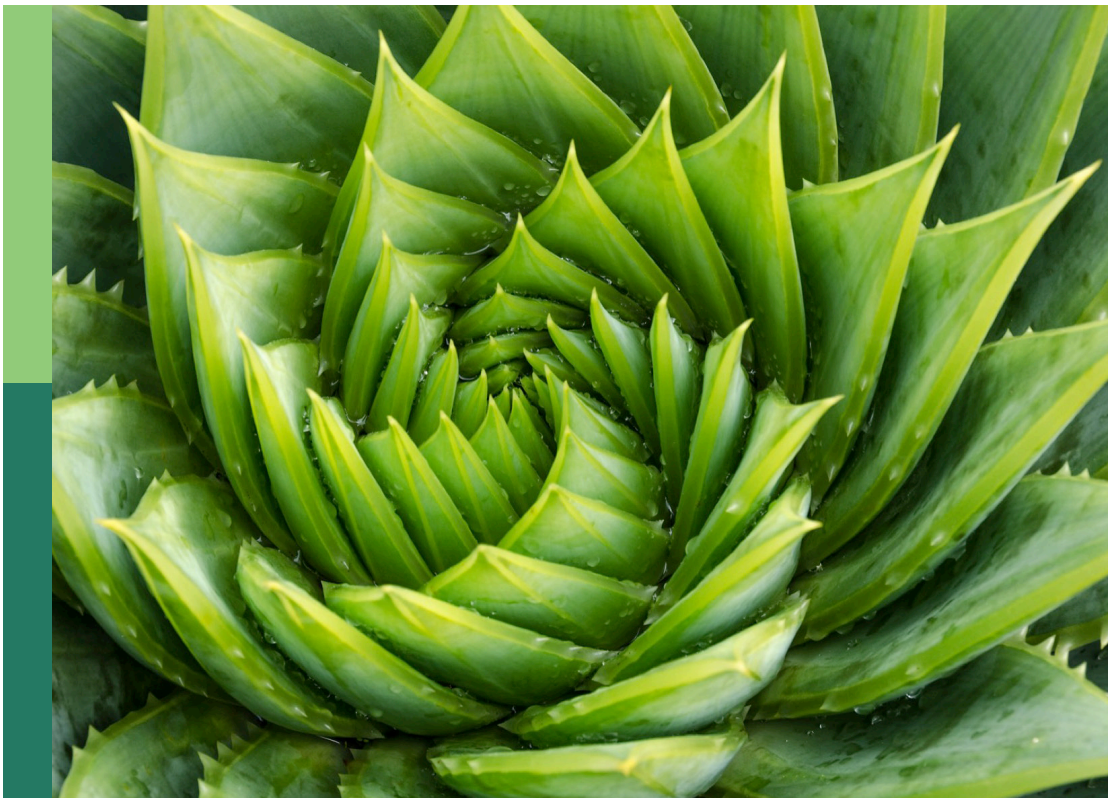
frontiers | Research Topics

Growth regulation in horticultural plants: New insights in the omics era

Edited by
Chenxia Cheng, Lin Xi, Nan Ma and Yi Zheng

Published in
Frontiers in Plant Science



ISSN 1664-8714
ISBN 978-2-8325-3863-0
DOI 10.3389/978-2-8325-3863-0

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

Growth regulation in horticultural plants: New insights in the omics era

Topic editors

Chenxia Cheng — Qingdao Agricultural University, China
Lin Xi — University of Hohenheim, Germany
Nan Ma — China Agricultural University, China
Yi Zheng — Beijing University of Agriculture, China

Citation

Cheng, C., Xi, L., Ma, N., Zheng, Y., eds. (2023). *Growth regulation in horticultural plants: New insights in the omics era*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-3863-0

Table of contents

frontiers | Frontiers in Plant Science

TYPE Editorial
PUBLISHED 20 October 2023
DOI 10.3389/fpls.2023.1305520

OPEN ACCESS

EDITED AND REVIEWED BY
Gregory Thyssen,
Agricultural Research Service (USDA),
United States

*CORRESPONDENCE
Chenxia Cheng
chengchenxia@qau.edu.cn
Lin Xi
lin.xi.260@uni-hohenheim.de

RECEIVED 01 October 2023
ACCEPTED 10 October 2023
PUBLISHED 20 October 2023

CITATION
Cheng C, Ma N, Zheng Y and Xi L (2023) Editorial: Growth regulation in horticultural plants: new insights in the omics era. *Front. Plant Sci.* 14:1305520. doi: 10.3389/fpls.2023.1305520

Editorial: Growth regulation in horticultural plants: new insights in the omics era

Chenxia Cheng[1]*, Nan Ma[2], Yi Zheng[3,4] and Lin Xi[5]*

[1]College of Horticulture, Qingdao Agricultural University, Qingdao, China, [2]Beijing Key Laboratory of Development and Quality Control of Ornamental Crops, Department of Ornamental Horticulture, China Agricultural University, Beijing, China, [3]Beijing Advanced Innovation Center for Tree Breeding by Molecular Design, Beijing University of Agriculture, Beijing, China, [4]Plant Science and Technology College, Beijing University of Agriculture, Beijing, China, [5]Department of Plant Systems Biology, University of Hohenheim, Stuttgart, Germany

KEYWORDS

horticultural plants, high-throughput sequencing technology, multi-omics, growth regulation, system biology

Editorial on the Research Topic
Growth regulation in horticultural plants: new insights in the omics era

In the recent decades, the rapid advancement of next-generation sequencing and "omics" technologies has brought about a revolutionary transformation of the speed and effectiveness of identifying valuable molecules during physiological changes in horticultural plants. A wealth of data on plant genomes, transcriptomes, proteomes, and metabolomes have been generated, enabling researchers to systematically decipher the regulatory mechanisms governing plant growth, development, and stress response. This comprehensive understanding facilitates plant breeding aimed at improving the agronomic traits of interest to breeders and scientists in the community, such as enhancing the yield and quality of horticultural products. Moreover, given the immense diversity of horticultural species and their diverse histories of cultivation, multi-omics studies can offer novel insights into the evolution of various traits and their domestication history. The current Research Topic, titled “Growth Regulation in Horticultural Plants: New Insights in the Omics Era” aims to leverage the potential of high-throughput omics approaches to uncover the molecular mechanisms governing growth and development in horticultural plants. Within this Research Topic, 13 studies authored by 111 researchers contribute valuable insights into multi-omics investigations concerning genes, gene networks or metabolites that regulate growth and development in horticultural plants. Among these studies, nine focus on fruit crops, one centers on vegetable and three delve into ornamental crops.

Research on fruit crops

Apples are highly popular and widely consumed fruits known for their nutritious, sweet and sour flavors. Consequently, traits associated with apples’ commercial values are of great interest to breeders. These traits encompass aspects such as fruit quality, resistance to pathogens, propagation efficiency, regulation of flowering, and more. (Luo et al., 2020; Liao

et al., 2021; Milyaev et al., 2021; Fan et al., 2022). Among these aspects, the appearance of fruit plays a crucial role in the evaluation of apples. In this Research Topic, Wang et al. conducted transcriptome analysis, chemical staining, and LC- and GC-MS analyses to uncover the factors responsible for the formation of russet apple fruit. Their findings highlighted that the elevated levels of lignin and suberin were key contribution to russet formation. Additionally, the study identified several metabolites and genes associated with this process. Virus infections have a significant impact on sustainable development of apple production as well as the fruit's quality (Hadidi et al., 2011). Apple scar skin viroid (ASSVd) leads to a decrease in apple fruit size, causes scarring and staining of the apple peel. These effects can subsequently diminish the marketability of apple fruits and have adverse implications for the apple industry. Li et al. shed light on the underlying mechanisms of ASSVd infections and emphasized the importance of virus-free cultivation for sustainable crop production, as demonstrated through leaf metabolomics. For many fruit varieties, including apple rootstocks, the lack of adventitious root formation presents a challenge for their asexual propagation. Tian et al. pointed out that *MdWRKY87* promoted adventitious rooting through regulating root-related gene involved in auxin signaling pathway by the comparative transcriptome profiling. Breeders have been intrigued by the regulatory mechanisms of floral initiation of cultivated apple and other perennial woody plants, with gibberellins (GAs) identified as key regulators. The function of GA2ox, an enzyme responsible for modulating the levels of active GAs, plays a crucial role in this process. Zhang et al. conducted a comprehensive investigation to elucidate the biological and evolutionary perspectives of GA2ox in apple, focusing specifically on its role in floral induction. The study revealed that the *MdGA2ox2A/2B* genes potentially participate in the repression of flowering.

Grape is among the top fruit crops cultivated worldwide. As a prominent fruit used in wine production, the diversity of compounds from various Vitis species has attracted cultivators and scientists attention (Lin et al., 2019). A noteworthy study by Cheng et al. focused on *Vitis adenoclada*, a unique wild species endemic to China. They successfully assembled a high-quality, chromosome-level genome for *V. adenoclada* with a 498.27 Mb in size and 28,660 protein-coding genes. Through comprehensive metabolome and transcriptome analysis, the research uncovered the higher expressions of COMT and STSs genes linked to increased phenolic compound content, subsequently impacting the sensory attributes of grapes and wine. Furthermore, when serving as table or dried fruit, grape quality primarily hinges on sugar composition and content. In China, the Eurasian grape resources from Xinjiang are very rich in diversity. Zhong et al. utilized GC-MS to determine the sugar components and content in 18 grape varieties, offering effective ways to enhance sugar content through breeding. Grapes are also among the oldest fruit crops cultivated worldwide. Consequently, a range of genomics knowledge is essential for either comprehending the domestication process or developing effective strategies for crop improvement. (Badouin et al., 2020; Navarro-Paya et al., 2021). Xiang et al. conducted a systematic genome-wide and expression analysis of sixty RNA directed DNA methylation (RdDM) pathway genes in grapes and predicted their involvement in multiple biological processes. This study introduced new candidate gene resources for further functional characterization and molecular breeding of grapes.

Citrus, a globally important evergreen fruit tree species, is another major consumable fruit provider, offering fruits as well as diverse by-products. However, it is susceptible to waterlogging, which leads to reduced yields. He et al. employed RNA-seq to identify 17 differentially expressed transcription factors (TFs) that play a pivotal role in enhancing citrus' tolerance to waterlogging.

Blueberries, known for their deliciousness and health benefits, are renowned for their high anthocyanin content. Light plays a significant role in plant growth and development. Zhang et al. conducted an integrated transcriptome and metabolome analysis and found that blue light stimulated the expression of LDOX, UFGT, and OMT genes, leading to the accumulation of cyanidin, pelargonidin, and malvidin anthocyanidins. On the other hand, the combination of red and blue light resulted in the up-regulation of DFR and OMT genes, promoting the accumulation of various downstream metabolites, including delphinidin, petunidin, and peonidin derivatives.

Research on vegetable crops

Chinese cabbage (*Brassica rapa ssp. pekinensis*) holds significant economic importance as a vegetable in Asia and is now widely available in markets around the world (Sun et al., 2022). Leaf size is a critical trait that directly affects Chinese cabbage yield. In a study conducted by Wang et al., transcriptome analysis revealed the pivotal roles played by cyclins (CYCB and CYCD) and transcription factors (MYB47 and MYB88) in regulating cell proliferation, thereby influencing leaf size variation. The findings of this study provide valuable insights for future endeavors aimed at improving Chinese cabbage yield through molecular breeding approaches.

Research on ornamentals

Chrysanthemum, renowned globally for its wide range of intricate flower types, holds immense value in molecular breeding for enhancing floral characteristics. It ranks second in the cut flower trade, following the rose (Su et al., 2019). Li et al. embarked on unraveling the molecular mechanisms underlying chrysanthemum flower development. They identified a total of 44 MIKCc-type MADS box genes throughout the whole genome-wide level in *Chrysanthemum lavandulifolium*. Through their investigation, they observed the expression of C-class genes in the corolla of disc florets, offering potential insights into the morphological distinctions between disc and ray florets. These findings contribute to the establishment of a tetrameric model elucidating floral organ development in *C. lavandulifolium*.

Another globally popular cut flower is the lily. However, it is sensitive to high temperatures (Ding et al., 2021). Among its cultivars, Lanzhou lily (*Lilium davidii* var. unicolor), is a well-

known edible crop cultivated in China. Cao et al. conducted an RNA-seq analysis to gain insights into the thermotolerance mechanisms of Lanzhou lily induced by trichokonins. Their findings suggest that LzHsfA2a-1 plays a crucial role in the acquisition of thermotolerance when exposed to trichokonins. This is attributed to its sustained response to heat stress and enhanced response to trichokonins treatment during prolonged heat stress.

Catalpa fargesii is a native tree species in China, valued for both its timber and ornamental qualities (Yu et al., 2022). Maiyuanjinqiu, a new variety of *C. fargesii*, exhibits significantly lower total chlorophyll content and reduced photosynthesis compared to the original species. Zhang et al. discovered that the m6A methylation level of total RNA was higher in yellow leaves than in green leaves. By silencing the methyltransferase gene *CfALKBH5*, the researchers induced a chlorotic phenotype and increased m6A methylation levels. These findings underscore the significance of m6A modification as an epigenetic mark associated with leaf color in *C. fargesii* and provide valuable insights for manipulating the epitranscriptome to enhance natural leaf color variation in forest breeding.

In summary, the studies mentioned above aim to achieve desirable traits, laying the groundwork and showcasing significant recent progress in identifying novel targets for genetic modification in horticultural plants. Multi-omics studies offer fresh perspectives into understanding the genetic mechanisms underlying crucial plant characteristics. The utilization of omics strategies in plant genetics and breeding research presents both promising opportunities and challenges.

Author contributions

CC: Conceptualization, Writing – original draft, Writing – review & editing. NM: Writing – review & editing. YZ: Writing – review & editing. LX: Conceptualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by National Natural Science Foundation of China (31902056) and Xinjiang Regional Collaborative Innovation Projects (2022E01066).

Acknowledgments

We greatly appreciate the invaluable contributions of all the authors, reviewers, and the Specialty Chief Editor for Functional and Applied Plant Genomics of Frontiers in Plant Science.

Conflict of interest

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

Publisher's note

References

Badouin, H., Velt, A., Gindraud, F., Flutre, T., Dumas, V., Vautrin, S., et al. (2020). The wild grape genome sequence provides insights into the transition from dioecy to hermaphroditism during grape domestication. *Genome Biol.* 21, 223. doi: 10.1186/s13059-020-02131-y

Ding, L., Wu, Z., Teng, R., Xu, S., Cao, X., Yuan, G., et al. (2021). LlWRKY39 is involved in thermotolerance by activating LlMBF1c and interacting with LlCaM3 in lily (Lilium longiflorum). *Hortic. Res.* 8, 36. doi: 10.1038/s41438-021-00473-7

Fan, X., Li, H., Guo, Y., Sun, H., Wang, S., Qi, Q., et al. (2022). Integrated multi-omics analysis uncovers roles of mdm-miR164b-MdORE1 in strigolactone-mediated inhibition of adventitious root formation in apple. *Plant Cell Environ.* 45, 3582–3603. doi: 10.1111/pce.14422

Hadidi, A. B., Candresse, T. M., and Jelkmann, W. (2011). Virus and virus-like diseases of pome and stone fruits. doi: 10.1094/9780890545010

Liao, L., Zhang, W., Zhang, B., Fang, T., Wang, X. F., Cai, Y., et al. (2021). Unraveling a genetic roadmap for improved taste in the domesticated apple. *Mol. Plant* 14, 1454–1471. doi: 10.1016/j.molp.2021.05.018

Lin, J., Massonnet, M., and Cantu, D. (2019). The genetic basis of grape and wine aroma. *Hortic. Res.* 6, 81. doi: 10.1038/s41438-019-0163-1

Luo, F., Evans, K., Norelli, J. L., Zhang, Z., and Peace, C. (2020). Prospects for achieving durable disease resistance with elite fruit quality in apple breeding. *Tree Genet. Genomes* 16, 21. doi: 10.1007/s11295-020-1414-x

Milyaev, A., Kofler, J., Klaiber, I., Czemmel, S., Pfannstiel, J., Flachowsky, H., et al. (2021). Toward systematic understanding of flower bud induction in apple: A multi-omics approach. *Front. Plant Sci.* 12, 604810. doi: 10.3389/fpls.2021.604810

Navarro-Paya, D., Santiago, A., Orduna, L., Zhang, C., Amato, A., D'inca, E., et al. (2021). The grape gene reference catalogue as a standard resource for gene selection and genetic improvement. *Front. Plant Sci.* 12, 803977. doi: 10.3389/fpls.2021.803977

Su, J., Jiang, J., Zhang, F., Liu, Y., Ding, L., Chen, S., et al. (2019). Current achievements and future prospects in the genetic breeding of chrysanthemum: a review. *Hortic. Res.* 6, 109. doi: 10.1038/s41438-019-0193-8

Sun, X., Li, X., Lu, Y., Wang, S., Zhang, X., Zhang, K., et al. (2022). Construction of a high-density mutant population of Chinese cabbage facilitates the genetic dissection of agronomic traits. *Mol. Plant* 15, 913–924. doi: 10.1016/j.molp.2022.02.006

Yu, X., Li, F., Zhao, Q., Wang, J., Liu, Y., Yi, F., et al. (2022). Primary selection of excellent catalpa fargesii clones based on growth and wood properties. *Forests* 13, 1659. doi: 10.3390/f13101659

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 08 November 2022
DOI 10.3389/fpls.2022.1057226

OPEN ACCESS

EDITED BY
Lin Xi,
University of Hohenheim, Germany

REVIEWED BY
Haipeng Zhang,
Henan Agricultural University, China
Peixian Nie,
Shandong Institute of Pomology, China

*CORRESPONDENCE
Yugang Zhang
jsh2862317@163.com
Shenghui Jiang
ygzhang@qau.edu.cn

SPECIALTY SECTION
This article was submitted to Functional and Applied Plant Genomics, a section of the journal Frontiers in Plant Science

RECEIVED 29 September 2022
ACCEPTED 21 October 2022
PUBLISHED 08 November 2022

CITATION
Wang Z, Liu S, Huo W, Chen M, Zhang Y and Jiang S (2022) Transcriptome and metabolome analyses reveal phenotype formation differences between russet and non-russet apples.
Front. Plant Sci. 13:1057226.
doi: 10.3389/fpls.2022.1057226

Transcriptome and metabolome analyses reveal phenotype formation differences between russet and non-russet apples

Ziqi Wang[1,2], Shasha Liu[1,2], Wenping Huo[1,2], Min Chen[3], Yugang Zhang[1,2]* and Shenghui Jiang[1,2]*

[1]College of Horticulture, Qingdao Agricultural University, Qingdao, China, [2]Engineering Laboratory of Genetic Improvement of Horticultural Crops of Shandong Province, Qingdao, China, [3]Yantai Institute of Coastal Zone Research, Chinese Academy of Sciences, Yantai, China

The apple is an economically important fruit, and fruit russeting is not conducive to its appearance. Although studies have examined fruit russeting, its mechanism remains unclear. Two apple strains of the F_1 hybrid population derived from 'Fuji' and 'Golden Delicious' were used in this study. We found that the skin of russet apples was rough and fissured, while that of non-russet apples was smooth and waxy. Chemical staining, LC- and GC-MS showed that both lignin and suberin were increased in russet apple skin. Meanwhile, genes involved in lignin and suberin synthetic pathways were upregulated in russet apple skin. Additionally, we found many differentially expressed genes (DEGs[1]) involved in hormone biosynthesis and signaling and stress responses in the two apple strains. We found that WRKY13 may influence russeting by regulating lignin synthesis. Our study identified several candidate metabolites and genes, which will provide a good foundation for further research.

KEYWORDS

transcriptome, metabolome, apple, fruit russeting, suberin

Abbreviations: DEGs, differentially expressed genes; DAFB, days after full bloom; SEM, scanning electron microscopy; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; DAMs, differentially accumulated metabolites; WGCNA, weighted gene co-expression network analysis; CK, cytokinin; ABA, abscisic acid; BR, brassinolide; MeJA, methyl jasmonate; HSPs, heat shock proteins; AQPs, aquaporins; PRXs, peroxidases.

Introduction

Apples are one of the most commonly consumed fruits in everyday life and are preferred by the public for their sweet and sour tastes. According to FAO 2020 (https://www.fao.org/faostat/en/#data/QCL/visualize), global apple exports reached 8.2 million tons in 2020, which holds great economic importance. In addition to taste and aroma, fruit appearance is an important criterion in apple evaluation. The skin of a normal apple is smooth and waxy, and the primary skin consists of cuticle, epidermis, and hypodermis (Khanal et al., 2013). Compared with normal fruits, russet fruits show yellow-brown rust spots on their surface.

Fruit russeting is a secondary metabolic process, where cuticle cracks and suberin rapidly form to replace the broken cuticle to prevent water dissipation and reduce pathogen infection (Lashbrooke et al., 2015). The causes of fruit russeting have been studied in fruit crops, such as apple, pear, and grape (Lashbrooke et al., 2015; Huang et al., 2020; Zhang et al., 2021). Apple russeting is influenced by environmental factors, and a prolonged period of moisture on the surface of the fruit skin or high humidity in the environment can cause the skin to crack, ultimately leading to russeting (Knoche et al., 2011).

With the development of omics technologies, fruit russeting research has shifted from physiological characteristics to molecular mechanisms. Transcriptome, metabolome, and proteome association analyses identified differences in suberin, phenylpropane, cutin, and wax biosynthesis between russet and non-russet pears (Shi et al., 2021). Therefore, it is speculated that phenylpropane and cuticle metabolism have important effects on fruit russeting. LC-MS and nuclear magnetic resonance identified three triterpene-caffeates extracted from russet apple skins of the 'Merton Russet' cultivar as oleanolic acid-3-trans-caffeate, betulinic acid-3-cis-caffeate, and betulinic acid-3-trans-caffeate, which were only found in russeting fruit (Andre et al., 2013). Triterpenes are included in the epicuticular and intracuticular wax of plant cuticles (embedded in cutin polymers). In apples, triterpenes are present in the cuticular skin layer, where they account for 60% of the wax content (Zhang et al., 2020). Thus, the presence of triterpene caffeates disrupts cuticle formation, which initiates fruit russeting. Furthermore, the transcription factor MYB66 can bind to *OSC5* to regulate triterpene caffeate formation, which disrupts normal cuticle synthesis and leads to fruit russeting in apple (Falginella et al., 2021). In apples, skin suberization has been linked to the presence of triterpene compounds (Legay et al., 2016). In addition, transcriptome results showed that suberin synthesis-related genes are abundantly expressed in russet fruits, including *LACS*, *FAR*, *KCS*, *CYP86*, *GPAT*, *ASFT*, *ABCG*, and *LTP* (Zhang et al., 2021). Transcription factors are also involved in suberin accumulation in fruit russeting, which is regulated by MYB93 in apple peels, thereby affecting fruit russeting (Legay et al., 2016). In pears, lignin also demonstrated an important effect on russeting because of its high expression of *PAL*, *4CL*, *HCT*, *C4H*, *COMT*, and *CcoAOMT* (Wang et al., 2020). In addition, *CCR* and *CAD4-* and *CAD9-like* genes were downregulated in russet apples (Legay et al., 2015). In addition to the upregulation of lignin synthesis genes, a substantial accumulation of phenolics related to lignin synthesis has been identified in russet grape fruits (Huang et al., 2020). Moreover, the genetic factors involved in fruit russeting were analyzed. The main effector gene controlling russet variation in 'Renetta Grigia di Torriana' apples is located on chromosome LG12 (Falginella et al., 2015). The *PpRus* gene, which controls the pear russet skin trait, was identified *via* Bulked Segregant analysis and is located on chromosome 8 of the pear genome (Ma et al., 2021).

In this study, we found that the skin of russet apples had more lignin and suberin than that of non-russet apples. Metabolome analysis showed that metabolites related to lignin and suberin synthesis in russet apple skin were higher than those in non-russet apple skin. Transcriptome data also demonstrated that the genes involved in the lignin and suberin synthetic pathways were upregulated in russet apple skin. Interestingly, we also found that a WRKY transcription factor, WRKY13, may influence russeting by regulating lignin synthesis. Our findings provide several candidate metabolites and genes to understand fruit russeting.

