. Natl. Acad. Sci. U. S. A.* 105 (3), 853–858. doi:10.1073/pnas.0708205105

Péret, B., Clement, M., Nussaume, L., and Desnos, T. (2011). Root developmental adaptation to phosphate starvation: Better safe than sorry. *Trends Plant Sci.* 16 (8), 442–450. doi:10.1016/j.tplants.2011.05.006

Pérez-Torres, C. A., Lopez-Bucio, J., Cruz-Ramirez, A., Ibarra-Laclette, E., Dharmasiri, S., Estelle, M., et al. (2008). Phosphate availability alters lateral root development in Arabidopsis by modulating auxin sensitivity via a mechanism involving the TIR1 auxin receptor. *Plant Cell* 20 (12), 3258–3272. doi:10.1105/tpc.108.058719

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155 (2), 945–959. doi:10.1093/genetics/155.2.945

Qi, W., Lu, L., Huang, S., and Song, R. (2019). Maize Dek44 encodes mitochondrial ribosomal protein L9 and is required for seed development. *Plant Physiol.* 180 (4), 2106–2119. doi:10.1104/pp.19.00546

Ramírez-González, R. H., Borrill, P., Lang, D., Harrington, S. A., Brinton, J., Venturini, L., et al. (2018). The transcriptional landscape of polyploid wheat. *Science* 361 (6403), eaar6089. doi:10.1126/science.aar6089

Ranocha, P., Dima, O., Nagy, R., Felten, J., Corratge-Faillie, C., Novak, O., et al. (2013). Arabidopsis WAT1 is a vacuolar auxin transport facilitator required for auxin homoeostasis. *Nat. Commun.* 4 (1), 2625–2629. doi:10.1038/ncomms3625

Reddy, V. S., and Reddy, A. S. N. (2004). Proteomics of calcium-signaling components in plants. *Phytochemistry* 65 (12), 1745–1776. doi:10.1016/j.phytochem.2004.04.033

Rengel, Z., and Marschner, P. (2005). Nutrient availability and management in the rhizosphere: Exploiting genotypic differences. *New Phytol.* 168 (2), 305–312. doi:10.1111/j.1469-8137.2005.01558.x

Rose, T. J., and Wissuwa, M. (2012). "Rethinking internal phosphorus utilization efficiency: A new approach is needed to improve PUE in grain crops," in *Advances in agronomy* (Elsevier), 185–217.

Sanjaya, Miller, R., Durrett, T. P., Kosma, D. K., Lydic, T. A., and Muthan, B. (2013). Altered lipid composition and enhanced nutritional value of Arabidopsis leaves following introduction of an algal diacylglycerol acyltransferase 2. *Plant Cell* 25 (2), 677–693. doi:10.1105/tpc.112.104752

Sattari, S. Z., Bouwman, A. F., Giller, K. E., and van Ittersum, M. K. (2012). Residual soil phosphorus as the missing piece in the global phosphorus crisis puzzle. *Proc. Natl. Acad. Sci. U. S. A.* 109 (16), 6348–6353. doi:10.1073/pnas.1113675109

Schröder, J. J., Smit, A. L., CorDell, D., and RosemArin, A. (2011). Improved phosphorus use efficiency in agriculture: A key requirement for its sustainable use. *Chemosphere* 84 (6), 822–831. doi:10.1016/j.chemosphere.2011.01.065

Shenoy, V. V., and Kalagudi, G. M. (2005). Enhancing plant phosphorus use efficiency for sustainable cropping. *Biotechnol. Adv.* 23 (7–8), 501–513. doi:10.1016/j.biotechadv.2005.01.004

Shi, Y., Phan, H., Liu, Y., Cao, S., Zhang, Z., Chu, C., et al. (2020). Glycosyltransferase OsUGT90A1 helps protect the plasma membrane during chilling stress in rice. *J. Exp. Bot.* 71 (9), 2723–2739. doi:10.1093/jxb/eraa025

Singh, S. K., Badgujar, G. B., Reddy, V. R., Fleisher, D. H., and Timlin, D. J. (2013). Effect of phosphorus nutrition on growth and physiology of cotton under ambient and elevated carbon dioxide. *J. Agron. Crop Sci.* 199 (6), 436–448. doi:10.1111/jac.12033

Sondergaard, T. E., Schulz, A., and Palmgren, M. G. (2004). Energization of transport processes in plants. Roles of the plasma membrane H+-ATPase. *Plant Physiol.* 136 (1), 2475–2482. doi:10.1104/pp.104.048231

Syers, J. K., Johnston, A. E., and Curtin, D. (2008). Efficiency of soil and fertilizer phosphorus use. *FAO Fertilizer plant Nutr. Bull.* 18 (108).

Tadesse, W., Ogbonnaya, F. C., Jighly, A., Sanchez-Garcia, M., Sohail, Q., Rajaram, S., et al. (2015). Genome-wide association mapping of yield and grain quality traits in winter wheat genotypes. *PloS one* 10 (10), e0141339. doi:10.1371/journal.pone.0141339

Team, R. C. (2013). *R: A language and environment for statistical computing*. Vienna, Austria': R Foundation for Statistical Computing.

Tola, A. J., Jaballi, A., Germain, H., and Missihoun, T. D. (2020). Recent development on plant aldehyde dehydrogenase enzymes and their functions in plant development and stress signaling. *Genes* 12 (1), 51. doi:10.3390/genes12010051

Tomar, V., Singh, D., Dhillon, G. S., Singh, R. P., Poland, J., Joshi, A. K., et al. (2021). New QTLs for spot blotch disease resistance in wheat (*Triticum aestivum* L.) using genome-wide association mapping. *Front. Genet.* 11, 613217. doi:10.3389/fgene.2020.613217

Vaid, N., Macovei, A., and Tuteja, N. (2013). Knights in action: Lectin receptor-like kinases in plant development and stress responses. *Mol. Plant* 6 (5), 1405–1418. doi:10.1093/mp/sst033

Valizadeh, G. R., Rengel, Z., and Rate, A. W. (2002). Wheat genotypes differ in growth and phosphorus uptake when supplied with different sources and rates of phosphorus banded or mixed in soil in pots. *Aust. J. Exp. Agric.* 42 (8), 1103–1111. doi:10.1071/ea01087

van de Wiel, C. C. M., van der Linden, C. G., and Scholten, O. E. (2016). *'Improving phosphorus use efficiency in agriculture: Opportunities for breeding'*, euphytica. Netherlands: Springer, 1–22. doi:10.1007/s10681-015-1572-3

Veneklaas, E. J., Lambers, H., Bragg, J., Finnegan, P. M., Lovelock, C. E., Plaxton, W. C., et al. (2012). Opportunities for improving phosphorus-use efficiency in crop plants. *New Phytol.* 195 (2), 306–320. doi:10.1111/j.1469-8137.2012.04190.x

Wang, G., Wei, X., Fan, R., Zhou, H., Wang, X., Yu, C., et al. (2011). Molecular analysis of common wheat genes encoding three types of cytosolic heat shock protein 90 (Hsp90): Functional involvement of cytosolic Hsp90s in the control of wheat seedling growth and disease resistance. *New Phytol.* 191 (2), 418–431. doi:10.1111/j.1469-8137.2011.03715.x

Wang, Q., Li, J., Li, Z., and Christie, P. (2005). Screening Chinese wheat germplasm for phosphorus efficiency in calcareous soils. *J. Plant Nutr.* 28 (3), 489–505. doi:10.1081/pln-200049186

Wang, X., Chen, Y., Thomas, C. L., Ding, G., Xu, P., Shi, D., et al. (2017). Genetic variants associated with the root system architecture of oilseed rape (Brassica napus L.)

under contrasting phosphate supply. *DNA Res.* 24 (4), 407–417. doi:10.1093/dnares/dsx013

Wang, X. X., Zhang, J., Wang, H., Rengel, Z., and Li, H. (2021). *Plasticity and co-variation of root traits govern differential phosphorus acquisition among 20 wheat genotypes*. Oikos.

Wang, X., Shen, J., and Liao, H. (2010). Acquisition or utilization, which is more critical for enhancing phosphorus efficiency in modern crops? *Plant Sci.* 179 (4), 302–306. doi:10.1016/j.plantsci.2010.06.007

Wissuwa, M., and Ae, N. (2001). Genotypic variation for tolerance to phosphorus deficiency in rice and the potential for its exploitation in rice improvement. *Plant Breed.* 120 (1), 43–48. doi:10.1046/j.1439-0523.2001.00561.x

Yadav, R., Kumar, L., Jat, M. L., and Gupta, R. (2011). "Tailoring wheat genotypes for conservation agriculture in different cropping systems: An innovative and much needed breeding paradigm," in *5 th world congress on conservation agriculture*.

Yadav, R., Gupta, S., Gaikwad, K. B., Bainsla, N. K., Kumar, M., Babu, P., et al. (2021). Genetic gain in yield and associated changes in agronomic traits in wheat cultivars developed between 1900 and 2016 for irrigated ecosystems of northwestern plain zone of India. *Front. Plant Sci.* 12, 719394. doi:10.3389/fpls.2021.719394

Yadav, R., Singh, S. S., Jain, N., Singh, G. P., and Prabhu, K. (2010). Wheat production in India: Technologies to face future challenges. *J. Agric. Sci. (Tor).* 2 (2). doi:10.5539/jas.v2n2p164

Yan, W., and Tinker, N. A. (2006). Biplot analysis of multi-environment trial data: Principles and applications. *Can. J. Plant Sci.* 86 (3), 623–645. doi:10.4141/P05-169