Materials and methods

Plant materials

Russet and non-russet apples were obtained from two strains of the F_1 hybrid population derived from 'Fuji' and 'Golden Delicious' in 2020. Apple trees were cultivated in the Jiaozhou Agricultural Science and Technology Demonstration Garden of Qingdao Agricultural University (Jiaozhou, Shandong Province). We chose fruits 120 days after full bloom (DAFB) as the material for our experiments because of their different phenotypes. Both fruit types were grown under the same conditions. Each type of fruit was similar in shape and size, without any mechanical damage or disease. The skin tissue was collected, immediately frozen in liquid nitrogen, and stored at -80 °C for further use. Fresh fruit pericarps (0.5 × 0.5 cm) were fixed in FAA fixative for microscopic observation.

Microscopy

A portion of the fruit peel was cut into paraffin sections, and the samples were sliced into 10 μm-thick sections using a Leica RM2245 paraffin slicer. Phloroglucinol HCl (Wiesner) was used for lignin deposition. Sudan red 7B was used to stain the deposited suberin. The stained sections were observed under a Leica microscope. Samples observed *via* scanning electron

microscopy (SEM) were cut into small pieces (0.5 × 0.5 cm), fixed using FAA, and dehydrated through an acetone gradient. The samples were then subjected to critical-point drying and gold spraying. The fabricated samples were then subjected to SEM analysis.

GC-MS sample extraction and determination

The peels were ground into a powder. Fruit peel (50 mg) was extracted with 150 μL of methanol, 200 μL of methyl tert-butyl ether, and 50 μL of 36% phosphoric acid. The solution was vortexed for 3 min and then centrifuged at 12000 r/min for 5 min at 4°C. Then, 200 μL of supernatant was dried on a nitrogen blower, and 300 μL 15% boron trifluoride methanol was added. The solution was vortexed for 3 min and incubated at 60°C for 30 min. Five hundred microliters of n-hexane and 200 μL of saturated NaCl were added to the solution at room temperature. Then, the solution was vortexed for 3 min and centrifuged at 12000 r/min for 5 min at 4°C. The n-hexane layer (100 μL) was used for further analysis. The sample derivates were analyzed using a GC-EI-MS system (GC, Agilent 8890; MS, 5977 B System). The analytical conditions were as follows, GC: column, DB-5MS capillary column (30 m to 0.25 mm × 0.25 μm, Agilent); carrier gas: high purity helium (purity > 99.999%); the heating procedure commenced at 40 °C (2 min), 30 °C/min to 200 °C (1 min), 10 °C/min to 240 °C (1 min), 5 °C/min to 285 °C (3 min); traffic: 1.0 mL/min; inlet temperature, 230 °C; injection volume: 1.0 μL. The EI-MS procedure utilized a Agilent 8890-5977B GC-MS System, temperature: 230°C, ionization voltage: 70 eV; transmission line temperature: 240 °C, four-stage rod temperature: 150 °C, solvent delay: 4 min, scanning mode: SIM. Fatty acids and their metabolites were detected using MetWare (http://www.metware.cn/) on an Agilent 8890-5977B GC-MS platform.

LC-MS sample extraction and detection

Fruit peels (100 mg) were ground into powder in liquid nitrogen, and 500 μL of 80% methanol was added. The solution was vortexed, incubated for 5 min on ice, and then centrifuged at 12000 r/min for 20 min at 4°C. The supernatant was diluted with mass spectrometry-grade water and then centrifuged at 12000 r/min for 5 min at 4°C. The second supernatant was used for LC-MS analysis. Separation was performed using an Xselect HSS T3 column (2.5 μm, 2.1 × 150 mm). The mobile phase was 0.1% aqueous formic acid solution (solvent A) and 0.1% formic acid in-acetonitrile (solvent B) at a flow rate of 0.4 ml min^{-1}. The linear gradient of phase B was as follows: 0–2 min, 2%; 2–15 min, 2%; 15–17 min, 100%; 17-7.1 min, 2%; 17.1–20 min, 2%. Mass spectrometry conditions for positive ions were as follows: curtain gas: 35 psi; collision gas: medium; IonSpray voltage: 5500 V; temperature: 550°C; Ion Source gas: 1:60; Ion Source gas: 2:60. 2. Mass spectrometry conditions for negative ions were as follows: curtain gas: 35 psi; collision gas: medium; IonSpray voltage: -4500 V; temperature: 550°C; ion source gas 1: 60; ion source gas 2: 60.

Based on the Novogene database, multiple reaction monitoring was used to detect the samples. The compounds were quantified using their daughter ions and characterized by retention time, declustering potential, and collision energy. SCIEX OSV1.4 software was used for integration, correction, and analysis.

RNA extraction, cDNA library construction, and qRT-PCR determination

Total RNA was extracted from apple fruit peels using the RNAprep Pure Plant Kit (Tiangen, Beijing, China), according to the manufacturer's instructions. RNA concentration and quality were assessed using an Agilent 2100 instrument (Agilent Technologies, Santa Clara, CA, USA). Then, cDNA libraries were constructed for Illumina sequencing, following the manufacturer's protocol. The libraries were sequenced on an Illumina PE150 platform (Illumina, San Diego, CA, USA). After removing the adapters and low-quality reads, the clean reads were aligned to the reference genome (https://iris.angers.inra.fr/gddh13/index.html) using the Hisat2 software. Gene expression was analyzed using the Counts and StrintTie tools. Differentially expressed genes (DEGs) were identified using the DESeq2 package. RNA-seq included three biological replicates for each sample. Transcriptome data were uploaded to the NCBI Short Read Archive (https://www.ncbi.nlm.nih.gov/sra/) under the accession number PRJNA871277.

First-strand cDNA was synthesized from RNA samples using the PrimeScript™ RT Reagent Kit (Takara RR047A, Japan), and cDNA was used as a template for gene expression analysis. Gene expression was measured using qRT-PCR on a CFX96 Touch Real-Time PCR Detection System (Bio-rad, CFX96 touch). The qRT-PCR was carried out with ChamQ SYBR Color qPCR Master Mix (Vazyme, China), primers, cDNA, and RNase-free water in a total volume of 20 μL. The primers used for qRT-PCR are listed in Table S1. Relative express was calculated by the cycle threshold (Ct) $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). The independent biological experiments were performed in triplicates for each sample.

Statistical analysis

Each result is the mean of three biological replicates for each experiment. One-way analysis of variance was performed *via* Student's t-test using the GraphPad Prism software. P values < 0.05 were considered significant.

Results

Microstructural peel observation and comparison of lignin and suberin content between russet and non-russet apples

Russet and non-russet apples were harvested 120 DAFB. Russet fruits had yellowish-brown rust spots on their pericarp surface, whereas non-russet fruits had a smooth surface with an intact cuticle (Figure 1A). To further investigate the structural differences in the pericarp, we observed the microstructure of the pericarp of both fruits using SEM. The russet fruit pericarp appeared fissured and had a laminar build-up, whereas the non-russet fruit pericarp had a smooth surface with a complete waxy texture (Figure 1B). The fruit peel was chemically stained to further analyze russeting composition. The sections were stained to determine the content and location of the peel, and Sudan Red 7B staining was used to determine the suberin content. High suberin content was observed in the russet sample (Figure 1C), and the lignin staining was darker in russet peels than in non-russet peels (Figure 1D). Moreover, lignin content was significantly higher in russet peels than in non-russet peels (Figure 1E).

Differential gene expression analysis and validation of RNA-seq data

As the fruits of the two strains exhibited different peel phenotypes, RNA-seq was performed to investigate the differences in gene expression between russet and non-russet peels. RNA-seq detected a total of 46,558 known genes and 704 new genes. The Pearson coefficient between biological replicates of all samples was greater than 0.75 (Figure 2A), which can be considered a high correlation coefficient between samples with correlated expression patterns, and the experimental results could be used to analyze DEGs. Screening based on $|\log_2(\text{FoldChange})| > 1$ and false discovery rate (FDR) ≤ 0.05, we identified 4718 genes with upregulated expression and 4448 genes with downregulated expression compared to non-russet fruit (Figure 2B).

Gene Ontology (GO) analysis showed DEG enrichment in molecular function, biological processes, and cellular components (Figure S1). The 20 terms with the highest number of enriched genes in the three ontologies were selected for analysis. For biological processes, genes were enriched in single-organism signaling (143), signal transduction (143), signaling (143), single-organism biosynthetic process (130), lipid metabolic process (120), and other terms. For cellular components, genes were enriched in the protein complex (105), ribonucleoprotein complex (94), and intracellular ribonucleoprotein complex (93). For molecular function, genes were enriched in ADP binding (184), cofactor binding (140), heme binding (133), and tetrapyrrole binding (133).

The DEGs were mapped to the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database (Figure 2C). KEGG enrichment analysis showed that the DEGs were significantly enriched in stilbenoid, diarylheptanoid, and gingerol biosynthesis (ko:mdm00945), flavonoid biosynthesis (ko:mdm00941), phenylpropanoid biosynthesis (ko:mdm00940), alpha-linolenic acid metabolism (ko:mdm00592), cutin, suberine, and wax biosynthesis (ko:mdm00073), fatty acid degradation (ko:mdm00071), fatty acid biosynthesis (ko:mdm00061), and fatty acid metabolism (ko:mdm01212).

Transcription factors play important roles in plant growth and development by activating or inhibiting gene expression.

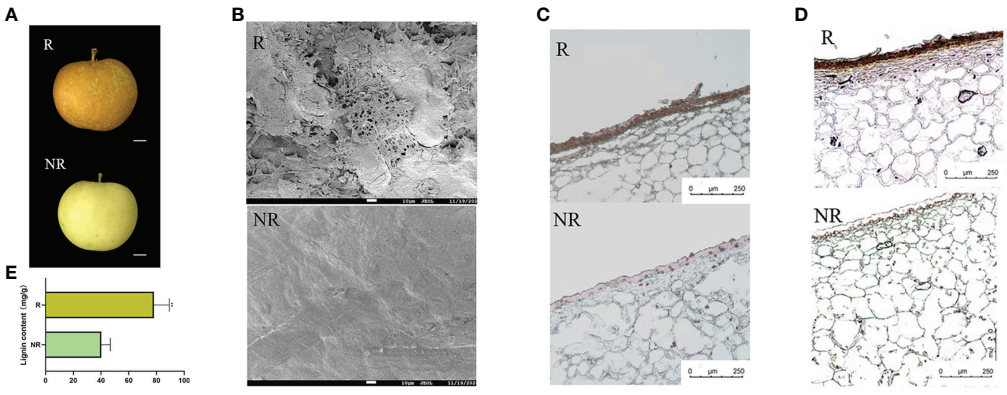


FIGURE 1
The different phenotypes between russet and non-russet apples at 120 DAFB. **(A)** Fruit phenotypes of the two apples. **(B)** SEM images of russet and non-russet apple fruit skins (scale bars: 10 μm). **(C)** Sudan red 7B staining of russet and non-russet fruit skins (scale bars: 250 μm). **(D)** Phloroglucinol staining (reddish color) of fruit skins (scale bars: 250 μm). **(E)** Lignin content between russet and non-russet fruits. Significance was calculated according to Student's *t-test*: $^{**}P < 0.01$.

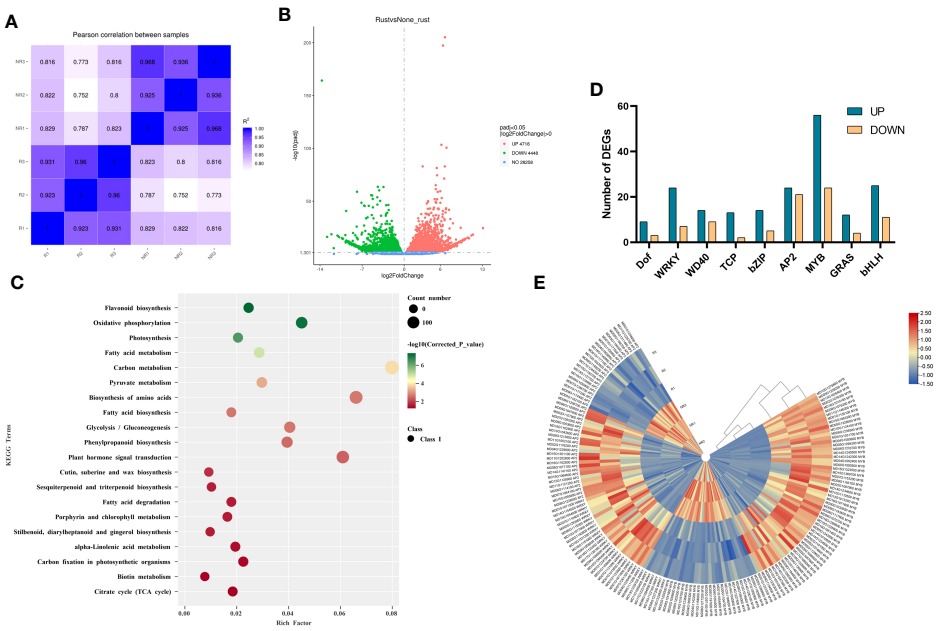


FIGURE 2
Identification and functional characterization of the DEGs in the two fruits. **(A)** Heatmap showing Pearson correlation coefficients between two samples. **(B)** Volcano plot filtering identified DEGs in the russet and non-russet fruit skin. **(C)** KEGG enrichment analysis of the DEGs in the two fruits. **(D)** Classification of transcription factors. **(E)** Expression heatmap of DEGs related to MYBs, WRKYs and AP2s.

DEG analysis identified 279 transcription factors, including nine common transcription factor families: Dof, WRKY, WD40, TCP, bZIP, AP2, MYB, GRAS, and bHLH (Figures 2D, E). Among them, MYB, WRKY, and AP2 were abundantly expressed in russet fruits, and we speculate that they may play an important role in fruit russeting. In the MYB transcription factor family, 56 genes were upregulated, and 24 genes were downregulated in russet fruit. *MYB6*, *MYB36*, *MYB66*, *MYB67*, *MYB73*, *MYB84*, *MYB93*, and *MYB4*, genes involved in suberin biosynthesis, were upregulated. In the WRKY transcription factor family, 24 transcription factors were upregulated, and seven were downregulated in russet apple. Furthermore, *WRKY13*, *WRKY24*, *WRKY51*, *WRKY56*, and *WRKY74* were significantly upregulated. In addition, 24 transcription factors were upregulated, and 21 were downregulated in the AP2 transcription factor family, these differential expressions in russet apples play a role in fruit russeting formation.

LC-MS and GC-MS profiling

To better explore the differentially accumulated metabolites (DAMs) between russet and non-russet pericarps, fruit peels were subjected to metabolome profiling *via* LC-MS and GC-MS. The DAMs were analyzed using the following parameters: VIP > 1, fold change > 1.5 or < 0.667, and *P* value < 0.05. In total, 165 DAMs were identified, of which 96 and 69 were upregulated and downregulated in russet apple, respectively (Figure 3A). All DAMs were divided into 20 categories (Table S2). GC-MS was used for targeted determination of the fatty acid content. A total of 32 fatty acids were detected, including medium- and long-chain fatty acids, 16 of which are involved in peel synthesis.

KEGG pathway enrichment analysis showed that the top enriched terms were metabolic pathways (map01100), purine metabolism (map00230), phenylpropanoid biosynthesis (map00940), flavonoid biosynthesis (map00941), and tropane, piperidine, and pyridine alkaloid biosynthesis (map00960) (Figure 3B). The results indicated that DAMs from different categories played important roles in fruit russet formation.

Interestingly, we found significant accumulation of cinnamic acid and its derivatives, phenylpropanoids, and phenolic acids in russet apples (Table S2). However, flavonoids, amino acids, and their derivatives were significantly enriched in both samples. Therefore, we noticed that during russeting formation, the plant secondary metabolism strengthened, and the metabolic changes in some amino acids and the flavonoids synthesis process also changed.

According to the GC-MS results (Figure 3C), we detected the main metabolites involved in cutin, suberine, and wax biosynthesis and biosynthesis of unsaturated fatty acids pathways in pericarp. C16:0 (palmitic acid), C16:1 (cis-9-palmitoleic acid), C16:2 (hexadecanedioic acid), C18:1 (cis-9-octadecenoic acid), and C22:2 (cis-13,16-docosadienoic acid) take part in suberin

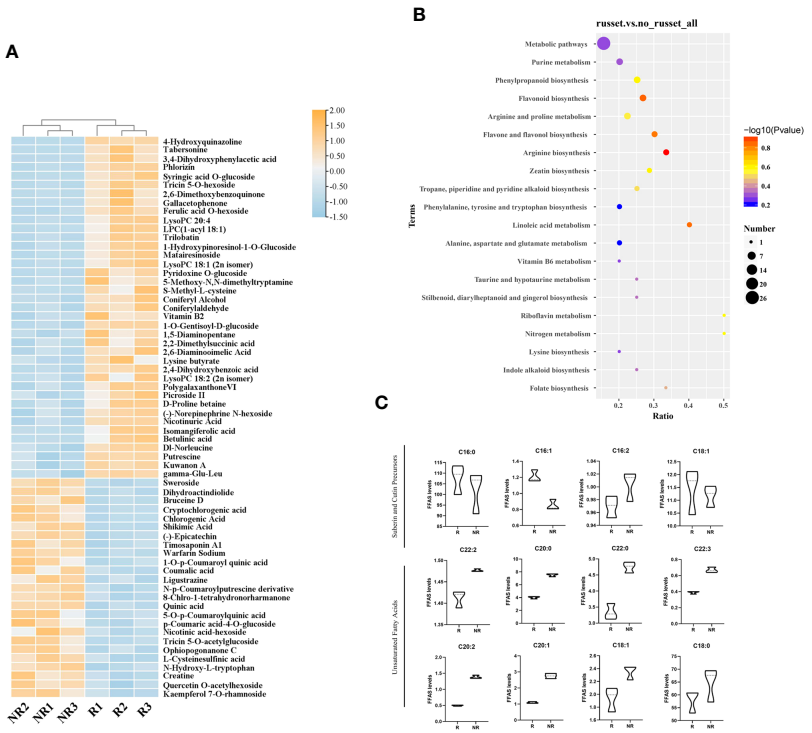
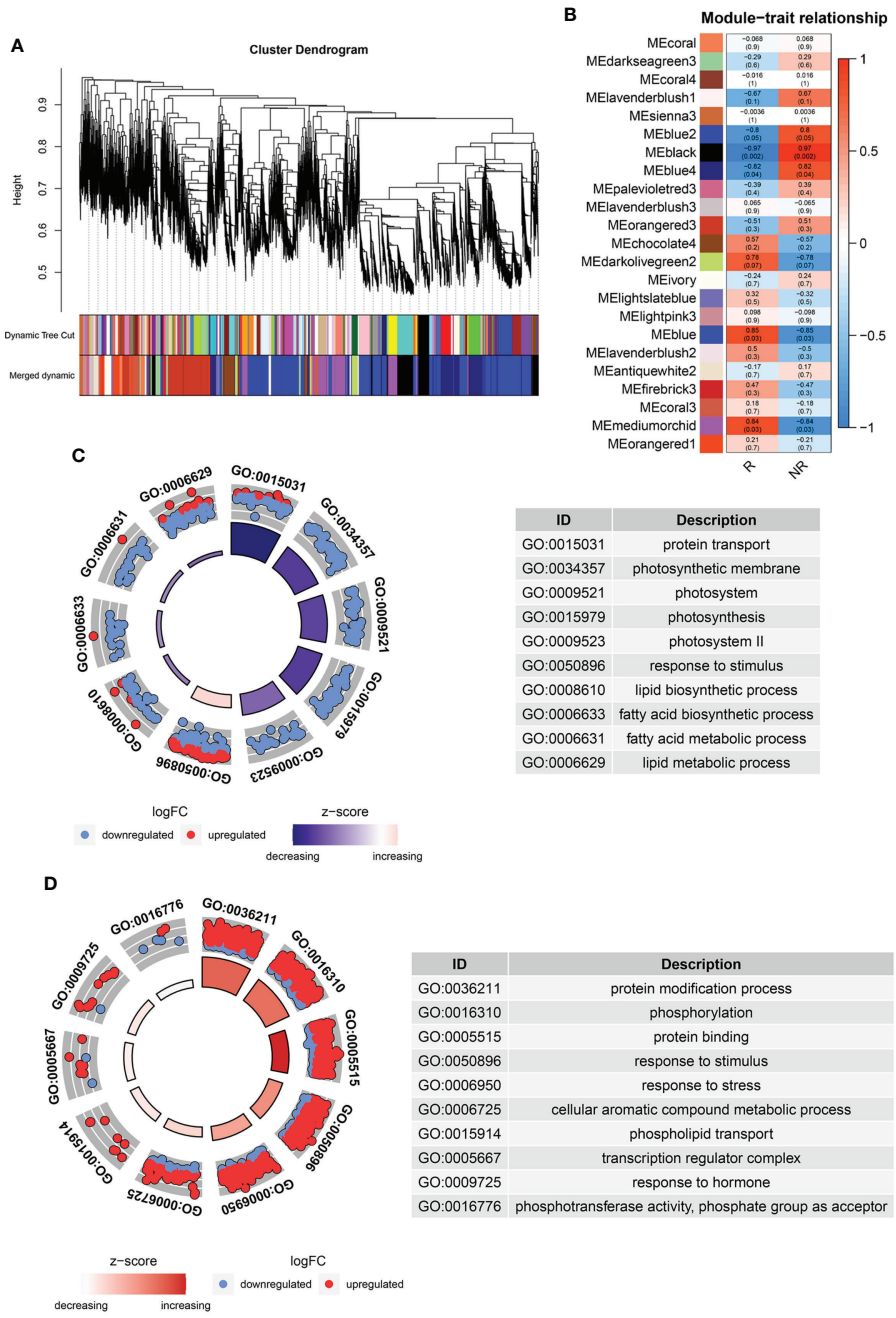


FIGURE 3
DAM analysis in russet and non-russet apples. **(A)** The heatmap shows the DAMs identified using LC-MS. Orange represents upregulated DAMs, and blue represents downregulated DAMs. **(B)** KEGG enrichment analysis of the DAMs in the two fruits. **(C)** Metabolite contents in the two fruits.

biosynthesis. C16:0, C16:1, and C18:1 were highly abundant in russet fruits, whereas C22:2 and C16:2 accumulated more in non-russet fruits. C20:0 (arachidic acid), C22:0 (behenic acid), C22:3 (erucic acid), C20:2 (cis-11,14-eicosadienoic acid), C20:1 (cis-11-eicosenoic acid), C18:0 (stearic acid), and C18:1 (trans-9-octadecenoic acid) participate in very long-chain fatty acid biosynthesis, thereby promoting wax synthesis and accumulation. The high levels of these substances may contribute to wax formation in non-russet fruits. The higher accumulation of waxy synthetic precursors in non-russet fruit, compared with russet fruit, suggests that the non-russet pericarp can form cuticles normally. We speculate that fruit russeting is a complex process that combines suberin accumulation and cuticle breakage.

Identification of weighted gene co-expression network analysis modules associated with russet and non-russet fruits

To identify the different gene expression and metabolite modules in the two kinds of fruits, we carried out a WGCNA of all DEGs with metabolites. The results identified 23 co-expression modules labeled with different colors in the cluster dendrogram (Figure 4A). Five modules (blue, blue2, blue4, black, and medium orchid) showed a significant correlation between the genes and metabolites ($P \geq 0.05$; Figure 4B). GO analysis of genes in the blue2, blue4 and black modules showed that lipid biosynthesis, lipid metabolism, fatty acid biosynthesis, and fatty acid metabolism were negatively related to fruit russeting (Figures 4C, S2B, D, E). This may be due to the breakdown of the peel cuticle in russet apples and the decrease in apple wax. We also found that some components of the photosystem (Figures 4C, S2B, D, E), indicating that light signals were involved in fruit russeting. Interestingly, the enriched processes detected in the blue and mediumorchid modules included phosphorylation and protein modification process, suggesting that protein modification may contribute to fruit russeting (Figures 4D, S2A, C). Furthermore, blue and mediumorchid modules also included response to stimulus, response to stress and response to hormone (Figures 4D, S2A, C), indicating that the environment and plant hormone may affect the formation of fruit russeting.