Yang, M., Wang, C. r., Hassan, M. A., Wu, Y. y., Xia, X. c., Shi, S. b., et al. (2021). QTL mapping of seedling biomass and root traits under different nitrogen conditions in bread wheat (*Triticum aestivum* L.) *J. Integr. Agric.* 20 (5), 1180–1192. doi:10.1016/s2095-3119(20)63192-6

Yoo, J. Y., Ko, K. S., Vu, B. N., Lee, Y. E., Yoon, S. H., Pham, T. T., et al. (2021). N-Acetylglucosaminyltransferase II is involved in plant growth and development under stress conditions. *Front. Plant Sci.* 12, 761064. doi:10.3389/fpls.2021.761064

Young, C. L., and Karbstein, K. (2011). The roles of S1 RNA-binding domains in Rrp5's interactions with pre-rRNA. *Rna* 17 (3), 512–521. doi:10.1261/rna.2458811

Yu, J., and Buckler, E. S. (2006). Genetic association mapping and genome organization of maize. *Curr. Opin. Biotechnol.* 17 (2), 155–160. doi:10.1016/j.copbio.2006.02.003

Yuan, Y., Gao, M., Zhang, M., Zheng, H., Zhou, X., Guo, Y., et al. (2017). QTL mapping for phosphorus efficiency and morphological traits at seedling and maturity stages in wheat. *Front. Plant Sci.* 8, 614. doi:10.3389/fpls.2017.00614

Yuan, Y., Zhang, J., Kallman, J., Liu, X., Meng, M., and Lin, J. (2019). Polysaccharide biosynthetic pathway profiling and putative gene mining of Dendrobium moniliforme using RNA-Seq in different tissues. *BMC Plant Biol.* 19 (1), 521. doi:10.1186/s12870-019-2138-7

Zadoks, J. C., Chang, T. T., and Konzak, C. F. (1974). A decimal code for the growth stages of cereals. *Weed Res.* 14 (6), 415–421. doi:10.1111/j.1365-3180.1974.tb01084.x

Zhang, Huanchao, Cao, F., Fang, S., Wang, G., and Cao, Z. (2005). Effects of agricultural production on phosphorus losses from paddy soils: A case study in the taihu lake region of China. *Wetl. Ecol. Manag.* 13 (1), 25–33. doi:10.1007/s11273-003-5033-8

Zhang, H., and Wang, H. (2015). QTL mapping for traits related to P-deficient tolerance using three related RIL populations in wheat. *Euphytica* 203 (3), 505–520. doi:10.1007/s10681-014-1248-4

Zhang, L., Xie, J., Wang, L., Si, L., Zheng, S., Yang, Y., et al. (2020). Wheat TabZIP8, 9, 13 participate in ABA biosynthesis in NaCl-stressed roots regulated by TaCDPK9-1. *Plant Physiol. biochem.* 151, 650–658. doi:10.1016/j.plaphy.2020.03.039

Zhou, W., Wei, L., Xu, J., Zhai, Q., Jiang, H., Chen, R., et al. (2010). Arabidopsis tyrosylprotein sulfotransferase acts in the auxin/PLETHORA pathway in regulating postembryonic maintenance of the root stem cell niche. *Plant Cell* 22 (11), 3692–3709. doi:10.1105/tpc.110.075721

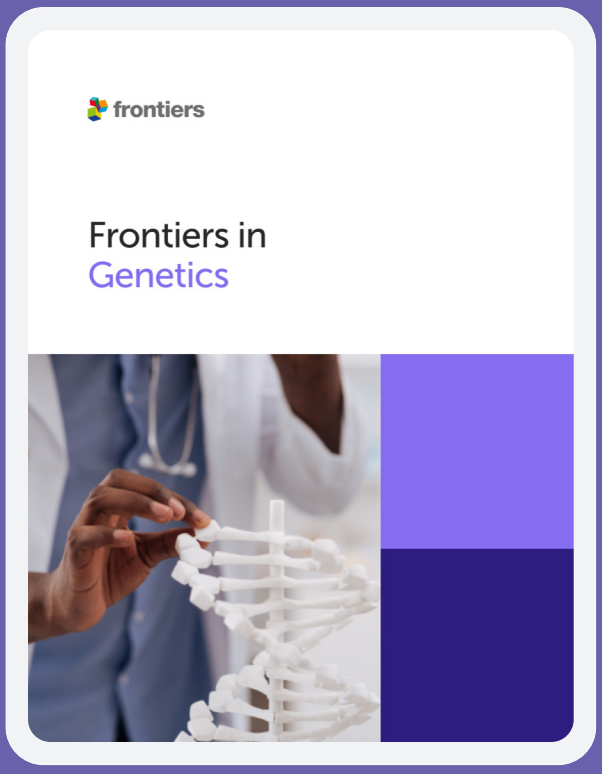
Frontiers in Genetics

Highlights genetic and genomic inquiry relating to all domains of life

The most cited genetics and heredity journal, which advances our understanding of genes from humans to plants and other model organisms. It highlights developments in the function and variability of the genome, and the use of genomic tools.

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