ID	Description
GO:0015031	protein transport
GO:0034357	photosynthetic membrane
GO:0009521	photosystem
GO:0015979	photosynthesis
GO:0009523	photosystem II
GO:0050896	response to stimulus
GO:0008610	lipid biosynthetic process
GO:0006633	fatty acid biosynthetic process
GO:0006631	fatty acid metabolic process
GO:0006629	lipid metabolic process

ID	Description
GO:0036211	protein modification process
GO:0016310	phosphorylation
GO:0005515	protein binding
GO:0050896	response to stimulus
GO:0006950	response to stress
GO:0006725	cellular aromatic compound metabolic process
GO:0015914	phospholipid transport
GO:0005667	transcription regulator complex
GO:0009725	response to hormone
GO:0016776	phosphotransferase activity, phosphate group as acceptor

FIGURE 4
WGCNA analysis of DAMs and DEGs. **(A)** Dendrogram with co-expressed gene modules. **(B)** Module-trait correlations and *P* values. The left panel shows 23 modules in different colors, and the right panel from blue to red represents the correlation between modules. GO analysis of blue2, blue4 and black modules **(C)**, blue and mediumorchid modules **(D)**.

Combined transcriptome and metabolome analysis and RNA-seq data validation

Apple fruit russeting involves many biosynthetic processes, including the suberin and lignin biosynthesis pathways (Jiang et al., 2022). We focused on the cuticle and suberin biosynthesis pathways, as well as phenylpropanoid metabolism. Moreover, the russeting process is regulated by several enzymes and genes. Based on the RNA-seq results, we identified 22 DEGs related to lignin biosynthesis and 62 DEGs related to suberin biosynthesis (Figures 5A, B).

The results showed that *PAL, C4H, 4CL, HCT, CCR, CAD, COMT*, and *CCoAOMT* were significantly differentially expressed (Figure 5A). Through LC-MS analysis, we found that phenylalanine, caffeic acid, caffeyl aldehyde, coniferyl aldehyde, and coniferyl alcohol were significantly enriched in russet samples (Figure 5C). Interestingly, 1-Caffeoylquinic acid and the gene expression of *hydroxylcinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase (HCT)* were both down-regulated in russeting fruits. The *HCT* is associated with *p-coumarate 3-hydroxylase (C3H)* and diverts the pathway from H-lignin to G- and S-lignin (Wagner et al., 2007). Based on the significant expressions of coniferyl aldehyde, coniferyl alcohol and sinapinic acid, the lignin type in fruit russeting is more likely to be G-S lignin.

We also detected some genes synthesized with suberin biosynthesis. The expression of *CER4*, *CYP86A1*, *CYP86A2*, *CYP86B1*, *CYP86B1*, *CYP86B2*, *ASFT*, *KCS4*, *GPAT5*, *GPAT6*, and *ABCGs* were up-regulated. However, *CER1*, *KCS6*, *KCS10*, *KCS19*, *CYP94B1*, *CYP86A2*, *CYP86A8*, *CYP77A4* and *GPAT3* were down-regulated (Figure 5B). We are also concerned about the differential accumulation of metabolites in suberin and cuticle, we found in LC-MS results high accumulation of betulinic acid, echinocystic acid, (9,10,13)-trihydroxyoctadec-11-enoic acid, and hexadecanedioic acid in russet fruit, while (9Z,11E)-octadecadienoic acid decreased. In addition, hexadecanedioic acid accumulated significantly. We identified significant *ASFT* expression with a high content of feruloyl-CoA (Figure 5B). ASFT usually catalyzes feruloyl CoA conversion into ω-OH-alkyl ferulates. The results showed a positive correlation between lignin synthesis and suberin content in this sample. The upregulated *C4H* expression is crucial. C4H greatly affects lignin synthesis and plays an important regulatory role in early lignin G/S differentiation (Hou et al., 2022). The downregulated *HCT* expression and 1-caffeoylquinic acid accumulation were also well demonstrated. In addition, the upregulated *COMT* and *CCoAOMT* expression also increased ferulic acid and coniferaldehyde accumulation and promoted lignin synthesis in russet fruit. In the suberin and cuticle synthesis pathway, we identified upregulated genes that promote suberin synthesis, including *CER4*, *CYP86As*, *ASFT*, and *KCS4*. Interestingly, some genes were downregulated in this pathway (mdm00730), such as *CER1* and *KCS6*, accompanied by downregulated expression of long-chain fatty acid synthesis precursors (Figure 5D). We speculate that lignin and suberin synthesis increased with fruit russeting, whereas cuticle and wax synthesis decreased.

We then performed qRT-PCR to determine the expression of these 12 genes and found that the quantitative results were consistent with the transcriptome results, indicating that the transcriptome results were credible (Figure 5E).

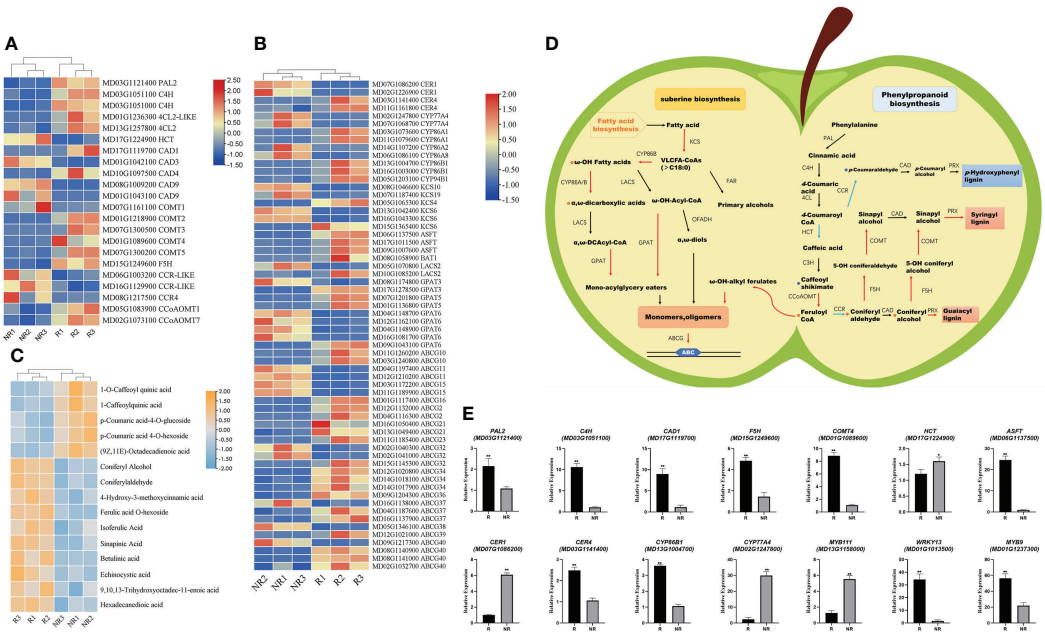


FIGURE 5
Integrated analysis of metabolomic and transcriptomic profiling. **(A)** DEGs involved in the lignin synthesis pathway. **(B)** DEGs involved in the suberin and cutin synthesis pathway. **(C)** Potential DAMs involved in fruit russeting. **(D)** Proposed model of russeting in apples (Red and blue arrows indicate upregulated and downregulated genes in the pathway, respectively; red and blue circles indicate the increased and reduced metabolites, respectively). **(E)** qRT-PCR validation of the transcriptome data; significance was calculated according to Student's *t*-test: $*P < 0.05$, $**P < 0.01$.

Plant hormone signals and expression of stress-related genes

Fruit russeting is affected by the environment and hormones. In this study, genes involved in hormone signaling and stress responses were analyzed in addition to the genes related to lignin and suberin biosynthesis. We identified the expression of multiple phytohormone pathway genes (Figure 6A). The results showed that the auxin synthesis pathway genes *IAA* and *AUX* were highly expressed in russet fruits. The cytokinin (CK) pathway genes *AHK* and *PYR1* were also expressed at higher levels in russet samples. *PP2C* in the abscisic acid (ABA) pathway was highly expressed in russet samples. *BSK* and *CYCD* were highly expressed in the brassinolide (BR) pathway. In addition, *JAR1, JAZ1, TGA9*, and *TGA10* were upregulated and involved in the methyl jasmonate (MeJA) pathway. Interestingly, some ERFs (*ERF1* and *ERF54*) in the ethylene synthesis pathway were downregulated in the russet samples.

To study the response of russet fruits to external environmental stimuli, we analyzed gene expression abundance under abiotic stress (Figure 6B). Heat shock proteins (HSPs), aquaporins (AQPs), and peroxidases (PRXs) exhibited different expression levels. We found that HSP (HSP70, *HSP70-1*, and *HSP90-1*), AQP (*PIP1, PIP2, SIP1*, and *TIP1*), and PRX (*PRXQ, PRX16, PRX47, PRX52, PRX53*, and *PRX72*) genes were upregulated in russet fruits.

Discussion

Physiological characteristics and composition analysis of fruit russeting

The skin of russet fruits is brown, rough, and covered with cracks. SEM showed that the cuticle was dramatically reduced, leading to a russet phenotype. For example, fruit russeting gradually appeared as cracks during the development period, although the integrity of the russeting cuticle was still maintained at 31 DAFB in apples (Falginella et al., 2021). SEM observations revealed V-shaped cracks in the peel of 150 DAFB russet apple (Yuan et al., 2019). Non-russet pears were smooth, with no cracks, but cuticle cracks were evident in russet pears (Shi et al., 2021). In this study, we found that russet apple skin was rough, whereas that of other apple was smooth (Figure 1A). SEM showed that the russet fruit surface was covered with cracks and gaps and the cuticle was broken, while the waxy layer of the non-russet peel was intact (Figure 1B). In addition, sections of russet kiwifruit and pear pericarp showed higher suberin and lignin contents than those of non-russet fruit (Wang et al., 2020; Macnee et al., 2021). These findings are consistent with the results of our study.

In previous studies, metabolome profiling of sand pears revealed that phenols, lignin, ω-unsaturated fatty acids, α,ω-diacids, ω-OH fatty acids, and glycerides were enriched in russet pears (Shi et al., 2021). These are important components of

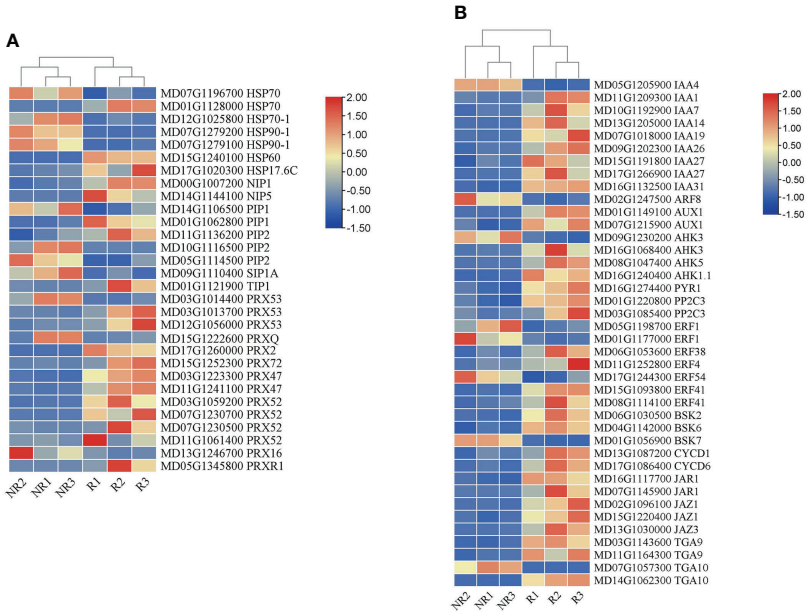


FIGURE 6
Expression heatmap of DEGs related to stress responses **(A)** and phytohormones **(B)** between the two fruit skins.

suberin and lignin. In 'Sunshine Muscat' grapes, lignin and quercetin play a role in berry russeting (Huang et al., 2020). Ursane-type triterpenes are predominant in waxy apple cuticles, but they shift towards lupane-type triterpenes with russet formation (Falginella et al., 2021). In this study, we detected many compounds that may influence fruit russeting. We found that C16:0, C16:1, and C18:1 contents were high in russet fruits, while C22:2 and C16:2 were higher in non-russet fruits (Figure 3C). This indicated that russeting decreased long-chain fatty acid biosynthesis and increased suberin biosynthesis. Moreover, two lignin synthesis-related metabolites, ferulic acid and coniferaldehyde, accumulated in russet apples in this study (Figure 5C). These results demonstrate that higher lignin and suberin contents contribute to fruit russeting in apples.

Molecular mechanism regulating fruit russeting

RNA-seq analysis was performed to reveal the molecular regulatory basis underlying russet fruits with contrasting phenotypes. Previous studies have shown that key genes, such as *CYP, KCS, GPAT, FAR*, and *ABCG*, play an important role in peel russeting (Pollard et al., 2008). In addition, key genes of the phenylpropane metabolism pathway also have an important effect on fruit russeting (Hou et al., 2022). The AP2-type transcription factor MdSHN3 promotes cuticle formation and inhibits fruit russet formation in apple pericarps (Lashbrooke et al., 2015). The transcription factor MdMYB66 can bind *MdOSC5* to promote lupane-type triterpene formation in russet fruits while reducing the ursane-type and oleanane-type formation, causing failure of the early development of apple cuticles and increasing fruit russeting (Falginella et al., 2021). A study on russet apple peels found that the normal development of apple fruit skin produces lenticels, which produce suberin (Straube et al., 2021). Moreover, the chemical phenotypes observed in most of the studied suberin mutants suggest the involvement of partially redundant enzymes from the CYP, GPAT, FAR, and KCS families (Li et al., 2007). In our study, many genes showed strong connections with suberin biosynthesis. For example, we found that *KCS1*, *KCS2*, and *KCS4* were upregulated, while *KCS6, KCS10, KCS11*, and *KCS19* were downregulated (Figure 5B). *KCS2* and *KCS20* promote suberin deposition in the periderm, whereas *KCS6* and *KCS11* promote wax formation (Lee et al., 2009; Lian et al., 2021; Li et al., 2021; Huang et al., 2022). These results indicate that the different expressions of *KCS2* and *KCS6* in our study may cause suberin increase and wax reduction. In addition, *PbABCGs* were always expressed at a high level in the 'Dangshan Su' russet pear mutant during different development periods (Hou et al., 2018). A putative ABCG family transporter has been identified near a significant quantitative trait locus (QTL) for apple fruit russeting on LG12, which is not collinear with LG8 in sand pear (Falginella et al., 2015). *ABCG2*, *ABCG6*, and *ABCG20* are involved in suberin synthesis in the roots and seed coats (Yadav et al., 2014). In this study, we identified 13 ABCGs with different expression patterns. The upregulated expression of *ABCG2*, *ABCG21*, *ABCG23*, *ABCG34*, *ABCG36*, *ABCG37*, *ABCG39*, and *ABCG40* in russet apples may promote fruit russeting (Figure 5B). GPAT is a major determinant of cuticle and suberin composition, catalyzing glycerol-3-phosphate acylation (Yang et al., 2012). *GPAT4* and *GPAT6* primarily esterify acyl groups at the sn-2 position of glycerol-3-phosphate during *Arabidopsis* cutin biosynthesis (Yang et al., 2010). In our study, GPAT6 expression was lower, suggesting that cutin synthesis is attenuated in russet fruit (Figure 5B).

Lignin is an important secondary metabolite that affects fruit russetting. PAL is a key enzyme in lignin synthesis, and the transcription factor LIM can interact with PAL-like enzymes to promote lignin synthesis in fruit russeting (Yuan et al., 2019). Studies on 'Dangshan Su' pear mutant varieties showed that increased *CCoAOMT* expression resulted in increased lignin content in the outer skin, which led to russet formation (Heng et al., 2016). Research on russet apples has highlighted that lignin synthesis genes, such as *COMT* and *C4H*, are expressed more frequently, contributing to russeting (Legay et al., 2015). In our study, we also found that the lignin synthesis genes *PAL*, *C4H*, *F5H*, *COMT*, and *CCoAOMT* were upregulated in russet apples. This indicated that lignification is an important process in fruit russeting. In addition, *CCR* and *CAD9* expression in russet apples was downregulated, consistent with previous results (Legay et al., 2015).

Fruit russeting in response to external environmental stress

Environmental factors, such as light, temperature, moisture, and endogenous hormones, influence russeting. In a previous study, light intensity influenced fruit russeting by affecting the levels of endogenous gibberellins (Noè and Eccher, 1996). In pears, rainfall treatment aggravated russeting both in 'Zaoshengxinshui' and 'Cuiguan' pears (Shi et al., 2021). Taken together, russeting formation may be a self-preservation mechanism in fruits. Thus, we analyzed the HSPs, AQPs, and PRXs involved in the response to biotic and abiotic stresses in the transcriptome data. We found three upregulated types of HSPs in russet apples (Figure 6A). Because HSPs respond to heat shock, higher HSP expression would increase heat tolerance in russet pears (Shi et al., 2021).

In addition, AQPs are non-enzymatic proteins that play roles in water stress by changing the cell wall structure and water channels of the cell membrane (Secchi et al., 2017). Three AQP genes, *PbTIP1-3*, *PbTIP2-1*, and *PbPIP2-5*, were highly expressed in rainfall-treated sand pears, indicating that these genes may be

involved in rainfall-aggravated russeting (Shi et al., 2021). In our study, we also identified nine AQP-related genes that were differentially expressed in the two apple phenotypes. *PIP1* and *PIP2* were expressed at higher levels in russet peels and may play an important role in regulating water transport against moisture (Figure 6A). Russeting was accompanied by cuticle damage, which could easily lead to water loss. Therefore, the high expression of AQP genes in russet fruits could explain the cuticle damage caused by russet formation, which may accelerate fruit water loss. We found that most *PRX* genes were highly expressed in russet apples (Figure 6A). PRX has been suggested to play a role in auxin metabolism, lignification, suberization, cross-linking of cell wall components, and defense against pathogens (Quiroga et al., 2000). This result indicated that *PRX* genes promoted lignin and suberin synthesis, enhancing the response to stress in russet apples.

Plant hormones play an important role in fruit russeting, which can be reduced by hormone application (Knoche et al., 2011). However, their direct effect on fruit russeting is not clear; thus, the molecular mechanisms, by which hormones affect fruit russeting, should be studied in the future. A total of 216 DEGs were identified and involved in pear browning spot formation, including the IAA, CTK, GA, ABA, JA, and SA pathways, showing that GA might promote browning-spot formation (Wang et al., 2020). In pears, 19 DEGs related to plant hormone biosynthesis and signaling were identified, including ABA, GA, IAA, and ethylene. ABA content was higher in russet pears, and ABA treatment could increase russeting (Wang et al., 2022). Many genes related to auxin and ethylene signaling pathways are highly expressed in russet apples (Legay et al., 2015). Moreover, it is believed that there are fewer DEGs for salicylic acid and jasmonic acid in russet fruits, and there may not be a common wounding and biotic stress response (Legay et al., 2015). In this study, we found that the IAA, CK, ABA, BR, MeJA, and ethylene pathways were significantly different (Figure 6B). A large number of auxin pathway genes were enriched, including *IAA1*, *IAA7*, *IAA14*, *IAA19*, *IAA26*, *IAA27*, *IAA27*, *IAA31*, *ARF8*, and *AUX1*. In addition, we identified higher expression levels of the CK and ABA pathway genes *AHK3*, *AHK5*, and *PP2C2*. We found that genes of the IAA pathway were expressed in the highest numbers in russet fruits and speculated that IAA may play an important role in fruit russeting. In this study, we first discovered the existence of multilayered cell stacks in russet peels (Figure 1D). It is speculated that IAA and CK promote russeting by controlling cell differentiation and proliferation, while ABA and MeJA further promote secondary metabolism by participating in defense responses. Moreover, a large number of ethylene pathway genes were identified in this study, and the expression of various ERFs was downregulated; however, the role of this downregulation in fruit russeting is unclear. In addition, we found upregulated expression of *BSK* and *CYCD* genes in the BR pathway. The relationship between BR and fruit russeting remains uncertain, which provides a new idea for future research.

Transcription factors involved in fruit russeting

Although fruit russeting involves metabolic synthesis and is influenced by the environment, transcription factor regulation is also important in russet formation. In plants, various MYB transcription factors regulate either suberin biosynthesis or extracellular lipid accumulation. In pears, 12 MYB genes involved in plant tissue suberization were differentially upregulated in the russet fruit skin, indicating that these MYB genes may affect russeting formation by regulating suberin synthesis (Ma et al., 2021). In addition, MYB9 and MYB107 (Lashbrooke et al., 2016), MYB6 and MYB67 (Falginella et al., 2021), and MYB93 (Legay et al., 2016) involvement has been reported in suberin in fruit russeting. Interestingly, several upregulated MYB transcription factors were identified in this study. In addition to transcription factors whose synthesis has previously been reported in fruit russeting and suberin, we found that MYB73 and MYB84 also had high levels of differential expression.

WRKY transcription factors can bind to stress-related promoters to regulate plant tolerance to abiotic stresses. According to the transcriptome results, we identified 31 WRKY transcription factors, only seven of which were downregulated. Among them, WRKY13 expression was the highest and its expression was higher in russet apple than that in non-russet apple (Figure 5E). AtWRKY13 was identified as a positive factor for lignin synthesis, and *AtWRKY13* overexpression could improve the mechanical strength of *Arabidopsis* stems (Li et al., 2015). Thus, WRKY13 may be a positive factor for lignin synthesis during russeting.

Conclusion

In this study, the high lignin and suberin content led to russet formation in the skin of the russet apple strain. Transcriptome and metabolome analysis revealed that the expression of suberin and lignin synthesis genes was upregulated in the skin of russet apples, and the intermediate metabolites involved in both pathways were abundantly accumulated. These results indicate a positive correlation between lignin and suberin synthesis and fruit russeting. Taken together, lignin and suberin accumulation in fruit skin is an important reason for fruit russet formation.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: NCBI, PRJNA871277.

Author contributions

SJ and YZ designed this experiment and revised the manuscript, ZW and SJ analyzed the data and wrote the manuscript. SL and WH helped to perform the qRT-PCR; MC helped ZW to analyzed the transcriptome data. All authors contributed to the article and approved the submitted version.

Funding

This work is supported by the National Natural Science Foundation of China (32102319), Natural Science Foundation of Shandong Province (ZR2021QC010), National Key Research and Development Program Foundation (2019YFD1001403), Taishan Scholar Foundation of Shandong Province (ts2022), Agricultural Variety Improvement Project of Shandong Province (2021LZGC024), Talents of High-level Scientific Research Foundation (6651121002) and Graduate Innovation Program of Qingdao Agricultural University (QNYCX21089).

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1057226/full#supplementary-material

References

Andre, C. M., Larsen, L., Burgess, E. J., Jensen, D. J., Cooney, J. M., Evers, D., et al. (2013). Unusual immuno-modulatory triterpene-caffeates in the skins of russeted varieties of apples and pears. *J. Agric. Food Chem.* 61, 2773–2779. doi: 10.1021/jf305190e

Falginella, L., Andre, C. M., Legay, S., Lin-Wang, K., Dare, A. P., Deng, C., et al. (2021). Differential regulation of triterpene biosynthesis induced by an early failure in cuticle formation in apple. *Hortic. Res.* 8, 75. doi: 10.1038/s41438-021-00511-4

Falginella, L., Cipriani, G., Monte, C., Gregori, R., Testolin, R., Velasco, R., et al. (2015). A major QTL controlling apple skin russeting maps on the linkage group 12 of 'Renetta grigia di torriana. *BMC Plant Biol.* 15, 150. doi: 10.1186/s12870-015-0507-4

Heng, W., Wang, M., Yang, J., Wang, Z., Jiang, X., and Zhu, L. (2016). Relationship between H2O2 in polyamine metabolism and lignin in the exocarp of a russet mutant of 'Dangshansuli' pear (Pyrus bretschneideri rehd.). *Plant Mol. Biol. Rep.* 34, 1056–1063. doi: 10.1007/s11105-016-0985-z

Hou, Z., Jia, B., Li, F., Liu, P., Liu, L., Ye, Z., et al. (2018). Characterization and expression of the ABC family (G group) in 'Dangshansuli' pear (*Pyrus bretschneideri* rehd.) and its russet mutant. *Genet. Mol. Biol.* 41, 137–144. doi: 10.1590/1678-4685-GMB-2017-0109

Hou, D., Lu, H., Zhao, Z., Pei, J., Yang, H., Wu, A., et al. (2022). Integrative transcriptomic and metabolomic data provide insights into gene networks associated with lignification in postharvest lei bamboo shoots under low temperature. *Food Chem.* 368, 130822. doi: 10.1016/j.foodchem.2021.130822

Huang, H., Ayaz, A., Zheng, M., Yang, X., Zaman, W., Zhao, H., et al. (2022). Arabidopsis KCS5 and KCS6 play redundant roles in wax synthesis. *Int. J. Mol. Sci.* 23, 4450. doi: 10.3390/ijms23084450

Huang, Y., Liang, D., Xia, H., Lin, L.-J., Wang, J., and Lv, X.-L. (2020). Lignin and quercetin synthesis underlies berry russeting in 'Sunshine muscat' grape. *Biomolecules* 10, 690. doi: 10.3390/biom10050690

Jiang, S., Chen, M., Wang, Z., Ren, Y., Wang, B., Zhu, J., et al. (2022). Advances in understanding the causes, molecular mechanism, and perspectives of russeting on tree fruit. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.834109

Khanal, B. P., Grimm, E., and Knoche, M. (2013). Russeting in apple and pear: a plastic periderm replaces a stiff cuticle. *AoB Plants* 5, pls048. doi: 10.1093/aobpla/pls048

Knoche, M., Khanal, B. P., and Stopar, M. (2011). Russeting and microcracking of 'Golden delicious' apple fruit concomitantly decline due to gibberellin A4+7 application. *J. Am. Soc Hortic. Sci. J. Amer Soc. Hort Sci.* 136, 159–164. doi: 10.21273/JASHS.136.3.159

Lashbrooke, J., Aharoni, A., and Costa, F. (2015). Genome investigation suggests MdSHN3, an APETALA2-domain transcription factor gene, to be a positive regulator of apple fruit cuticle formation and an inhibitor of russet development. *J. Exp. Bot.* 66, 6579–6589. doi: 10.1093/jxb/erv366

Lashbrooke, J., Cohen, H., Levy-Samocha, D., Tzfadia, O., Panizel, I., Zeisler, V., et al. (2016). MYB107 and MYB9 homologs regulate suberin deposition in angiosperms. *Plant Cell* 28, 2097–2116. doi: 10.1105/tpc.16.00490

Lee, S.-B., Jung, S.-J., Go, Y.-S., Kim, H.-U., Kim, J.-K., Cho, H.-J., et al. (2009). Two arabidopsis 3-ketoacyl CoA synthase genes, KCS20 and KCS2/DAISY, are functionally redundant in cuticular wax and root suberin biosynthesis, but differentially controlled by osmotic stress. *Plant J.* 60, 462–475. doi: 10.1111/j.1365-313X.2009.03973.x

Legay, S., Guerriero, G., André, C., Guignard, C., Cocco, E., Charton, S., et al. (2016). MdMyb93 is a regulator of suberin deposition in russeted apple fruit skins. *New Phytol.* 212, 977–991. doi: 10.1111/nph.14170

Legay, S., Guerriero, G., Deleruelle, A., Lateur, M., Evers, D., André, C. M., et al. (2015). Apple russeting as seen through the RNA-seq lens: strong alterations in the exocarp cell wall. *Plant Mol. Biol.* 88, 21–40. doi: 10.1007/s11103-015-0303-4

Lian, X.-Y., Gao, H.-N., Jiang, H., Liu, C., and Li, Y.-Y. (2021). MdKCS2 increased plant drought resistance by regulating wax biosynthesis. *Plant Cell Rep.* 40, 2357–2368. doi: 10.1007/s00299-021-02776-4

Li, Y., Beisson, F., Koo, A. J. K., Molina, I., Pollard, M., and Ohlrogge, J. (2007). Identification of acyltransferases required for cutin biosynthesis and production of cutin with suberin-like monomers. *Proc. Natl. Acad. Sci.* 104, 18339–18344. doi: 10.1073/pnas.0706984104

Li, Z., Ma, S., Song, H., Yang, Z., Zhao, C., Taylor, D., et al. (2021). A 3-ketoacyl-CoA synthase 11 (KCS11) homolog from malania oleifera synthesizes nervonic acid in plants rich in 11Z-eicosenoic acid. *Tree Physiol.* 41, 331–342. doi: 10.1093/treephys/tpaa125

Li, W., Tian, Z., and Yu, D. (2015). WRKY13 acts in stem development in arabidopsis thaliana. *Plant Sci.* 236, 205–213. doi: 10.1016/j.plantsci.2015.04.004

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

Macnee, N., Hilario, E., Tahir, J., Currie, A., Warren, B., Rebstock, R., et al. (2021). Peridermal fruit skin formation in actinidia sp. (kiwifruit) is associated with genetic loci controlling russeting and cuticle formation. *BMC Plant Biol.* 21, 334. doi: 10.1186/s12870-021-03025-2

Ma, C., Wang, X., Yu, M., Zheng, X., Sun, Z., Liu, X., et al. (2021). PpMYB36 encodes a MYB-type transcription factor that is involved in russet skin coloration in pear (Pyrus pyrifolia). *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.776816

Noè, N., and Eccher, T. (1996). 'Golden delicious' apple fruit shape and russeting are affected by light conditions. *Sci. Hortic.* 65, 209–213. doi: 10.1016/0304-4238(95)00850-0

Pollard, M., Beisson, F., Li, Y., and Ohlrogge, J. B. (2008). Building lipid barriers: biosynthesis of cutin and suberin. *Trends Plant Sci.* 13, 236–246. doi: 10.1016/j.tplants.2008.03.003

Quiroga, M., Guerrero, C., Botella, M. A., Barceló, A., Amaya, I., Medina, M. I., et al. (2000). A tomato peroxidase involved in the synthesis of lignin and Suberin1. *Plant Physiol.* 122, 1119–1128. doi: 10.1104/pp.122.4.1119

Secchi, F., Pagliarani, C., and and Zwieniecki, M. A. (2017). The functional role of xylem parenchyma cells and aquaporins during recovery from severe water stress. *Plant Cell Environ.* 40, 858–871. doi: 10.1111/pce.12831

Shi, C.-H., Wang, X.-Q., Xu, J.-F., Zhang, Y.-X., Qi, B., and Jun, L. (2021). Dissecting the molecular mechanism of russeting in sand pear (Pyrus pyrifolia nakai) by metabolomics, transcriptomics, and proteomics. *Plant J.* 108, 1644–1661. doi: 10.1111/tpj.15532

Straube, J., Chen, Y.-H., Khanal, B. P., Shumbusho, A., Zeisler-Diehl, V., Suresh, K., et al. (2021). Russeting in apple is initiated after exposure to moisture ends: Molecular and biochemical evidence. *Plants* 10, 65. doi: 10.3390/plants10010065

Wagner, A., Ralph, J., Akiyama, T., Flint, H., Phillips, L., Torr, K., et al. (2007). Exploring lignification in conifers by silencing hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyltransferase in pinus radiata. *Proc. Natl. Acad. Sci.* 104, 11856–11861. doi: 10.1073/pnas.0701428104

Wang, Y., Dai, M., Cai, D., and Shi, Z. (2020). Proteome and transcriptome profile analysis reveals regulatory and stress-responsive networks in the russet fruit skin of sand pear. *Hortic. Res.* 7, 16. doi: 10.1038/s41438-020-0242-3

Wang, Q., Liu, Y., Wu, X., Wang, L., Li, J., Wan, M., et al. (2022). MYB1R1 and MYC2 regulate ω-3 fatty acid desaturase involved in ABA-mediated suberization in the russet skin of a mutant of 'Dangshansuli' (Pyrus bretschneideri rehd.). *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.910938

Yadav, V., Molina, I., Ranathunge, K., Castillo, I. Q., Rothstein, S. J., and Reed, J. W. (2014). ABCG transporters are required for suberin and pollen wall extracellular barriers in arabidopsis. *Plant Cell* 26, 3569–3588. doi: 10.1105/tpc.114.129049

Yang, W., Pollard, M., Li-Beisson, Y., Beisson, F., Feig, M., and Ohlrogge, J. (2010). A distinct type of glycerol-3-phosphate acyltransferase with sn-2 preference and phosphatase activity producing 2-monoacylglycerol. *Proc. Natl. Acad. Sci.* 107, 12040–12045. doi: 10.1073/pnas.0914149107

Yang, W., Simpson, J. P., Li-Beisson, Y., Beisson, F., Pollard, M., and Ohlrogge, J. B. (2012). A land-Plant-Specific glycerol-3-Phosphate acyltransferase family in arabidopsis: Substrate specificity, sn-2 preference, and evolution. *Plant Physiol.* 160, 638–652. doi: 10.1104/pp.112.201996

Yuan, G., Bian, S., Han, X., He, S., Liu, K., Zhang, C., et al. (2019). An integrated transcriptome and proteome analysis reveals new insights into russeting of bagging and non-bagging "Golden delicious". *Apple Int. J. Mol. Sci.* 20, 4462. doi: 10.3390/ijms20184462

Zhang, Y.-L., You, C.-X., Li, Y.-Y., and Hao, Y.-J. (2020). Advances in biosynthesis, regulation, and function of apple cuticular wax. *Front. Plant Sci.* 11, 1165. doi: 10.3389/fpls.2020.01165

Zhang, J., Zhang, Y.-F., Zhang, P.-F., Bian, Y.-H., Liu, Z.-Y., Zhang, C., et al. (2021). An integrated metabolic and transcriptomic analysis reveals the mechanism through which fruit bagging alleviates exocarp semi-russeting in pear fruit. *Tree Physiol.* 41, 1306–1318. doi: 10.1093/treephys/tpaa172

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 08 December 2022
DOI 10.3389/fpls.2022.1073332

OPEN ACCESS

EDITED BY
Chenxia Cheng,
Qingdao Agricultural University, China

REVIEWED BY
Minjie Qian,
Hainan University, China
Yu Zong,
Zhejiang Normal University, China

*CORRESPONDENCE
Haishan An
anhaishan@saas.sh.cn
Xueying Zhang
zhangxueying@saas.sh.cn

†These authors have contributed equally to the work

SPECIALTY SECTION
This article was submitted to Functional and Applied Plant Genomics,
a section of the journal Frontiers in Plant Science

RECEIVED 18 October 2022
ACCEPTED 21 November 2022
PUBLISHED 08 December 2022

CITATION
Zhang J, Li S, An H, Zhang X and Zhou B (2022) Integrated transcriptome and metabolome analysis reveals the anthocyanin biosynthesis mechanisms in blueberry (*Vaccinium corymbosum* L.) leaves under different light qualities.
Front. Plant Sci. 13:1073332.
doi: 10.3389/fpls.2022.1073332

Integrated transcriptome and metabolome analysis reveals the anthocyanin biosynthesis mechanisms in blueberry (*Vaccinium corymbosum* L.) leaves under different light qualities

Jiaying Zhang[1,2†], Shuigen Li[1,2†], Haishan An[1,2*], Xueying Zhang[1,2*] and Boqiang Zhou[1,2]

[1]Forestry and Pomology Research Institute, Shanghai Academy of Agricultural Sciences, Shanghai, China, [2]Shanghai Key Lab of Protected Horticultural Technology, Shanghai Academy of Agricultural Sciences, Shanghai, China

Introduction: Blueberry (*Vaccinium corymbosum* L.) is a popular fruit with an abundance of anthocyanins in its leaves and fruits. Light is one of the pivotal environmental elements that affects plant growth and development, but the regulatory mechanism between light quality and anthocyanin formation is poorly understood.

Methods: An integrated transcriptome and metabolome analysis was performed to investigate the effects of white (control), blue (B), red (R), and red/blue (60R/40B) light on blueberry growth and reveal the potential pathway controlling anthocyanin biosynthesis in blueberry leaves.

Results: The anthocyanin content was significantly improved by the blue and red/blue light when compared with white light, whereas there was a significant reduction in the photosynthesis under the blue light, showing an inverse trend to that of anthocyanin accumulation. Transcriptomic analysis resulted in the assembly of 134,709 unigenes. Of these, 22 were differentially expressed genes (DEGs) that participate in the anthocyanin biosynthesis pathway, with the majority being significantly up-regulated under the blue light. Most of the photosynthesis-related genes that were down-regulated were expressed during anthocyanin accumulation. Targeted metabolome profiling identified 44 metabolites associated with anthocyanin biosynthesis. The contents of most of these metabolites were higher under blue light than the other light conditions, which was consistent with the transcriptome results. The integrated transcriptome and metabolome analysis suggested that, under blue light, leucoanthocyanidin dioxygenase (LDOX), O-methyltransferase (OMT), and UDP-glucose flavonoid glucosyltransferase (UFGT) were the most significantly expressed, and they promoted the synthesis of cyanidin (Cy),

malvidin (Mv), and pelargonidin (Pg) anthocyanidins, respectively. The expression levels of dihydroflavonol 4-reductase (DFR) and OMT, as well as the accumulation of delphinidin (Dp), peonidin (Pn), and petunidin (Pt), were significantly increased by the red/blue light.

Discussion: The blue and red/blue lights promoted anthocyanin biosynthesis via inducing the expression of key structural genes and accumulation of metabolites involved in anthocyanin synthesis pathway. Moreover, there was a possible feedback regulating correlation between anthocyanin biosynthesis and photosynthesis under different light qualities in blueberry leaves. This study would provide a theoretical basis for elucidating the underlying regulatory mechanism of anthocyanin biosynthesis of V. corymbosum.

KEYWORDS

Vaccinium corymbosum L., anthocyanin, light quality, transcriptome, metabolome

Introduction

Blueberry (*Vaccinium corymbosum* L.) is a fruit-bearing shrub in the genus *Vaccinium*, which is part of the heath family (Ericaceae). It is a commercially important small fruit crop owing to its healthy and flavorful bioactive compounds, including vitamins, anthocyanins, and other phenolic compounds (Zhang L. et al., 2019; Kalt et al., 2020). Blueberries are inherently high in anthocyanins and can be used to fight some human cancers (Seeram et al., 2006) and regulate oxidative stress in organs (Song et al., 2016), blood glucose levels (Fallah et al., 2020a), and inflammatory responses (Fallah et al., 2020b). Due to these potential health benefits, interest in blueberry anthocyanins has been increasing (Chai et al., 2021). The previous studies on anthocyanin biosynthesis have mainly focused on blueberry fruits and discovered some crucial differentially expressed genes or transcription factors related to anthocyanin biosynthesis pathway (Lin et al., 2018), but reports on anthocyanin accumulation in the leaves have been limited. Leaves are abundant in metabolites and the site of secondary plant metabolite production. Anthocyanins are one of the most important classes of secondary metabolite, and the anthocyanins in leaves may promote plant growth and development (Guo et al., 2020). Secondary metabolites, including flavonoids and anthocyanins, extracted from *Lithocarpuspo lystachyus Rehd* and blueberry leaves, can play a crucial role in the treatment of diabetes, hypertension, and chronic illnesses (Ehlenfeldt and Prior, 2001; Takeshita et al., 2009; Hou et al., 2011; Hou et al., 2012; Li et al., 2021). Anthocyanins in blueberry leaves may also be involved in the scavenging of reactive oxygen species (ROS), which could enhance disease resistance in humans (Sakaida et al., 2007). Additionally, the blueberry leaf is a commonly used drug for treating human thrombotic stroke in clinical trials (Nagao et al., 2008). Nevertheless, the true value of the blueberry plant including its leaves has not been fully explored yet, resulting in the unnecessary waste of this resource. Exploring and studying nutritious compounds, such as the anthocyanins in blueberry leaves, will lead to more efficient utilization of fruit resources and help to satisfy the growing demand for natural foods and medicines. Therefore, there is substantial interest in understanding the anthocyanin metabolism in blueberry leaves. Regrettably, our current understanding of the molecular mechanisms underlying anthocyanin biosynthesis is limited and how environmental factors regulate anthocyanin accumulation in blueberry leaves is unclear.

Anthocyanins are the most conspicuous class of flavonoid metabolic branches involved in the phenylpropane metabolic pathway, which presents numerous secondary metabolic pathways in plants (Allan et al., 2008; Wang et al., 2017). The anthocyanin biosynthesis pathway in plants is well understood and key enzyme genes involved in anthocyanin biosynthesis, including cinnamate 4-hydroxylase (*C4H*), 4-coumarate-CoA ligase (*4CL*), chalcone isomerase (*CHI*), chalcone synthase (*CHS*), flavanone 3-hydroxylase (*F3H*), flavonoid 3′-monooxygenase (*F3'H*), flavonoid 3′, 5′-hydroxylase (*F3'5H*), dihydroflavonol-4-reductase (*DFR*), phenylalanine ammonialyase (*PAL*), anthocyanidin synthase (*ANS*), and UDP-glucose flavonoid glucosyltransferase (*UFGT*), have been identified in the colored tissues of several plants. Moreover, the pathway is regulated by the interaction of DNA-binding R2R3MYB transcription factors and MYC-like basic helix–loop–helix (bHLH) and WD40-repeat proteins, i.e., MYB-bHLH-WD40 transcription factor complexes (Jaakola, 2013). The up-regulation of these genes could promote anthocyanin accumulation (Liu et al., 2016; Gao et al., 2021). Aside from these

genes and transcription factors, anthocyanin biosynthesis is also influenced by environmental factors (Liu et al., 2016; Lin et al., 2018). A considerable amount of new research has focused on elucidating the environmental regulations controlling anthocyanin biosynthesis, specifically the impact of light-mediated regulation (Jaakola, 2013). Light exposure can increase anthocyanin concentrations, especially in the fruit skin, and the shading of fruits can have the opposite effect (Takos et al., 2006). In the pathway of light-controlled anthocyanin biosynthesis, light-receptor could interact with the CONSTITUTIVE PHOTOMORPHOGENIC1 (COP1), which regulated the expression level of the ELONGATED HYPOCOTYL5 (HY5), or interact directly with certain anthocyanin biosynthesis-related MYB transcription factors to promote the transcription of structural pathway genes regulating anthocyanin biosynthesis (Stracke et al., 2010; Li et al., 2012). In recent decades, researchers have found that light quality can significantly affect the biosynthesis of anthocyanins; UV and other types of light (e.g., blue light) have been associated with the regulation of anthocyanin biosynthesis (Ordidge et al., 2011; Li et al., 2013). Additionally, light quality was found to regulate the gene expression patterns related to anthocyanin synthesis and the regulation of anthocyanin accumulation and stability (Briggs and Olney, 2001; Zoratti et al., 2014). During anthocyanin biosynthesis, the regulatory effects of anthocyanin pigmentation were found to vary according to the different light quality, which influences the levels of key enzymes that are involved in anthocyanin biosynthesis (Mol et al., 1996). For most plants, blue light induces an increase in the transcript levels for the *PAL1*, *CHS*, *CHI*, and *DFR* genes, which encode anthocyanin biosynthesis enzymes (Feinbaum et al., 1991; Batschauer et al., 1996; Noh and Spalding, 1998). Red light (600–700nm) may markedly increase peroxidase (POD) and phenylalanine ammonialyase (PAL) activity in a phytochrome-type response, irritating the aggrandizement of anthocyanin content in the flavonoid metabolic pathways of rice, maize, and turnips (Reddy et al., 1994; Sharma et al., 1999). Some studies have elucidated the effects of different light qualities on anthocyanin accumulation in certain plants and the transcriptome profiling analysis of blueberry fruit has revealed the mechanisms of red and blue light-mediated anthocyanin biosynthesis regulations (Li and Yang, 2007; Lu et al., 2015; Ottosen et al., 2015; Samkumar et al., 2021). However, none of these studies have demonstrated the regulatory mechanism underlying anthocyanin biosynthesis is response to different spectra by integrating the transcriptome and metabolome analysis. This approach may allow us to gain a better understanding of the key genes and metabolites involved in anthocyanin formation under different light qualities in blueberry leaves.

RNA-sequencing is a powerful tool that can be used to unravel novel genes, identify gene expression levels, and facilitate the study of mechanisms underlying metabolite variations (Lv et al., 2019; Zhang Y. et al., 2019). However, it is still difficult to determine a direct correlation between transcript abundance and the associated levels of the respective metabolites since numerous variables are often considered (Lou et al., 2014). Metabolomics is an important part of systematic biology that focuses on the quantitative analysis of all metabolites in an organism to understand the correlations between phenotype and the metabolite (Zhang X. et al., 2019; Guo et al., 2020). An integrative analysis of the transcriptome and metabolome, to investigate the transcript levels in conjunction with the metabolic products, could contribute to the identification of functional genes and the elucidation of pathways involved in plant metabolism processes (Lou et al., 2014).

In this study, an integrated transcriptome and metabolome analysis was used to understand the key genes and metabolites associated with anthocyanin biosynthesis in blueberry leaves under different light qualities. The differential levels of anthocyanin metabolites and their regulatory genes in blueberry leaves were identified under four light qualities, i.e., blue (B), red (R), red/blue (60R/40B), and white light. The connection network was mapped based on a correlation analysis between transcript expression and metabolite levels to highlight the regulator genes and metabolites related to anthocyanin accumulation under different light qualities. Moreover, a correlation analysis between anthocyanin accumulation and photosynthesis was also conducted. The data presented herein will not only provide novel insights into the molecular mechanisms underlying the biosynthesis and regulation of anthocyanins in blueberry, but also useful insights to aid in the breeding of new blueberry varieties with enhanced anthocyanin content.

Materials and methods

Plant materials

In this study, three-year-old southern-highbush blueberries of the cultivar 'Misty' were utilized, and the plants were grown in a plant factory system with artificial lighting (PFAL) at Zhuanghang Comprehensive Experimental Station of Shanghai Academy of Agricultural Sciences, Shanghai, China. The growth conditions in the PFAL were as follows: the temperature was set at 25 ± 2°C during the daytime and 22 ± 2°C at night, the photoperiod was 16-h light and 8-h dark, and the CO_2 concentration and relative humidity were set at approximately 400 μmol·mol^{-1} and 65% ± 5%, respectively. The blueberry plants were treated respectively with four light qualities: white (W), full blue (B), full red (R), and 60% red + 40% blue (60R/40B), and the white light treatment was considered the control. The light intensity of all treatments was equal at 200 ± 5 μmol·m^{-2}·s^{-1}, each treatment with four experimented blueberry plants. Approximately 30 days after the

treatments, the fifth to tenth mature leaves, which were similar developmental stage, had no disease, no mechanical damage, and complete form, were collected from the top branch of each blueberry plant. Approximately 3 g of leaves were collected from each plant and immediately frozen in liquid nitrogen and stored at −80°C for transcriptome profiling and metabolite quantitative analysis. The determination of physiological indicators was performed with four replicates and each plant as a replicate. The transcriptome and metabolome were analyzed with three biological replicates.

Determination of anthocyanin content

Anthocyanin content in the fresh leaf samples was determined according to the method improved by Li et al. (2016). Leaves were harvested from the blueberry plants, quickly weighed, and homogenized with a mortar and pestle; they were then treated with 0.1 mol·L^{-1} acidized ethanol (10mL) in a microcentrifuge tube and then with water at 60°C for 30 min. The extraction procedure was performed twice and then the samples were centrifuged at 1,500g for 5 min at 4°C, and the supernatants were then collected. The absorbance of the samples was measured at 530, 620, and 650 nm, respectively, using an ultraviolet–visible spectrophotometer (UV-2000) (Spark, TECAN, Switzerland). The values are reported as (1) ΔOD = (OD530 − OD620) − 0.1 (OD650 − OD620) and (2) anthocyanin content = (ΔOD × V × 1000000)/(ξ× m), where V is the dilution volume, ξ is the molar absorption coefficient, and m is the fresh weight of the sample. All measurements were performed with four replicates.

Leaf photosynthesis measurements

The photosynthesis of the blueberry leaves was measured using a leaf gas exchange system (CIRAS-3; PP Systems). During the measurements, the gas exchange system controlled the environmental conditions within the curette. The temperature of the leaf was 25.0 ± 0.2°C and the CO_2 concentration was 390 ± 5 μmol·mol^{-1}. To discover the photosynthetic efficiency of the different spectra, light response curves were constructed for the blueberry leaves under corresponding light qualities. Photosynthesis indicators, including the maximum gross assimilation rate ($A_{g,\ max}$) and maximum quantum yield for CO_2 assimilation ($QY_{m,\ inc}$), were calculated according to the light response curves.

Total RNA extraction, RNA-seq, and transcript profile analysis

Total RNA from the blueberry leaves was isolated using the Polysaccharides and Polyphenolics-rich RNA Prep Pure Plant Kit (TIANGEN, China), following the manufacturer's instructions. The integrity of the RNA was verified using an Agilent 2100 Bio analyzer (Agilent Technologies, USA). The RNA-Seq and libraries were constructed using the Biomarker Technologies Corporation (Beijing, China) and the Vcorymbosum_v1.0 genome was used as the reference genome in this study. Before conducting data analysis, it is vital to ensure that the red reads are of sufficient quality. Remove reads containing joints and low-quality reads (N was more than 10%; base number with mass value Q≤ 10 accounted for more than 50%), clean reads with high quality were obtained. StringTie was used to assemble the clean reads above and construct a traffic network using the alignment information to construct a multivariable shear map through using the maximum flow algorithm to assemble reads and evaluate its expression. Using the FPKM (Fragments Per Kilobase of transcript per Million fragments mapped) as an indicator to measure the expression level of transcripts or genes through the maximum flow algorithm. The edgeR was used to analysis significant difference of genes transcript level and gene with an adjusted FDR < 0.01& Fold Change ≥2 was identified as differentially expressed. The functions of the unigenes were determined by aligning them to protein databases using BLASTx, including NCBI non-redundant (Nr) and nucleotide (Nt), Swiss-Prot, Kyoto Encyclopedia of Genes and Genomes (KEGG), Clusters of Orthologous Groups of proteins (COGs), and Gene Ontology (GO) databases. GO classification was performed by mapping the relationship between the Swiss-Prot and GO terms, and the genes were mapped to the KEGG database to annotate their potential metabolic pathways (Kanehisa et al., 2008).

Expression validation of the differentially expressed genes with real-time quantitative RT-PCR

Total RNA was isolated from the blueberry leaves collected under the four light-quality treatments. For the qRT-PCR analysis, the SYBR RNA RT-PCR Kit (TaKaRa, Japan) was used to synthesis the first-strand of cDNA and the Light Cycler 480 (Roche, USA) was used to perform the qRT-PCR procedure. qRT-PCR was performed with the SYBR-Green PCR kit (TaKaRa, Japan) and the reaction mixture system was prepared according to the manufacturer's instructions. The qPCR program was previously described by Zhang et al. (2021b). The expression level of each gene was calculated in triplicate based on the $2^{-\Delta\Delta Ct}$ algorithm, and *VcGAPDH* was used as the internal reference gene. Snap Gene software (www.snapgene.cn) was used to design the primers for the target genes (Table S1).

Metabolite profiling using ultraperformance liquid chromatography tandem mass spectrometry

Anthocyanins are secondary metabolites in plants. The qualitative and quantitative analysis of the anthocyanins could be measured using the LC-MS/MS system. Freeze-dried blueberry leaves (50 mg) were crushed and dissolved in 500 μL of extract. They were then centrifuged at 1200 g/min for 10min; the extracts were filtrated and absorbed for LC-MS/MS analysis. The instrument system used for data acquisition included an Ultra Performance Liquid Chromatography (UPLC) and Tandem Mass Spectrometry (MS/MS). Qualitative analysis of the anthocyanins detected using mass spectrometry was performed based on the construction of standard databases and the quantification of anthocyanins was identified by employing scheduled multiple reaction monitoring (MRM). The Analyst 1.6.3 software (Sciex) was used to acquire and dispose data and the quantities of all metabolites were analyzed using Multiquanta 3.0.3 software (Sciex). To draw standard curves for the different substances the concentration was used as the abscissa and the peak area as the ordinate. The integrated peak areas for all detected samples were substituted into the standard curve linear equation for calculation and the absolute content values for each sample were obtained.

Targeted selection of anthocyanins and their intermediates

Targeted anthocyanins and their intermediates were selected based on the following criteria: (1) the obtained compounds should be related to the phenylpropanoid and anthocyanin biosynthesis pathways and their molecular formula and mass information was established using the KEGG and plant metabolic network databases; (2) certain filtrated daughter ions in the anthocyanins could be identified as similar compounds as described in the literary references; and (3) the MS^2 spectra of the selected metabolites and standard compounds was as found in the METLIN and LIPDMAPS databases. The obtained anthocyanin-related compounds are shown in Supplementary Table S2.

Integrated transcriptome and metabolome analysis

Pearson's correlation coefficient (PCC), a statistical method used to measure the linear relationship between two variables with values between −1 and 1, was used to calculate the correlation coefficient between the transcriptome and metabolome data and further integration to analyze the relationships between the gene transcript and metabolite content. Associations between PCC > 0.8and $P < 0.05$ were selected and the network of genes and metabolites was constructed using Cytoscape software (Cytoscape Consortium, San Diego, CA). Additionally, genes and metabolites related to anthocyanin biosynthesis were mapped using the KEGG pathway database.

Data analysis

For this study, all experiments were performed in triplicate and the results are expressed as the mean ± standard deviation (SD). Analysis of variance (ANOVA) was performed to estimate the significant differences between the treatment means at $P < 0.05$. *A* heat map was constructed to exhibit the gene expression pattern and metabolite accumulation and to separate the groups of genes and metabolites with similar expressions and accumulation values. Principal component analysis (PCA) was performed to realize the metabolite variability among the white, blue, red, and red/blue light treatment groups.

Results

Phenotypic characteristics, anthocyanin content, and photosynthesis of blueberry leaves under different light qualities

The relative spectra under different wavelengths were significantly different (Figure 1I) and the different light qualities influenced the color gradation and anthocyanin synthesis in the blueberry leaves. The color gradation in the blueberry leaves was deepest with the blue light, followed by the red/blue, red, and white lights, and the corresponding anthocyanin content was highest under the blue light, followed by red/blue, red, and white lights. The anthocyanin content of the leaves under blue, red/blue, and red lights was approximately 4.6-, 3.7-, and 1.6-fold higher, respectively, than that under the white light (the control group) (Figure 1II). A phenomenon was identified in which the blue and red lights could promote the accumulation of anthocyanin, but differed significantly in their impact levels. The promotion of anthocyanin biosynthesis was more prominent under blue light than red light, while the accumulation was in between these two levels with the red/blue light combination. Leaf photosynthesis exhibited adverse effects on anthocyanin accumulation, which were strongest under white light but lowest under blue light. The photosynthesis indices, including $A_{g,\ max}$ and $QY_{m,\ inc}$, presented a consistent trend, with the greatest index under the white light, followed by red, red/blue, and blue (Figure 1III). Blue and red lights can promote anthocyanin accumulation but impair photosynthesis in blueberry leaves. In contrast, white light prejudiced anthocyanin biosynthesis but facilitated photosynthesis.

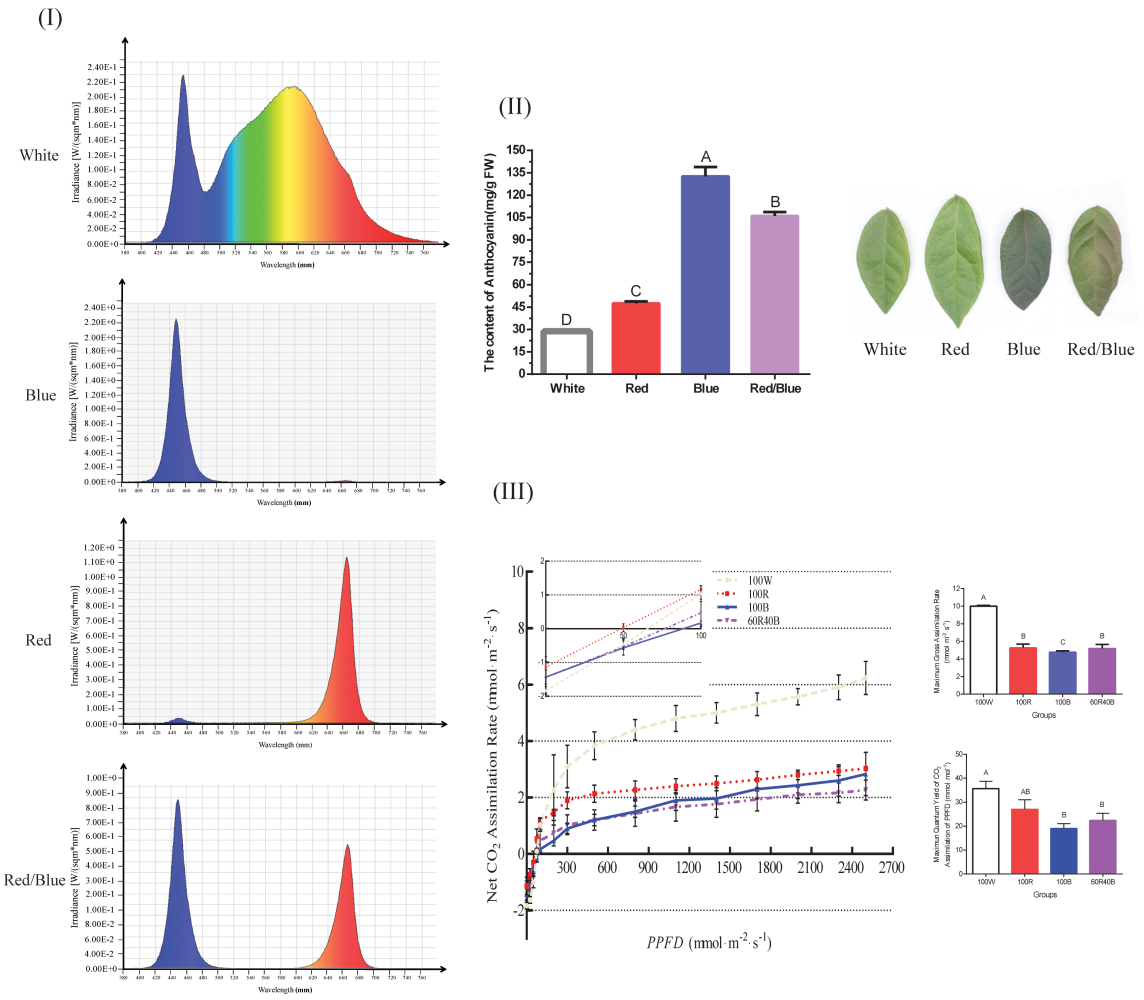


FIGURE 1
Spectrum plot showing the different light qualities and anthocyanin content, light response curves, and photosynthesis of blueberry leaves under four light spectra. **(I)** Dynamic distributions of the relative spectra under different wavelengths; **(II)** Anthocyanin content and phenotypic comparisons of blueberry leaves under four light spectra; **(III)** Net assimilation (*An*)-light response curves, $A_{g,\ max}$, and $QY_{m,\ inc}$ for blueberry leaves under different light spectra. $A_{g,\ max}$, maximum gross assimilation rate; $QY_{m,\ inc}$, maximum quantum yield of CO_2 assimilation; Capital letters indicate significant differences at $P < 0.01$ among the different treatments. Error bars indicate the standard deviation based on three replicates.

Transcriptome data and screenings for the DEGs under different light qualities

To elucidate the molecular genetics mechanisms underlying anthocyanin biosynthesis in blueberry leaves when under different light conditions, a comparative transcriptome analysis was performed. The Illumina sequencing platform was used to construct and sequence twelve RNA-Seq libraries (each treatment with three biological replicates). The rigorous quality estimation and data cleaning resulted in 21.48–26.23 M clean reads with Q30 bases that were identified as high-quality reads for further analysis. Their GC contents and Q30 ranges were 46.93%–47.36% and 94.69%–94.93%, respectively (Table S3). Moreover, the N percentage was 0.00% in all samples.

To substantiate the effects of light quality on the transcript levels of anthocyanin biosynthesis genes in blueberry leaves, DEGs between samples from the light quality treatments were identified using pairwise comparisons of the four cDNA libraries (blue vs. white, red vs. white, red/blue vs. white, and blue vs. red), and the fragments per kilo-base per million mapped reads (FPKMs) of all DEGs were calculated. A total of 1812 (640 up-regulated and 1172 down-regulated), 596 (250 up-regulated and 346 down-regulated), 3283 (1123 up-regulated and 2160 down-regulated), and 3428 (1974 up-regulated and 1457 down-regulated) differentially expressed genes ($P < 0.05$, fold change > 2) were found during the comparisons of the blue and white, red and white, red/blue and white, and blue and red libraries,

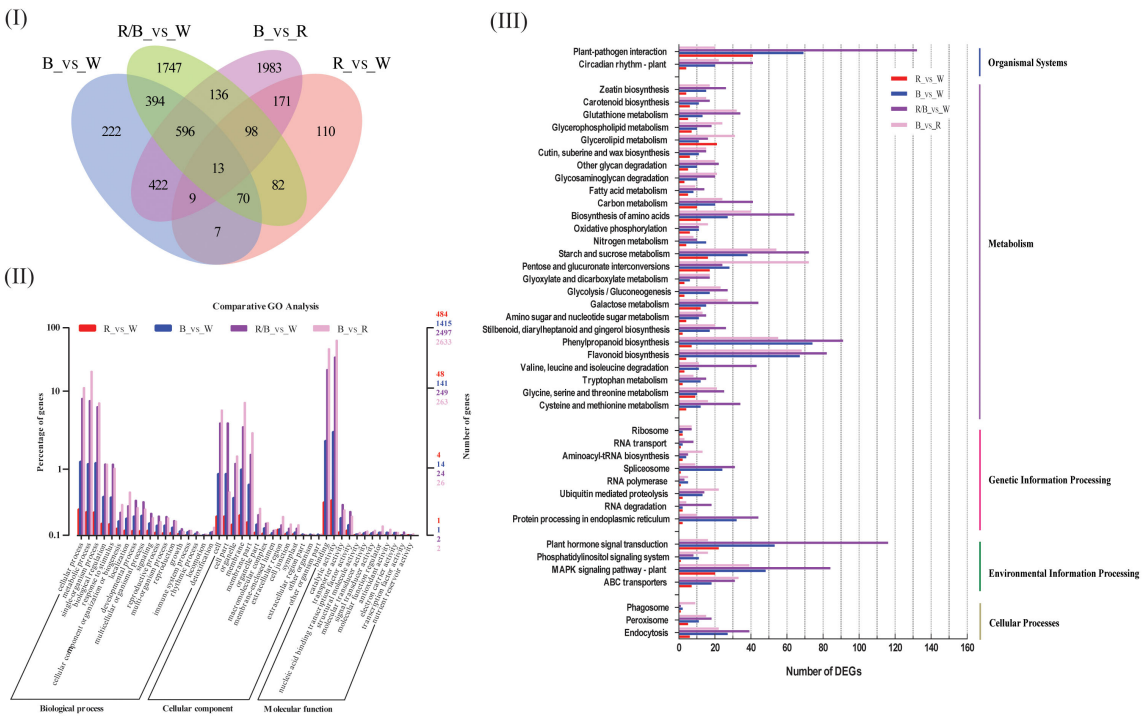


FIGURE 2
Functional annotation results of the genes differentially expressed in blue **(B)**, red **(R)**, and red/blue (60R/40B) when compared with those of the corresponding control (white light) and the comparison of B vs. R using the transcriptome data from blueberry 'Misty' leaves; **(I)** Venn diagram showing the shared and common DEGs for all light quality treatments; **(II)** Functional annotation of the DEGs based on the Clusters of Orthologous Groups (COG) database; **(III)** Annotation information obtained from the Gene Ontology (GO) database; **(IV)** Functional classification of DEGs based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) dataset.

respectively, whereas 13 DEGs were shared among the four comparisons (Figure 2I; Table S4). The number of DEGs was substantially higher in the blue vs. white comparison when compared with red vs. white, which indicated a strong gene expression disturbance experienced by the blueberry leaves under blue light. The blueberry leaves were more sensitive and actively responded to blue light to promote anthocyanin accumulation when compared to red light. GO term annotation was performed to improve our understanding of the functions of these DEGs in blueberry leaves under different light conditions. The DEGs were considered enriched in GO-terms if at least one of the terms was categorized as a biological process, cellular component, or molecular function (Figure 2II). The results of the GO-biological process (GO-BP) analysis revealed that significantly enriched DEGs were associated with metabolic processes, mainly phenylalanine and secondary metabolisms, among the four comparisons. To further explore the enriched metabolic pathways related to anthocyanin accumulation in blueberry leaves, we mapped all DEGs to reference canonical pathways in the KEGG database. A total of 4097 DEGs (blue vs. white:844; red vs. white:307; red/blue vs. white:1440; and blue vs. red:1506) with annotated KEGG results were classified into five main categories, including organismal systems, metabolism, genetic information processing, environmental information processing, and cellular processes, based on the type of KEGG pathways (Figure 2III). During metabolism classification, phenylpropanoid biosynthesis and starch and sucrose metabolism (photosynthesis) were the primary enriched pathways throughout the four comparisons. Additionally, DEGs were annotated by searching against the COG dataset. COG-annotated putative proteins were functionally classified into at least 26 categories, including cellular structure, biochemical metabolism, molecular processing, and signal transduction (Figure S1). In addition, the DEGs were collected and classified into 17 groups based on their annotated functions (Table 1). Among all the categories, phenylalanine metabolism, secondary metabolism, starch and sucrose metabolism, and carbohydrate metabolism (photosynthesis) were highly significant and enriched pathways, as well as being enriched they also had a greater number of DEGs. The results indicated that these DEGs were associated with metabolic pathways that could contribute to the biosynthesis and accumulation of anthocyanin in blueberry leaves under different light qualities.

TABLE 1 Functional categories of the genes differentially expressed in the blue, red, and red/blue light when compared with white light and the comparison of blue vs. red in blueberry leaf.

Annotation	Blue_vs_White				Red_vs_White				Red/Blue_vs_White				Blue_vs_Red			
	Up		Down		Up		Down		Up		Down		Up		Down	
	Number	%	Number	%	Number	%	Number	%	Number	%	Number	%	Number	%	Number	%
Photosynthesis	5	0.74	1	0.08	2	0.73	1	0.29	5	0.40	2	0.09	3	0.22	8	0.39
Cell part	160	23.53	290	23.09	68	24.82	72	20.69	307	24.74	516	22.64	340	25.70	577	27.79
Secondary metabolism	23	3.38	103	8.20	22	8.03	21	6.03	37	2.98	167	7.33	67	5.06	117	5.64
Plant hormone	17	2.50	36	2.87	12	4.38	10	2.87	45	3.63	71	3.12	7	0.53	9	0.43
Signaling	20	2.94	39	3.11	3	1.09	18	5.17	28	2.26	65	2.85	26	1.97	13	0.63
Cell	160	23.53	290	23.09	68	24.82	72	20.69	307	24.74	516	22.64	340	25.70	577	27.80
Development	40	5.88	85	6.77	10	3.65	21	6.03	80	6.45	148	6.49	83	6.27	121	5.83
Growth	3	0.44	16	1.27	2	0.73	8	2.30	12	0.97	23	1.01	7	0.53	23	1.11
Carbohydrate metabolism	43	6.32	58	4.62	31	11.31	11	3.16	89	7.17	118	5.18	65	4.91	134	6.45
Transport activity	43	6.32	85	6.77	5	1.82	19	5.46	52	4.19	174	7.63	71	5.37	116	5.59
Starch and sucrose metabolism	19	2.79	19	1.51	4	1.46	12	3.45	31	2.50	41	1.80	20	1.51	34	1.64
Phenylalanine metabolism	5	0.74	12	0.96	1	0.36	0	0.00	2	0.16	16	0.70	7	0.53	4	0.19
Biological regulation	105	15.44	199	15.84	32	11.68	70	20.11	189	15.23	380	16.67	232	17.54	289	13.92
Molecular function regulator	21	3.09	17	1.35	7	2.55	11	3.16	26	2.10	25	1.10	37	2.80	32	1.54
Function unknown	8	1.18	4	0.32	5	1.82	0	0.00	15	1.21	4	0.18	10	0.76	9	0.43
RNA regulation	8	1.18	2	0.16	2	0.73	2	0.57	16	1.29	13	0.57	8	0.60	13	0.63
Total	680	100.0	1256	100.0	274	100.0	348	100.0	1241	100.0	2279	100.0	1323	100.0	2076	100.0

10.3389/fpls.2022.1073332

Transcriptional profiles of the DEGs related to photoreceptors, photosynthesis, and anthocyanin metabolism in blueberry leaves under different light qualities

The transcriptome analysis under light conditions showed that the DEGs related to phenylalanine, secondary, starch and sucrose, and carbohydrate (photosynthesis) metabolism processes may help blueberry leaves to accumulate anthocyanins. The physiological indicators suggested that there were significant differences in anthocyanin accumulation and photosynthesis in blueberry leaves due to the distinct light qualities. The DEGs implicated in the crucial steps of these processes were consequently screened. Moreover, the blueberry leaf was treated with distinct spectra, the DEGs associated with photoreceptor responses to the different light qualities were searched for. Core genes related to anthocyanin biosynthesis, photosynthesis, and photoreceptors were studied in detail, and the results demonstrated that most of the uni-transcripts showed significant changes in their expression levels (Figure 3; Table S5). An opposite dynamic trend was observed between the expression patterns of the anthocyanin accumulation-related and photosynthesis-related DEGs. The expression levels for most of the anthocyanin biosynthesis-related DEGs were higher in the blue light when compared with the other lights, but the DEGs associated with photosynthesis were down-regulated with the blue light. There may be feedback regulation among the DEGs related to anthocyanin accumulation and photosynthesis under blue light. DEGs encoding phytochrome, as receptor proteins of red and far-red light (600–750 nm), were up-regulated under red or red/blue light. Consistently, DEGs encoding cryptochrome, which mediated several responses to blue light and UV-A (320–500 nm), were up-regulated under blue light (Figure 3I). In relation to photosynthesis, there were DEGs among the different spectra. Photosynthesis-related DEGs were down-regulated under blue light when compared with the other treatments (Figure 3II), which was

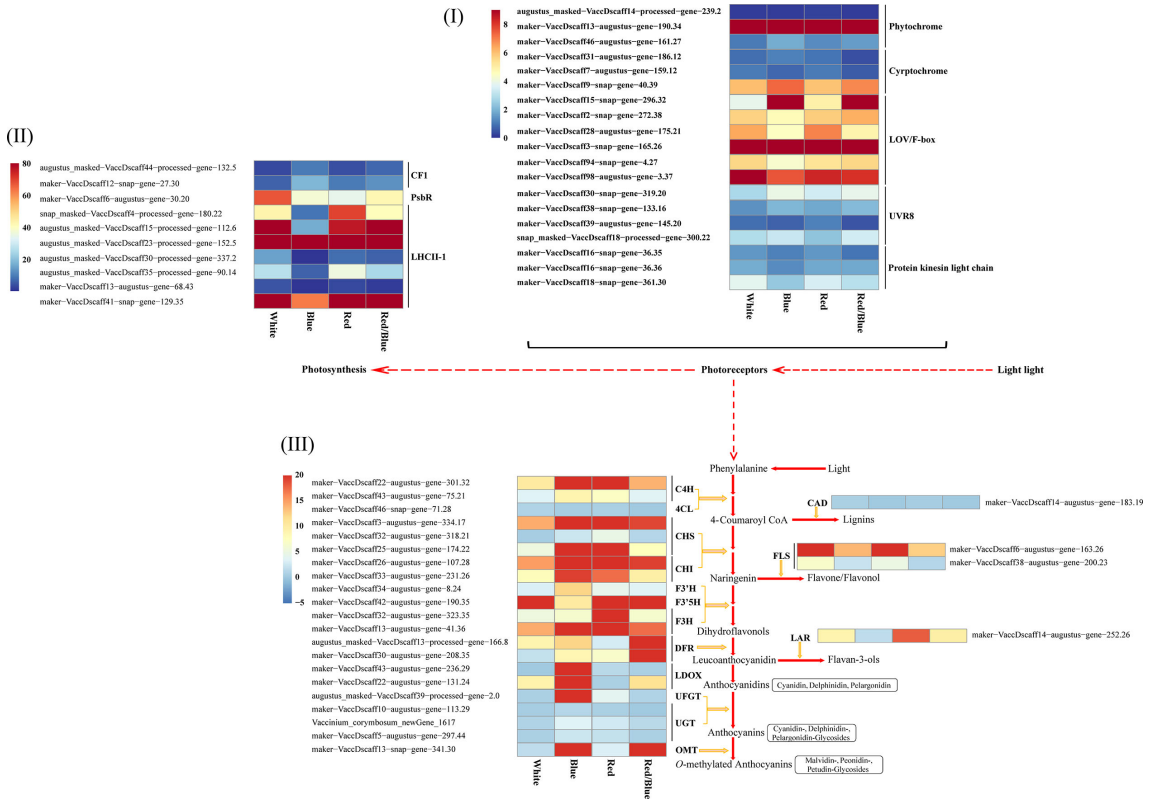


FIGURE 3
Expression heatmap showing the differentially expressed genes related to photoreceptor, photosynthesis, and anthocyanin and flavonoid biosynthesis pathway under different light qualities for the blueberry leaf. **(I)** Differential expression of the genes related to the photoreceptor, including phytochrome, cryptochrome, LOV/F-box, UVR8, and protein kinases in the light chain; **(II)** Expression levels of the DEGs from CF1, PsBR, and LHCII-1; **(III)** Expression patterns of the DEGs involved in the anthocyanin and flavonoid biosynthesis pathway. Enzyme names and expression levels are shown at the side of each step. C4H, cinnamate 4-hydroxylase; 4CL, 4-coumarate:CoA ligase; CHS, chalcone synthase; CHI, chalcone isomerase; OMT, *O*-methyltransferase; F3′H, flavonoid3′-monooxygenase; F3′5H, flavonoid 3′, 5′-hydroxylase; DFR, dihydroflavonol-4-reductase; LDOX, leucoanthocyanidin dioxygenase; LAR, leucoanthocyanidin reductase; UFGT, UDP-glucose flavonoid glucosyltransferase; UGT, anthocyanidin 3-O-glucoside 5-O-glucosyltransferase; CAD, cinnamyl alcohol dehydrogenase; FLS, flavonolsynthase; Red and blue, respectively, indicated high and low expression according to the color bar using the unit variance scaling method.

consistent with the weakness of photosynthesis under blue light. These results suggest that blue light prejudiced photosynthesis in the blueberry leaves, which is in sharp contrast to the results for anthocyanin accumulation under blue light. In many cases, changes in anthocyanin biosynthesis correspond to changes in the expression of genes that encode pathway enzymes. Among the related anthocyanin accumulation processes, 22 DEGs (including *C4H*, *4CL*, *CHS*, *DFR*, *F3'H*, *F3'5H*, *F3H*, *LDOX*, *OMT*, and *UFGT*) were found to participate in vital steps in the anthocyanin pathway. Among these, *LDOX*, *UFGT*, and *OMT* were expressed well under blue light, whereas *DFR* and *OMT* were highly expressed under red/blue light. These DEGs contributed to the cyanidin, pelargonidin, malvidin, delphinidin, petunidin, and peonidin accumulation in the anthocyanin biosynthesis pathway (Figure 3III). Additionally, the expression of the DEGs related to lignin and flavonoid metabolism, which are also significant secondary metabolite pathways in plants, decreased in the blue light in comparison with that in white, red, and red/blue lights. These results indicate that blue light could promote the accumulation of anthocyanin but hinder the photosynthesis and secondary metabolism process of lignin and flavonoids in blueberry leaves when compared with other light sources *via* the direct or indirect regulation of the expression levels of related DEGs.

Differentially expressed transcription factors and the co-expression network between DETFs and DEGs involved in anthocyanin and flavonoid accumulation

Blueberry leaves under different light qualities showed significantly different TFs, belonging primarily to 20 different families (Figure 4I; Table S6). The differentially expressed TFs belonged mostly to the AP2/ERF-ERF, bZIP, bHLH, MYB, and NAC families. The number distribution mole for the TFs in the blueberry leaves under different light qualities was consistent with the quantitative distribution of the gene function annotation. There were more differentially expressed genes or TFs under the red/blue light, indicating that the red/blue light combination was beneficial for changes in gene or TF expression levels.

TFs regulate gene expression patterns by acting as activators or repressors to induce or inhibit gene promoter activity. There is a potential regulatory mechanism by which differentially expressed TFs (DETFs) may influence anthocyanin biosynthesis by adjusting the expression levels of the DEGs implicated in anthocyanin accumulation under different light qualities. To verify this hypothesis, co-expression networks were established between the TFs and genes using gene expression profile data based on the correlation coefficient. The resulting co-expression networks indicated that different TFs connected distinct genes, and there were correlations among them (Figure 4II). TFs (*MYB*, *NAC*, and *bHLH*) connected more genes related to anthocyanin accumulation, and these TFs were highly expressed in blueberry leaves under blue light.

Validation of key DEGs and DETFs involved in anthocyanin accumulation by qRT-PCR

To further validate the comparative transcriptome results, the transcript level variances of some putative genes or TFs related to anthocyanin accumulation were evaluated by qRT-PCR analysis

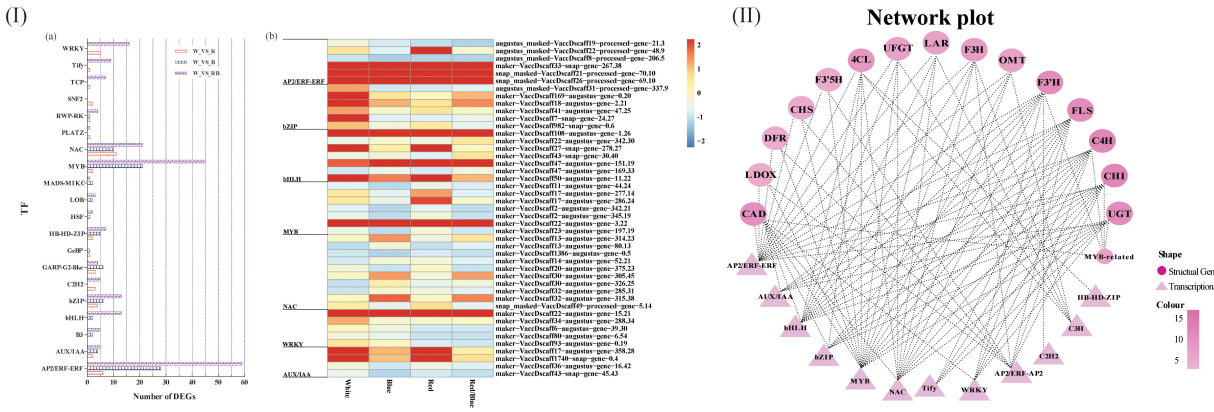


FIGURE 4
Distribution of the differentially expressed transcription factors (TFs) and co-expression networks between candidate genes and transcription factors. **(I)** Quantitative statistics for the differentially expressed TFs **(A)** and a heatmap showing the candidate TFs under various light qualities **(B)**; **(II)** Co-expression networks between differentially expressed structural genes involved in the anthocyanin and flavonoid biosynthetic pathway and differentially expressed TFs under different light spectra.

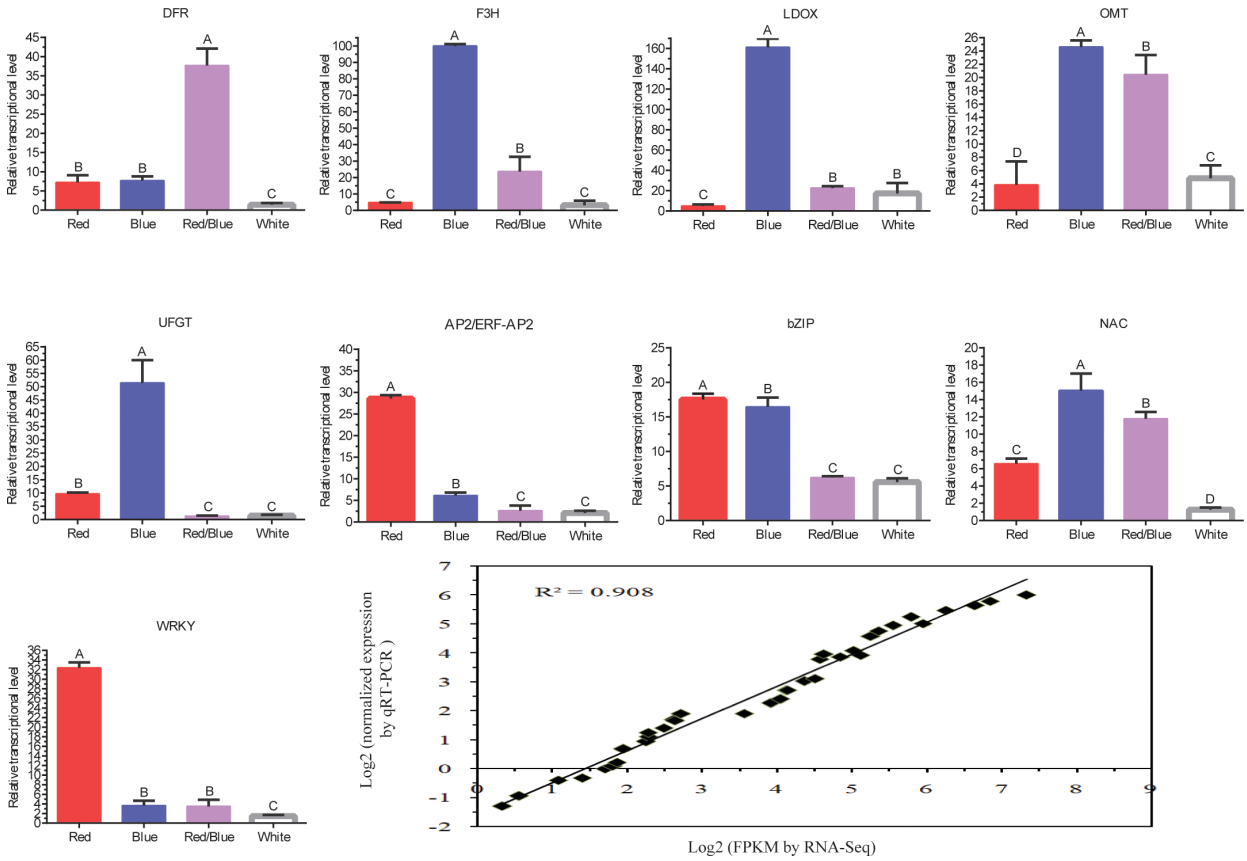


FIGURE 5
Quantitative RT-PCR validation of the structural genes involved in the anthocyanin biosynthesis pathway and transcription factors and the correlation analysis of the gene expression results obtained from the qRT-PCR and RNA-Seq under different light qualities (red, blue, red/blue, and white). The x-axis represents the different light qualities while the y-axis represents the relative expression of each gene. Capital letters indicate significant differences at $P < 0.01$ among the different treatments. Error bars represent the standard deviation based on three replicates.

with three biological replicates (Figure 5). Structural genes, including *F3H*, *LDOX*, *OMT*, and *UFGT*, involved in the anthocyanin biosynthesis pathway exhibited similar dynamic trends, with the highest transcript levels under blue light, followed by red/blue, red, and white. The expression level of *DFR* was the highest under red/blue light, and the effect of inducing *DFR* transcripts was the least significant under white light. The transcript level for *UFGT* was the lowest under red/blue light, and its peak occurred under blue light. Moreover, the transcript level for the *OMT* was second highest under the red/blue light. These expression trends agreed with the alterations in gene expression detected by the transcriptome analysis, corroborating the RNA-sequencing data. These results coincided with the dynamic changes in anthocyanin content under different light qualities (they were highest in blue light, followed by red/blue, red, and white light), indicating that these structural genes may play a role in anthocyanin biosynthesis in blueberry leaves in response to the different spectra. The expression levels of the TFs (*AP2/ERF-AP2*, *bZIP*, *NAC*, and *WRKY*) were higher in the blue, red, and red/blue lights than in the corresponding control (white light). Moreover, a correlation analysis of the gene expression results was obtained using qRT-PCR and RNA-Seq data under different light qualities and a close connection was identified between them (correlation coefficient: $R^2 = 0.908$). The high consistency between the RNA-Seq and qRT-PCR results suggested that the RNA-Seq data were reliable for evaluating the regulation of gene expression under different light quality treatments in blueberry leaves.

Metabolome profiling analysis of blueberry leaves under different light qualities

To compare the metabolite compositions associated with anthocyanin accumulations in the blueberry leaves under different light conditions, the major anthocyanins were determined using LC-MS/MS. In total, 108 anthocyanidins were detected and grouped into eight categories (Figure 6I). The largest category

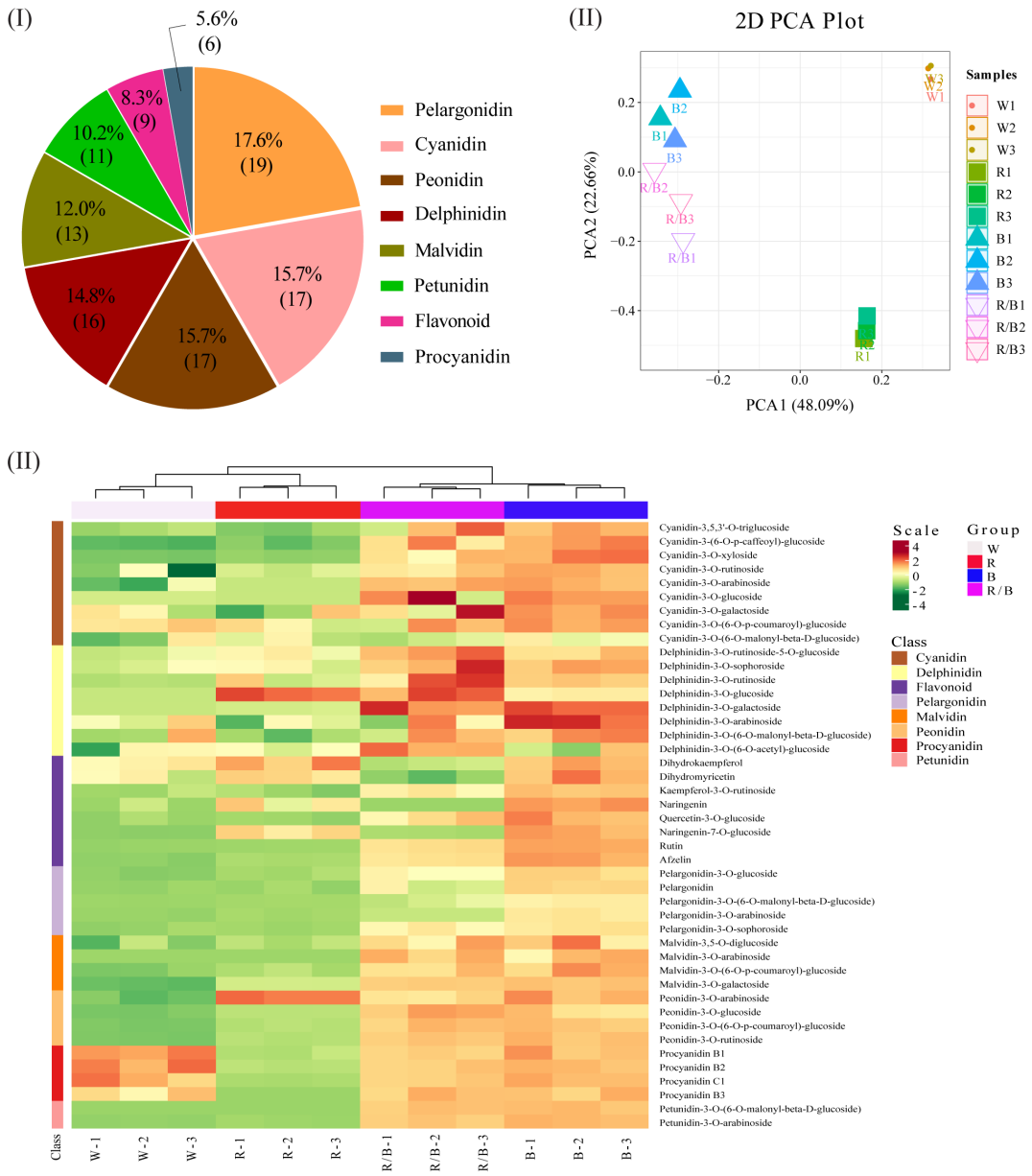


FIGURE 6
Metabolome profiling of 'Misty' blueberry leaves under different light qualities: white (W), red (R), blue (B), and red/blue (R/B). **(I)** Primary classification of the metabolites identified in the leaves. **(II)** Principal component analysis (PCA) scores for the leaves under different light qualities for the first component (PC1) and second component (PC2). **(III)** Hierarchical clustering and correlation analysis of the leaf metabolites under different light qualities.

(pelargonidin) contained 19 metabolites, followed by cyanidin (n=17), peonidin (n=17), delphinidin (n=16), and malvidin (n=13). The PCA for the anthocyanin derivatives was used to classify twelve samples into four distinct clusters, accurately reflecting the four light quality treatments for the blueberry leaves (Figure 6II). In this model, the secondary principal component (PC2; 22.66% of the total variables) was clearly separated between blue and red/blue light. The differences between the white and red light resulted from PC1 and PC2 in this model (48.09% and 22.66% of the variables, respectively). An Orthogonal Partial Least Squares-Discriminant Analysis (OPLS-DA) model was performed to identify the differentially accumulated metabolites responsible for metabolic differentiation among the various light quality treatments for the blueberry leaves. A total of 43 (25 up-accumulated; 3 down-accumulated), 41 (7 up-accumulated; 4 down-accumulated), 44 (24 up-accumulated; 3 down-accumulated), and 42 (22 up-

accumulated; 1 down-accumulated) metabolites were selected for comparison analysis as follows: blue vs. white, red vs. white, red/blue vs. white, and blue vs. red lights, respectively (Table S7). The hierarchical clustering and correlation analysis exhibited a compact dependence of the anthocyanin metabolites on the different light qualities of the blueberry leaves (Figure 6III). The overall accumulation of the anthocyanin metabolites was higher under blue light than under the other light conditions, and the accumulation of cyanidin, pelargonidin, malvidin, peonidin, and petunidin was significantly higher under the blue and red/blue lights when compared with the white and red lights.

Differential accumulation of derivatives related to anthocyanin biosynthesis under different light qualities

Based on the KEGG database annotation, 44 metabolites were found to be involved in the anthocyanin biosynthesis pathway and the identified anthocyanin derivatives and their relevant compounds were rearranged based on their corresponding positions in the anthocyanin biosynthesis pathway, which were established based on the KEGG, PMN, and literature references. Among them, two naringenins (Ng), one dihydrokaempferol (Dk), one kaempferol (Ka), three flavones, nine cyanidins (Cy), four peonidins (Pn), five pelargonidins (Pg), one dihydromyricetin, eight delphinidins (Dp), two petunidins (Pt), four malvidins (Mv), and four proanthocyanidins (Pa) were identified (Figure 7). Each metabolite in the anthocyanin biosynthesis pathway differed significantly under the different light conditions in the blueberry leaves. Interestingly, the content for most metabolites was highest under the blue light but lowest under the white light, suggesting that the blue light was beneficial in promoting anthocyanin accumulation in blueberry leaves. The dynamic changes in the content of metabolites in the anthocyanin biosynthesis pathway were consistent with the variations in the expression levels for the genes related to

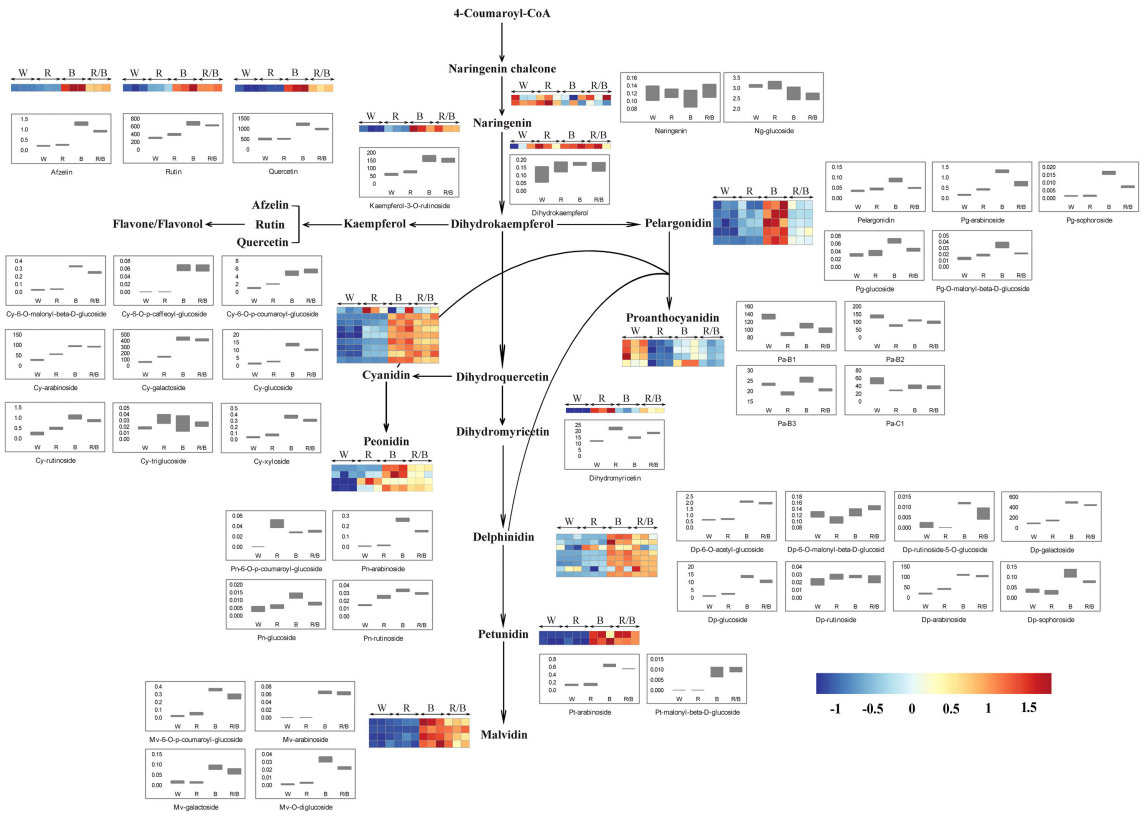


FIGURE 7
Model showing the changes in the abundance of anthocyanin and flavonoid biosynthesis pathway metabolites in blueberry leaves under different light qualities. The model was constructed based on the KEGG pathway and a literary reference. Ng, naringenin; Pg, pelargonidin; Cy, cyanidin; Pa, proanthocyanidin; Pn, peonidin; Dp, delphinidin; Pt, petunidin; Mv, malvidin; W, white; R, red; B, blue; R/B, Red/Blue; Each colored unit represents the normalized intensity of each compound ion according to the color scale (three biological replicates × four treatments, *n*=12). Box-and-whisker plots show the changes in the metabolites for each treatment. Maximum and minimum values of a metabolite among three biological replicates are represented at the upper and lower ends of the whisker, respectively.

anthocyanin accumulation under the different light qualities. Among all anthocyanin derivatives, the quantification of Cy-type, Pg-type, and Mv-type derivatives were more advantageous under blue light than under other light qualities. However, the combination of blue and red light could dramatically improve the Dp-type, Pn-type, and Pt-type derivative content.

Integrated analysis of differentially accumulated metabolites and DEGs in response to different light qualities

To understand the regulatory networks for the anthocyanins that were implicated in the differential distribution of anthocyanin derivatives under the different light qualities, we performed an integrated analysis of the core differentially accumulated metabolites (DAMs) and DEGs related to anthocyanin accumulation in the four comparisons (blue vs. white, red vs. white, red/blue vs. white, and blue vs. red). Based on the above results, the DEGs related to secondary metabolism, carbohydrate metabolism processes(photosynthesis), anthocyanin biosynthesis, and anthocyanin derivatives of Cy, Dp, Ng, Pg, Pa, Pn, Pt, and Mv were used to perform an interaction network based on the Pearson correlation analysis. In this interaction network, genes and metabolites were closely related, and positive and negative correlations between gene expression levels and anthocyanin accumulation content were discovered (Figure 8; Table S8). In contrast to the comparison between red and white (28), there were much stronger correlations between DAMs and DEGs in both comparisons with blue vs. white (57) and red/blue vs. white (56). This result demonstrated that the blue light or blue/red combination may

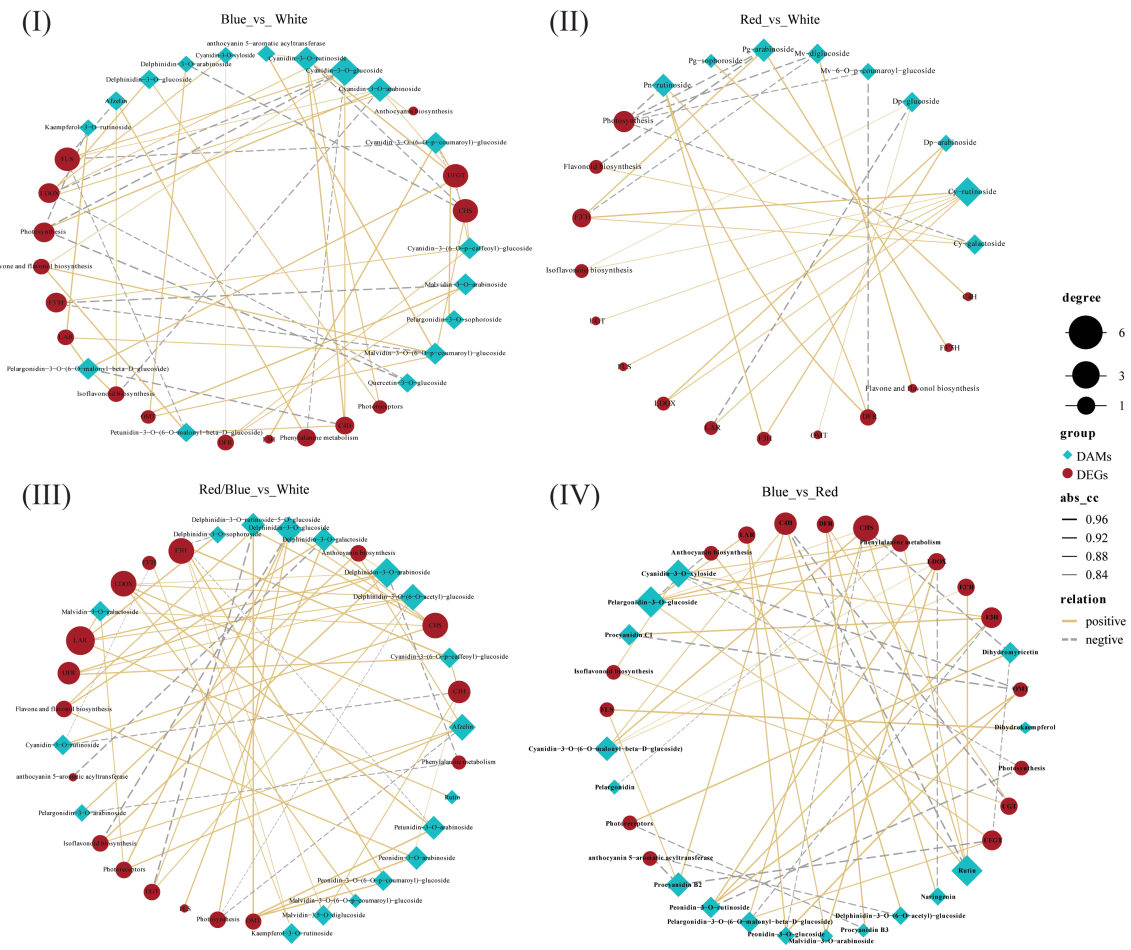


FIGURE 8
Results of the correlation analysis for the differentially expressed genes (DEGs) and differentially accumulated metabolites (DAMs) involved in the anthocyanin biosynthesis pathway of blueberry leaves based on the correlation coefficient (CC) and the p-value for the correlation coefficient (CCP) between DAMs and DEGs. Analysis of **(I)** blue vs. white, **(II)** red vs. white, **(III)** red/blue vs. white, and **(IV)** blue vs. red. The size of the black dots indicates the number of connected DAMs or DEGs; the width of the connecting lines indicates the strength of the correlation; solid lines indicate positive correlations and dashed lines indicate negative correlations.

play a more prominent role in promoting the biosynthesis and accumulation of anthocyanins because there were more DAMs and DEGs related to anthocyanin formation. Because of the changes in DAM content and the DEG transcriptional level, the color and anthocyanin content of the blueberry leaves differed significantly under different light qualities. In the blue vs. white correlation network, there was a close correlation between Cy and *LDOX*, Mv and *OMT*, and Pg and *UFGT* levels. *LDOX*, *OMT*, and *UFGT* are upstream regulatory genes for Cy, Mv, and Pg synthesis in the anthocyanin pathway, and their high expression contributes to Cy, Mv, and Pg accumulation, respectively. Based on the results of the analysis of the DEG transcriptional profile and DAM metabolic profile, the expression levels of *LDOX*, *OMT*, and *UFGT* and the contents of the Cy, Mv, and Pg derivatives were highest under blue light. This validated the idea that blue light induced substantial accumulation of the Cy, Mv, and Pg derivatives *via* the promotion of *LDOX*, *OMT*, and *UFGT* expression, respectively, further facilitating anthocyanin synthesis. A similar phenomenon was also found in the red/blue vs. white correlation network: there were higher expression levels of *DFR* and *OMT*, greater accumulation of Dp, Pn, and Pt derivatives, and an intimate connection between *DFR* and Dp, *OMT* and Pn, and *OMT* and Pt under a combination of red and blue lights. *DFR* and *OMT* induced the synthesis of Dp, Pn, and Pt derivatives, leading to an increase in anthocyanin content under the red/blue light in the blueberry leaves. Additionally, in the blue vs. red correlation network, *UFGT* and *OMT* were closely connected with Pg and Pn, respectively. Among the DEG transcriptional profiles and DAM metabolic profiles with the comparison of blue vs. red, the transcript levels of *UFGT* and *OMT* and the accumulations of Pg and Pn were higher in blue when compared with red. This further indicates that *UFGT* and *OMT* play vital roles in promoting Pg and Pn accumulation, respectively.

Discussion

Blueberry is a commercially popular small fruit crop with leaves and fruit that are abundant in anthocyanins, which are secondary metabolites considered beneficial for human health (Hou et al., 2011; Norberto et al., 2013; Zhou et al., 2013). Under natural growth conditions, the biosynthesis of anthocyanins is dramatically affected by light, including both light quality and intensity. This occurs *via* regulatory gene activity and key metabolites level associated with anthocyanin biosynthesis that can modulate anthocyanin content and pigmentation activity (Dixon and Paiva, 1995; Chalker-Scott, 1999; Jaakola, 2013; Ottosen et al., 2015; Ma et al., 2019), among which, the light quality plays a more critical role (Briggs and Olney, 2001; Zoratti et al., 2014). However, little is known about the molecular mechanisms underlying the changes in anthocyanin content in response to different light qualities in blueberry leaves. In this study, an integrated analysis of the transcriptome and metabolome during anthocyanin biosynthesis under different light qualities was carried out, which allowed us to gain insight into the key metabolites, genes, and metabolic pathways involved in anthocyanin biosynthesis and accumulation in blueberry.

During the development of green tissues (such as leaves, flower buds, and stems) and tissue-cultured cells, light promotes the accumulation of anthocyanins by activating gene transcript levels related to anthocyanin metabolic pathways (Uleberg et al., 2012; Zhang et al., 2021a). In the early stages of anthocyanin synthesis, the expression patterns of related genes can be regulated by different light qualities and intensities, and this expression pattern is dramatically changed and accompanied by distinct species. For example, anthocyanin accumulation within the blueberry peel is significantly affected by light, the content of anthocyanin and expression level of genes related to anthocyanin synthesis pathway were significantly different under various light treatments including: intensity 100%, 50%, and 20% (Guo et al., 2021). Moreover, the differences in the light qualities on the anthocyanin biosynthesis also were significant. Compared with the white control film, the red and yellow films led to a substantial increase in the total anthocyanin content (TAC), while the green and blue films caused a decrease in TAC. The colored film treatment also significantly affects the related enzyme activity and the expression of structural genes and transcription factors in strawberry fruits (Miao et al., 2016). Other studies have shown that the anthocyanin content of purple celery was significantly higher with the blue light treatment when compared with the other light quality conditions (red, green, yellow, and fluorescent lamps) (Field et al., 2001; Neill and Gould, 2003; Hughes et al., 2005). In this study, we measured the anthocyanin content in blueberry leaves under different light qualities and found that the maximal accumulations occurred with blue light, followed by red/blue, red, and white (control). Compared with the white light, the blue, red, and red/blue combination could markedly enhance the accumulation of anthocyanin, but the effects of the other light qualities varied. Blue light was more favorable for the accumulation of anthocyanin than red, red/blue, and white light in blueberry leaves, which is consistent with a previous report on wild-type petunias, in which blue light radiation significantly increased the anthocyanin content in leaves (Nick et al., 2009; Tao et al., 2018), while red light has been reported to promote anthocyanin accumulation in strawberry fruit (*Fragaria ananassa*) (Miao et al., 2016). The impacts of the different light qualities on the anthocyanin content differed markedly between plants and tissues. Additionally, the effects of the red/blue combination were intermediated between blue and red lights, indicating that the functions among the blue and red lights on anthocyanin biosynthesis were redundant.

In this investigation, we discovered that there was an opposing trend between anthocyanin synthesis and photosynthesis: blue light promoted anthocyanin accumulation, but decreased photosynthesis; while photosynthesis was highest and the anthocyanin content lowest in white light when compared with the other light conditions. An increase in the accumulated anthocyanin causes the leaf color to turn from green to light-red as it decreases the chlorophyll content, and reduces the chlorophyll *a/b* ratios, resulting in reduced photosynthesis in broadleaf evergreen species (Hughes and Smith, 2007). In contrast, anthocyanin biosynthesis indirectly affects the circadian leaf starch metabolism and attenuates the sugar-promoted feedback resulting in the down-regulation of photosynthesis. The lower the observed photosynthesis rate decrease in red leaves with higher anthocyanin content than in the green leaves with lower anthocyanin accumulation (Landi et al., 2015; Piccolo et al., 2018). To some degree, the accumulation of anthocyanin inhibits photosynthesis (Piccolo et al., 2020). The blueberry leaves were treated with different light qualities and this affected the anthocyanin content which was significantly higher, but photosynthesis was lower under the blue and red/blue combination lights than under white and red lights. This is in accordance with previous reports, in which an increase in anthocyanin accumulation was accompanied by a reduction in photosynthesis (Murakami et al., 2008). Another explanation is that the abundant accumulation of anthocyanin in the leaves appeared to screen underlying photosynthetic tissues, increasing light saturation and compensation points, reducing the maximal photosynthetic assimilation rate (A_{max}) (Nick et al., 2009). Notwithstanding the insights proposed by these investigations, the correlation and regulatory mechanism between anthocyanin accumulation and photosynthesis is yet to be defined under different light qualities, and future studies on blueberry leaves and other plant tissues will be required.

Recently, the integration of large-scale datasets derived from high-throughput functional genomics techniques has been successfully applied to study the functions of genes regulating tissue development, environmental responses, and plant metabolism (An et al., 2020; Raza et al., 2021). Comparative transcriptome analysis of bilberry (*V. myrtillus* L.), for example, under different light qualities (red and blue) has provided information about the gene transcript level and regulation related to anthocyanin biosynthesis (Samkumar et al., 2021). Our results showed that the majority of the DEGs in the flavonoid and anthocyanin biosynthesis pathways were up- and down-regulated under different light qualities in blueberry leaves using RNA-Seq. Anthocyanins are biosynthesized *via* a branch of the flavonoid pathway, and their synthesis is affected by the presence or absence of upstream genes. The anthocyanin biosynthesis pathway is relatively clear in plants, and related key structural genes, including *C4H*, *4CL*, *CHS, CHI*, *OMT*, *F3'H*, *F3'5H*, *DFR*, *LDOX*, *LAR*, *UFGT*, *UGT*, and *FLS*, which have been identified in the colored tissues of several plants (Gao et al., 2021). Blue light has been identified as a strong positive influence on anthocyanin biosynthesis in numerous fruit crops, for instance, in sweet cherries (Kokalj et al., 2019). Red light can also do it in similar ways (Zhang et al., 2018). Our results showed that blue and the red/blue light combination up-regulated most of the anthocyanin biosynthesis genes, including *CHS*, *CHI, F3H*, *DFR*, *LDOX*, *OMT*, and *UFGT*, in blueberry leaves, and this led to a higher accumulation of anthocyanin under both blue and red/blue light combination treatments when compared with white light. It should be noted that multiple unigenes were annotated as the same enzyme in this study. For example, three CHS-encoding genes and two DFR-encoding genes were identified. The main reason for this was that these unigenes belong to different selective splicing transcripts as well as a specific gene family (Yang et al., 2005).

The photoreceptor can induce the accumulation of anthocyanin by promoting the expression of anthocyanin biosynthesis genes. Red light, acting *via* phytochromes, stimulates PAL activity in the cotyledons and hypocotyls of tomato seedlings, and exposure to UV-B has been shown to stimulate PAL activity, thus, increasing anthocyanin content in rice, maize, and turnip plants (Brodenfeldt and Mohr, 1988; Reddy et al., 1994; Sharma et al., 1999). Phytochrome B has a specialized function in red light, whereas cryptochrome senses blue light to promote secondary metabolite synthesis and anthocyanin accumulation. Previous studies have reported that UV light regulates the synthesis of anthocyanins through different photoreceptors, including the UV-B receptor UV RESISTANCE LOCUS 8 (UVR8) (Brown et al., 2005; Christie et al., 2012; Giliberto et al., 2005). In this study, genes related to photoreceptors differed in their expression levels under different light qualities and subsequently played a potential role in anthocyanin accumulation. The DEGs related to phytochromes were highly expressed in response to red light, whereas cryptochrome-related DEGs were substantially expressed under blue light. These results indicate that the transcript levels of these photoreceptor genes were relatively high under their corresponding light quality, while the lowest expression level of the photoreceptor-related genes was observed under white light. Blue and red light may promote anthocyanin biosynthesis by inducing the expression of photoreceptor-related genes to stimulate the gene promoter activity involved in the anthocyanin biosynthesis pathway. Photosynthesis is an important physiological process in plants under light conditions, and the photosynthesis-related DEGs showed an opposite dynamic trend with DEGs related to anthocyanin synthesis in gene expression patterns, which was consistent with their corresponding physiological index content. This result may further confirm the above conjecture that anthocyanin accumulation may restrain photosynthesis to a certain extent in blueberry leaves.

In several cases, changes in the expression of genes related to anthocyanin biosynthesis correspond to changes in the

content of anthocyanin derivatives in this pathway (Kyoungwon et al., 2016). Consistent with the transcriptome results, the level of metabolite accumulation in the anthocyanin biosynthesis pathway differed substantially under different light qualities. Anthocyanins are glycosides and acylglycosides of anthocyanidin aglycones that are biosynthesized through the flavonoid pathway *via* the phenylpropanoid pathway (Stushnoff et al., 2010; Jaakola,

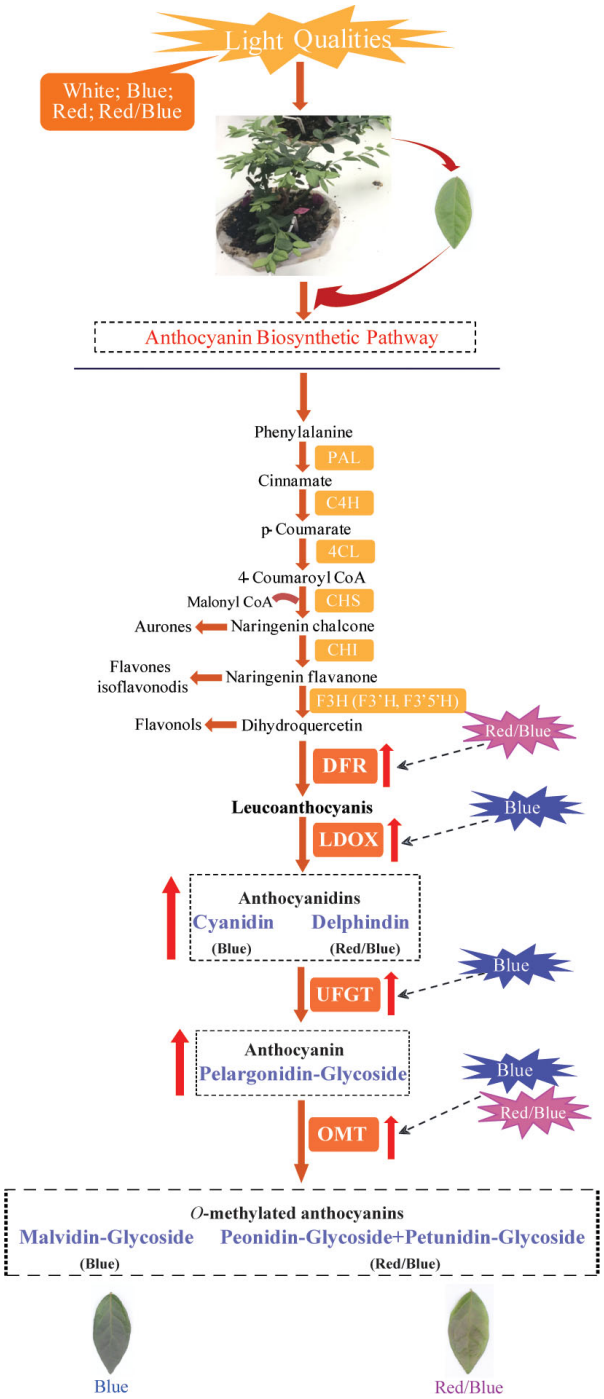


FIGURE 9
Proposed model of the molecular regulatory mechanism underlying anthocyanin biosynthesis in blueberry leaves in response to different light qualities. The red up-arrow indicates a significant increase; the dotted line arrow indicates a potential role for positive regulation.

2013). Among these metabolic pathways, cyanidin (Cy), delphinidin (Dp), pelargonidin (Pg), peonidin (Pn), petunidin (Pt), and malvidin (Mv) are the primary metabolic derivatives that play vital roles in anthocyanin pigmentation (Tahara, 2007). In this study, the accumulation levels of these derivatives differed under different light qualities; blue and red/blue could promote the synthesis of more metabolites than red and white lights could, which was consistent with the total anthocyanin content and leaf color under different light qualities. Among these derivatives, the accumulation of Cy, Pg, and Mv derivatives and Dp, Pn, and Pt derivatives were significantly higher under blue light and the red/blue light combination, respectively (Figure 7). It has been reported that the six anthocyanidins that generally occur in colored tissues are Cy, Pg, Mv, Dp, Pn, and Pt and that they play vital roles in anthocyanin pigmentation and biosynthesis. Additionally, in the correlation network between the transcriptome and metabolome of blue vs. white, Cy, Pg, and Mv were found to be closely associated with *LDOX*, *UFGT*, and *OMT*, respectively. Previous studies have reported that individual enzymes also have an impact on the overall stability of the biosynthesis pathway, as the silencing, over-expression, or heterologous expression of single enzyme genes often leads to substantial changes in the anthocyanin composition of the target tissue (Griesser et al., 2008; Han et al., 2010; Han et al., 2012). Therefore, as regulative downstream materials, the high expression of *LDOX*, *UFGT*, and *OMT* resulting in the abundant accumulation of Cy, Pg, and Mv under blue light in blueberry leaves. Similarly, in the red/blue vs. white correlation network, similar relationships were observed between Dp, Pn, and Pt and *DFR* and *OMT*. The expression levels of *DFR* and *OMT* were higher, when accompanied by the abundant synthesis of Dp, Pn, and Pt with the combination of the red and blue lights. This is consistent with a previous report by Kyoungwon et al. (2016), in which the Dp derivative was shown to positively correlate with the expression level of genes, including *DFR*, *CHS*, and *CHI*; consequently, the more significant up-regulation of these genes has a role in the more significant accumulation of the Dp derivative. *OMT*, a vital enzyme gene in the anthocyanin biosynthesis pathway, catalyzes the formation of *O*-methylated anthocyanins, such as Mv, Pn, and Pt (Koes et al., 2005). High levels of *OMT* expression can induce the accumulation of Mv, Pn, and Pt under blue and red/blue light. Thus, we concluded that blue and red/blue lights induce the activity of the DEGs associated with anthocyanin biosynthesis and consequently, this facilitates the accumulation of anthocyanins in blueberry leaves under blue and red/blue lights, respectively.

Conclusion

In this study, blue and red/blue, and red lights were found to promote anthocyanin accumulation in blueberry leaves, and there was a redundant function that facilitated anthocyanin biosynthesis and accumulation between the blue and red lights. Additionally, correlation analysis of the changes in gene expression and metabolite levels suggested that blue light induced substantial *LDOX*, *UFGT*, and *OMT* expression to promote the accumulation of cyanidin, pelargonidin, and malvidin anthocyanidins; while the combination of red and blue light facilitated the up-regulation of *DFR* and *OMT*, leading to the accumulation of large amounts of downstream metabolites, including delphinidin, petunidin, and peonidin derivatives (Figure 9). Meanwhile, there was a sharp adverse dynamic trend between anthocyanin accumulation and photosynthesis in the blueberry leaves under different light quality treatments. The results obtained in this study provide new insights into the molecular mechanisms underlying anthocyanin biosynthesis and how the inter-correlation between anthocyanin accumulation and photosynthesis is regulated.

Data availability statement

The data presented in the study are deposited in the NCBI repository with the link: https://www.ncbi.nlm.nih.gov/, accession number PRJNA880737.

Author contributions

XZ and HA organized the entire project. JZ, HA, and SL performed the experiments and data analysis SL and BZ helped to review and modify the manuscript. XZ and HA edited the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by Key Scientific Project of Shanghai Science and Technology Commission (No. 17391900800) and Excellent Team Program of Shanghai Academy of Agricultural Sciences (No. 2022-022).

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1073332/full#supplementary-material

References

Allan, A. C., Hellens, R. P., and Laing, W. A. (2008). MYB transcription factors that colour our fruit. *Trends Plant Sci.* 13, 99–102. doi: 10.1016/j.tplants.2007.11.012

An, H. S., Zhang, J. Y., Xu, F. J., Jiang, S., and Zhang, X. Y. (2020). Transcriptomic profiling and discovery of key genes involved in adventitious root formation from green cuttings of highbush blueberry (*Vaccinium corymbosum* l.). *BMC Plant Biol.* 20, 182–195. doi: 10.1186/s12870-020-02398-0

Batschauer, A., Rocholl, M., Kaiser, T., Nagatani, A., Furuya, M., and Schafer, E. (1996). Blue and UV-a light-regulated CHS expressionin arabidopsis independent of phytochrome a and phytochromeB. *Plant J.* 9, 63–69. doi: 10.1046/j.1365-313X.1996.09010063.x

Briggs, W. R., and Olney, M. A. (2001). Photoreceptors in plant photomorphogenesis to date. five phytochromes, two cryptochromes, one phototropin, and one superchrome. *Plant Physiol.* 125, 85–88. doi: 10.1104/pp.125.1.85

Brodenfeldt, R., and Mohr, H. (1988). Time courses for phytochrome-induced enzyme levels in phenylpropanoid metabolism (phenylalanine ammonia-lyase, naringenin-chalcone synthase) compared with time courses for phytochrome-mediated end-product accumulation (anthocyanin, quercetin). *Planta.* 176, 383–390. doi: 10.1007/BF00395419

Brown, B. A., Cloix, C., Jiang, G. H., Kaiserli, E., Herzyk, P., and Kliebenstein, D. J. (2005). A UV-b-specific signalling component orchestrates plant UV protection. *P. Nat. Acad. Sci. U.S.A.* 102, 18225–18230. doi: 10.1073/pnas.0507187102

Chai, Z., Herrera-Balandrano, D. D., and Yu, H. A. (2021). Comparative analysis on the anthocyanin composition of 74 blueberry cultivars from China. *J. Food Compos. Anal.* 4, 104051. doi: 10.1016/j.jfca.2021.104051

Chalker-Scott, L. (1999). Environmental significance of anthocyanins in plant stress responses. *Photochem. Photobiol.* 70, 1–9. doi: 10.1111/j.1751-1097.1999.tb01944.x

Christie, J. M., Arvai, A. S., Baxter, K. J., Heilmann, M., Pratt, A. J., and O'Hara, A. (2012). Plant UVR8 photoreceptor senses UV-b by tryptophan-mediated disruption of cross-dimer salt bridges. *Sci.* 335, 1492–1496. doi: 10.1126/science.1218091

Dixon, R. A., and Paiva, N. L. (1995). Stress-induced phenylpropanoid metabolism. *Plant Cell.* 7, 1085–1097. doi: 10.1105/tpc.7.7.1085

Ehlenfeldt, M. K., and Prior, R. L. (2001). Oxygen radical absorbance capacity (ORAC) and phenolic and anthocyanin concentrations in fruit and leaf tissues of highbush blueberry. *J. Agric. Food Chem.* 49, 2222–2227. doi: 10.1021/jf0013656

Fallah, A. A., Sarmast, E., Fatehi, P., and Jafari, T. (2020b). Impact of dietary anthocyanins on systemic and vascular inflammation: systematic review and meta-analysis on randomized clinical trials. *Food Chem.* 135, 110922. doi: 10.1016/j.fct.2019.110922

Fallah, A. A., Sarmast, E., and Jafari, T. (2020a). Effect of dietary anthocyanins on biomarkers of glycemic control and glucose metabolism: a systematic review and meta-analysis of randomized clinical trials. *Food Res. Int.* 137, 109379. doi: 10.1016/j.foodres.2020.109379

Feinbaum, R. L., Storz, G., and Ausbel, F. M. (1991). High intensity and blue light regulated expression of chimeric chalcone synthase genes in transgenic arabidopsis thaliana plants. *Mol. Gen. Genet.* 226, 449–456. doi: 10.1007/BF00260658

Field, T., Lee, D., and Holbrook, N. (2001). Why leaves turn red in autumn. the role of anthocyanins in senescing leaves of red-osier dogwood. *Plant Physiol.* 127, 566–574. doi: 10.1104/pp.900027

Gao, X., Chen, Q., and Yang, Y. (2021). Transcriptomic analysis reveals anthocyanin biosynthesis regulation in blueberry (*Vaccinium ashei*) fruit. *Cana. J. Plant Sci.* 102, 195–206. doi: 10.1139/CJPS-2020-0118

Giliberto, L., Perrotta, G., Pallara, P., Weller, J. L., Fraser, P. D., Bramley, P. M., et al. (2005). Manipulation of the blue light photoreceptor cryptochrome 2 in tomato affects vegetative development, flowering time, and fruit antioxidant content. *Plant Physiol.* 137, 199–208. doi: 10.1104/pp.104.051987

Griesser, M., Hoffmann, T., Bellido, M. L., Rosati, C., Fink, B., Kurtzer, R., et al. (2008). Redirection of flavonoid biosynthesis through the down-regulation of an anthocyanidin glucosyltransferase in ripening strawberry fruit. *Plant Physiol.* 146, 1528–1539. doi: 10.1104/pp.107.114280

Guo, X. L., Wang, D. L., and Shakeel, M. (2021). Transcriptome analysis reveals light-induced anthocyanin synthesis candidate genes in rabbiteye blueberry (*Vaccinium ashei*: Reade). *Biotechnol. Biotec. Eq.* 35 (1), 747–758. doi: 10.1080/13102818.2021.1924078

Guo, J. Y., Wu, G., Wang, T., and Wang, F. (2020). Integrated analysis of the transcriptome and metabolome in young and mature leaves of ginkgo biloba l. *Ind. Crops Prod.* 143, 111906. doi: 10.1016/j.indcrop.2019.111906

Han, Y. P., Vimolmangkang, S., Soria-Guerra, R. E., and Korban, S. S. (2012). Introduction of apple *ANR* genes into tobacco inhibits expression of both *CHI* and *DFR* genes in flowers, leading to loss of anthocyanin. *J. Exp. Bot.* 63, 2437–2447. doi: 10.1093/jxb/err415

Han, Y., Vimolmangkang, S., Soria-Guerra, R. E., Rosales-Mendoza, S., Zheng, D. M., and Lygin, V. A. (2010). Ectopic expression of apple *F30H* genes contributes to anthocyanin accumulation in the arabidopsis tt7 mutant grown under nitrogen stress. *Plant Physiol.* 153, 806–820. doi: 10.1104/pp.109.152801

Hou, S. Z., Chen, S. X., Huang, S., Jiang, D. X., Zhou, C. J., Chen, C. Q., et al. (2011). The hypoglycemic activity of *Lithocarpuspo lystachyus rehd.* leaves in the experimental hyperglycemic rats. *J. Ethnopharmacol.* 138, 142–149. doi: 10.1016/j.jep.2011.08.067

Hou, S. Z., Xu, S. J., Jiang, D. X., Chen, S. X., Wang, L. L., Huang, S., et al. (2012). Effect of the flavonoid fraction of *Lithocarpuspo lystachyus rehd.* on spontaneously hypertensive and normotensive rats. *J. Ethnopharmacol.* 143 (2), 441–447. doi: 10.1016/j.jep.2012.06.016

Hughes, N. M., Neufeld, H. S., and Burkey, K. O. (2005). Functional role of anthocyanins in high-light winter leaves of the evergreen herb galax urceolata. *New Phytol.* 168, 575–587. doi: 10.1111/j.1469-8137.2005.01546.x

Hughes, N. M., and Smith, W. K. (2007). Seasonal photosynthesis and anthocyanin production in 10 broad leaf evergreen species. *Funct. Plant Biol.* 34, 1072–71079. doi: 10.1071/FP07205

Jaakola, L. (2013). New insights into the regulation of anthocyanin biosynthesis in fruits. *Trends Plant Sci.* 18 (9), 477–483. doi: 10.1016/j.tplants.2013.06.003

Kalt, W., Cassidy, A., Howard, L. R., Krikorian, R., Stull, A. J., and Tremblay, F. (2020). Recent research on the health benefits of blueberries and their anthocyanins. *Adv. Nutr.* 11, 224–236. doi: 10.1093/advances/nmz065

Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., et al. (2008). KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* 36 (746), 480–484. doi: 10.1093/nar/gkm882

Koes, R., Verweij, W., and Quattrocchio, F. (2005). Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. *Trends Plant Sci.* 10, 236–242. doi: 10.1016/J.TPLANTS.2005.03.002

Kokalj, D., Zlatic, E., Cigic, B., and Vidrih, R. (2019). Postharvest light-emitting diode irradiation of sweet cherries (*Prunus avium* l.) promotes accumulation of anthocyanins. *Postharvest Biol. Tec.* 148, 192–199. doi: 10.1016/j.postharvbio.2018.11.011

Kyoungwon, C., Kwang-Soo, C., and Hwang-Bae, S. (2016). Network analysis of the metabolome and transcriptome reveals novel regulation of potato pigmentation. *J. Exp. Bot.* 5, 1519–1533. doi: 10.1093/jxb/erv549

Landi, M., Tattini, K., and Gould, S. (2015). Multiple functional roles of anthocyanins in plant environment interactions. *Environ. Exp. Bot.* 119, 4–17. doi: 10.1016/j.envexpbot.2015.05.012

Li, K. X., Liu, K., and Chen, Y. (2021). Comprehensive transcriptome and metabolome analysis of *Lithocarpuspo lystachyus* leaf revealed key genes in flavonoid biosynthesis pathways. *J. Amer. Soc Hortic. Sci.* 146, 147–157. doi: 10.21273/JASHS05020-20

Li, Y. Y., Mao, K., Zhao, C., and Zhao, X. Y. (2013). Molecular cloning and functional analysis of a blue light receptor gene *MdCRY2* from apple (*Malus domestica*). *Plant Cell Rep.* 32, 555–566. doi: 10.1007/s00299-013-1387-4

Li, Y. Y., Mao, K., Zhao, C., Zhao, X. Y., Zhang, H. L., Shu, H. R., et al. (2012). MdCOP1 ubiquitin E3 ligases interact with MdMYB1 to regulate light-induced anthocyanin biosynthesis and red fruit coloration in apple. *Plant Physiol.* 160 (2), 1011–1022. doi: 10.1104/pp.112.199703

Lin, Y., Wang, Y., Li, B., Tan, H., Li, D. N., and Li, L. (2018). Comparative transcriptome analysis of genes involved in anthocyanin synthesis in blueberry. *Plant Physiol. Biochem.* 127, 561–572. doi: 10.1016/j.plaphy.2018.04.034

Liu, Y., Kui, L. W., Espley, R. V., Li, W., Yang, H., and Yu, B. (2016). Functional diversification of the potato R2R3MYB anthocyanin activators AN1, MYBA1, and MYB113 and their interaction with basic helix-loop-helix cofactors. *J. Exp. Bot.* 67, 2159–2176. doi: 10.1093/jxb/erw014

Li, Q. H., and Yang, H. Q. (2007). Cryptochrome signaling in plants. *Photochem. Photobiol.* 83, 94–101. doi: 10.1562/2006-02-28-IR-826

Li, L., Zhang, H., Liu, Z., Cui, X., Zhang, T., Li, Y., et al. (2016). Comparative transcriptome sequencing and *de novo* analysis of *Vaccinium corymbosum* during fruit and color development. *BMC Plant Biol.* 16 (1), 223. doi: 10.1186/s12870-016

Lou, Q., Liu, Y., and Qi, Y. (2014). Transcriptome sequencing and metabolite analysis reveals the role of delphinidin metabolism in flower colour in grape hyacinth. *J. Exp. Bot.* 12, 3157–3164. doi: 10.1093/jxb/eru168

Lu, X. D., Zhou, C. M., Xu, P. B., Luo, Q., Lian, H. L., and Yang, H. Q. (2015). Red-light-dependent interaction of phyB with SPA1 promotes COP1-SPA1 dissociation and photomorphogenic development in *Arabidopsis*. *Mol. Plant* 8, 467–478. doi: 10.1016/j.molp.2014.11.025

Lv, Y., Xu, L., Dossa, K., Zhou, K., Zhu, M., Xie, H., et al. (2019). Identification of putative drought-responsive genes in rice using gene coexpression analysis. *Bioinformation.* 15, 480–489. doi: 10.6026/97320630015480

Ma, Z. H., Li, W. F., Mao, J., Li, W., Zuo, C. W., Zhao, X., et al. (2019). Synthesis of light-inducible and light-independent anthocyanins regulated by specific genes in grape 'Marselan' (*V. vinifera l.*). *Peer J.* 3, 1–24. doi: 10.7717/peerj.6521

Miao, L., Zhang, Y., and Yang, X. (2016). Colored light-quality selective plastic films affect anthocyanin content, enzyme activities, and the expression of flavonoid genes in strawberry (*Fragaria ananassa*) fruit. *Food Chem.* 207, 93–100. doi: 10.1016/j.foodchem.2016.02.077

Mol, J., Jenkins, G. I., Schafer, E., and Weiss, D. (1996). Signal perception, transduction, and gene expression involved in anthocyanin biosynthesis. *Crit. Rev. Plant Sci.* 15, 525–557. doi: 10.1080/07352689609382369

Murakami, P. F., Schaberg, P. G., and Shane, J. B. (2008). Stem girdling manipulates leaf sugar concentrations and anthocyanin expression in sugar maple trees during autumn. *Tree Physiol.* 28, 1467–1473. doi: 10.1093/treephys/28.10.1467

Nagao, K., Higa, K., Shirouchi, B., Nomura, S., Inoue, N., and Inafuku, M. (2008). Effect of *Vaccinium ashei reade* leaves on lipid metabolism in otsuka long-Evans tokushima fatty rats. *Biosci. Biotechno. Biochem.* 72, 1619–1622. doi: 10.1271/bbb.80036

Neill, S. O., and Gould, K. S. (2003). Anthocyanins in leaves: light attenuators or antioxidants. *Funct. Plant Biol.* 30, 865–873. doi: 10.1071/fp03118

Nick, W. A., David, H. L., Zhang, H. B., Louis, J. I., Paula, E. J., and Kevin, M. D. (2009). Light-induced vegetative anthocyanin pigmentation in petunia. *J. Exp. Bot.* 60, 2191–2202. doi: 10.1093/jxb/erp097

Noh, B., and Spalding, E. P. (1998). Anion channels and the stimulation of anthocyanin accumulation by blue light in *Arabidopsis* seedlings. *Plant Physiol.* 116, 503–509. doi: 10.1104/pp.116.2.503

Norberto, S., Silva, S., Meireles, M., Faria, A., Pintado, M., and Calhau, C. (2013). Blueberry anthocyanins in health promotion: a metabolic overview. *J. Funct. Foods.* 5, 1518–1528. doi: 10.1016/j.jff.2013.08.015

Ordidge, M., García-Macías, P., Battey, N. H., Gordon, M. H., John, P., Vysini, E., et al. (2011). Development of colour and firmness in strawberry crops is UV light sensitive, but colour is not a good predictor of several quality parameters. *J. Sci. Food Agric.* 92, 1597–1604. doi: 10.1002/jsfa.4744

Ottosen, C. O., Ouzounis, T., and Rosenqvist, E. (2015). Spectral effects of artificial light on plant physiology and secondary metabolism: A review. *Hortic. Sci.* 50, 1128–1135. doi: 10.21273/HORTSCI.50.8.1128

Piccolo, E. L., Landi, M., and Massai, R. (2020). Girled-induced anthocyanin accumulation in red-leafed *Prunus cerasifera*: Effect on photosynthesis, photoprotection and sugar metabolism. *Plant Sci.* 294, 110456. doi: 10.1016/j.plantsci.2020.110456

Piccolo, E. L., Landi, M. E., Pellegrini, G., Agati, C., Giordano, T., Giordani, G., et al. (2018). Multiple consequences induced by epidermally-located anthocyanins in young, mature and senescent leaves of *Prunus*. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.00917

Raza, A., Su, W., and Hussain, M. A. (2021). Integrated analysis of metabolome and transcriptome reveals insights for cold tolerance in rapeseed (*Brassica napus l.*). *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.721681

Reddy, V. S., Goud, K. V., Sharma, R., and Reddy, A. R. (1994). UV-B responsive anthocyanin production in a rice cultivar is associated with a specific phase of phenylalanine ammonia lyase biosynthesis. *Plant Physiol.* 105, 1059–1066. doi: 10.1104/pp.105.4.1049

Sakaida, H., Nagao, K., Higa, K., Shirouchi, B., Inoue, N., and Hidaka, F. (2007). Effect of *Vaccinium ashei reade* leaves on angiotensin converting enzyme activity *in vitro* and on systolic blood pressure of spontaneously hypertensive rats *in vivo*. *Biosci. Biotechno. Biochem.* 71, 2335–2337. doi: 10.1271/bbb.70277

Samkumar, A., Jones, D., Karppinen, J., Andrew, P. D., Nina, S., and Richard, V. E. (2021). Red and blue light treatment of ripening bilberry fruits reveal differences in signaling through abscisic acid-regulated anthocyanin biosynthesis. *Plant Cell Environ.* 44, 3227–3245. doi: 10.1111/pce.14158

Seeram, N. P., Adams, L. S., Zhang, Y. J., Lee, R., Sand, D., Scheuller, H. S., et al. (2006). Blackberry, black raspberry, blueberry, cranberry, red raspberry, and strawberry extracts inhibit growth and stimulate apoptosis of human cancer cells *in vitro*. *J. Agric. Food Chem.* 54, 9329–9339. doi: 10.1021/jf061750g

Sharma, R., Singh, A., and Selvi, M. (1999). Sun light-induced anthocyanin pigmentation in maize vegetative tissues. *J. Exp. Bot.* 50, 1619–1625. doi: 10.1093/jexbot/50.339.1619

Song, Y., Huang, L., and Yu, J. (2016). Effects of blueberry anthocyanins on retinal oxidative stress and inflammation in diabetes through *Nrf2/HO-1* signaling. *J. Neuroimmunol.* 301, 1–6. doi: 10.1016/j.jneuroim.2016.11.001

Stracke, R., Favory, J. J., Gruber, H., Bartelniewoehner, L., and Ulm, R. (2010). The arabidopsis bZIP transcription factor HY5 regulates expression of the PFG1/MYB12 gene in response to light and ultraviolet-b radiation. *Plant Cell Environ.* 33, 88–103. doi: 10.1111/j.1365-3040.2009.02061.x

Stushnoff, C., Ducreux, L. J., Hancock, R. D., Hedley, P. E., Holm, D. G., McDougall, G. J., et al. (2010). Flavonoid profiling and transcriptome analysis reveals new gene-metabolite correlations in tubers of *Solanum tuberosum l. J. Exp. Bot.* 61, 1225–1238. doi: 10.1093/jxb/erp394

Tahara, S. (2007). A journey of twenty-five years through the ecological biochemistry of flavonoids. *Biosci. Biotechnol. Biochem.* 71, 1387–1404. doi: 10.1271/bbb.70028

Takeshita, M., Ishida, Y. I., and Akamatsu, E. (2009). Proanthocyanidin from blueberry leaves suppresses expression of subgenomic hepatitis c virus RNA. *J. Biol. Chem.* 284, 21165–21176. doi: 10.1074/jbc.M109.004945

Takos, A. M., Felix, W., Jacob, S. R., Bogs, J., Robinson, S. P., and Walker, A. R. (2006). Light-induced expression of a MYB gene regulates anthocyanin biosynthesis in red apples. *Plant Physiol.* 142, 1216–1232. doi: 10.1104/PP.106.088104

Tao, R., Bai, S., Ni, J., Yang, Q., Zhao, Y., and Teng, Y. (2018). The blue light signal transduction pathway is involved in anthocyanin accumulation in 'Red zaosu' pear. *Planta.* 248, 37–48. doi: 10.1007/s00425-018-2877-y

Uleberg, E., Rohloff, J., and Jaakola, L. (2012). Effects of temperature and photoperiod on yield and chemical composition of northern and southern clones of bilberry (*Vaccinium myrtillus l.*). *J. Agric. Food Chem.* 60, 10406–10414. doi: 10.1021/jf302924m

Wang, H., Guo, X., Hu, X., Li, T., Fu, X., and Liu, R. H. (2017). Comparison of phytochemical profiles, antioxidant and cellular antioxidant activities of different varieties of blueberry (*Vaccinium* spp.). *Food Chem.* 217, 773–781. doi: 10.1016/j.foodchem.2016.09.002

Yang, T., Li, J., Wang, H. X., and Zeng, Y. (2005). A geraniol-synthase gene from cinnamomum tenuipilum. *Phytochem.* 66, 285–293. doi: 10.1016/j.phytochem.2004.12.004

Zhang, G., Cui, X., and Niu, J. (2021a). Visible light regulates anthocyanin synthesis *via* malate dehydrogenases and the ethylene signaling pathway in plum (*Prunus salicina l.*). *Physiol. Plantarum.* 83, 1–11. doi: 10.1111/ppl.13383

Zhang, X., Hong, M., Wan, H., Luo, L., Yu, Z., and Guo, R. (2019). Identification of key genes involved in embryo development and differential oil accumulation in two contrasting maize genotypes. *Genes (Basel).* 10, 993. doi: 10.3390/genes10120993

Zhang, Y., Jiang, L., Li, Y., Chen, Q., Ye, Y., Zhang, Y., et al. (2018). Effect of red and blue light on anthocyanin accumulation and differential gene expression in strawberry (*Fragaria ananassa*). *Molecules.* 23, 820. doi: 10.3390/molecules23040820

Zhang, L. Q., Jiang, S., Meng, J. J., An, H. S., and Zhang, X. Y. (2019). First report of leaf spot caused by nigrosporaoryzae on blueberry in shanghai. *China Plant Dis.* 103, 2473. doi: 10.1094/PDIS-02-19-0242-PDN

Zhang, Y., Lin, L., Long, Y., Guo, H., Wang, Z., Cui, M., et al. (2019). Comprehensive transcriptome analysis revealed the effects of the light quality, light intensity, and photoperiod on phloriz in accumulation in *Lithocarpuspo lystachyus* rehd. *Forests.* 10, 995. doi: 10.3390/f10110995

Zhang, J. Y., Li, J. M., and Xue, C. (2021b). The variation of stone cell content in 236 germplasms of sand pear (*Pyrus pyrifolia*) and identification of related candidate genes. *Hortic. Plant J.* 7, 108–116. doi: 10.1016/j.hpj.2020.09.003

Zhou, C. J., Huang, S., Liu, J. Q., Qiu, S. Q., Xie, F. Y., Song, H. P., et al. (2013). Sweet tea leaves extract improves leptin resistance in diet-induced obese rats. *J. Ethnopharmacol.* 145, 386–392. doi: 10.1016/j.jep.2012.09.057

Zoratti, L., Sarala, M., Carvalho, E., Karppinen, K., Martens, S., Giongo, L., et al. (2014). Monochromatic light increases anthocyanin content during fruit development in bilberry. *BMC Plant Biol.* 14, 377. doi: 10.1186/s12870-014-0377-1

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 09 December 2022
DOI 10.3389/fpls.2022.1089392

OPEN ACCESS

EDITED BY
Chenxia Cheng,
Qingdao Agricultural University, China

REVIEWED BY
Xiangpeng Leng,
Qingdao Agricultural University, China
Songtao Jiu,
Shanghai Jiao Tong University, China

*CORRESPONDENCE
Li Wang
vivi@hebau.edu.cn
Guoqiang Du
gdu@hebau.edu.cn

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

SPECIALTY SECTION
This article was submitted to Functional and Applied Plant Genomics, a section of the journal Frontiers in Plant Science

RECEIVED 04 November 2022
ACCEPTED 25 November 2022
PUBLISHED 09 December 2022

CITATION
Xiang R, Ahmad B, Liang C, Shi X, Yang L, Du G and Wang L (2022) Systematic genome-wide and expression analysis of RNA-directed DNA methylation pathway genes in grapes predicts their involvement in multiple biological processes. *Front. Plant Sci.* 13:1089392. doi: 10.3389/fpls.2022.1089392

Systematic genome-wide and expression analysis of RNA-directed DNA methylation pathway genes in grapes predicts their involvement in multiple biological processes

Rui Xiang[1†], Bilal Ahmad[1,2†], Chen Liang[1], Xiaoxin Shi[1], Lili Yang[3], Guoqiang Du[1*] and Li Wang[1*]

[1]College of Horticulture, Hebei Agricultural University, Baoding, China, [2]Department of Horticulture, Muhammad Nawaz Sharif (MNS)-University of Agriculture Multan, Multan, Pakistan, [3]Shijiazhuang Fruit Research Institute, Hebei Academy of Agricultural and Forestry Sciences, Shijiazhuang, China

RNA-directed DNA methylation (RdDM) is an important epigenetic pathway in plants and mediates transcriptional silencing by siRNAs. Different gene families have role in the regulation of the RdDM pathway and there is a lack of information about these gene families in the grapes (*Vitis vinifera* L.). Here, we mentioned the genome-wide identification, bioinformatics analysis, evolutionary history, and expression profiling of *VvRdDM* pathway genes against various stresses, hormonal treatments as well as in different organs. Sixty *VvRdDM* genes belonging to fourteen different families were identified. All the genes were unevenly distributed and chromosome 4 contained the highest number of genes (7). Most of the genes showed similar exon-intron and motif distribution patterns within the same subfamilies. Out of 14 families, only members of 4 families underwent duplication events during the evolutionary process and 50% of members of the AGO family are the result of duplication events. Based on Ka/Ks ratio all duplicated gene pairs have a negative mode of selection. *VvRdDM* pathway genes showed differential spatiotemporal expression patterns against different hormone and stress treatments. Further, with multiple transcriptome analysis, some *VvRdDM* genes showed a broad spectrum of high expression in different organs at various stages, and *VvRdDM* genes also displayed different expression in seeded and seedless cultivars during different phases of seed development. This proposed that *VvRdDM* genes may play multiple roles in grape growth and development, especially in seed development. qRT-PCR analysis of selected genes further verified the critical roles of RdDM genes in multiple biological processes, especially in seed development/ovule abortion i.e., *VvIDN2a*, *VvDRD1a*, *VvRDR1a*, and *VvRDR6*. Our study provides detailed information about *VvRdDM* genes in perspective of

gene structure and evolution, as well as expression pattern against different stress, hormones and in different plants parts. It provides new candidate gene resources for further functional characterization and molecular breeding of grapes.

KEYWORDS

RdDM, seed development, grapes, DNA methylation, expression analysis

Introduction

Epigenetics investigates heritable changes in gene function that occur without a change in DNA sequence (Holliday, 2006). Most of the epigenetic changes (Histone modification, DNA methylation, chromatin remodeling, and chemical modification) occur due to microRNAs (miRNAs). The epigenetic mechanisms are tightly connected to cellular development and differentiation (Barter et al., 2012). DNA methylation is a conserved epigenetic silencing process that is involved in a variety of biological functions, including the regulation of gene expression, prevention of the spread of transposons, and defense against these organisms. Additionally, changes in DNA methylation can result in abnormal development (Zhang et al., 2018). Mostly DNA methylation in plants occurs in the CG, CHG, and CHH sequences (where "H" can be either an A, C, or T), and it is especially prevalent over heterochromatic transposable elements (TEs) and repetitions, where it is crucial for transcriptional gene silencing (TGS) (Zhang et al., 2018). However, many plant species have DNA methylation over gene bodies but in this case precise role is still unknown (Bewick and Schmitz, 2017).

In plants, DNA methylation usually happens by RNA-directed DNA methylation (RdDM). RdDM is a biological mechanism in which DNA methylation is controlled by non-coding RNA molecules. RdDM-induced DNA methylation is usually linked to transcriptional suppression of the genomic regions targeted by the pathway (Erdmann and Picard, 2020). DNA methylation can change gene expression through direct and indirect mechanisms. Further, when it occurs in promoter region of a gene it inhibits transcription (Wierzbicki et al., 2008; Zhang et al., 2018). The key players of RdDM pathway are Pol IV (RNA Polymerase 4), RDR2 (RNA-dependent RNA polymerase 2), DCL3 (Dicer-like 3), and AGO4 (Argonaute 4) (Dinh et al., 2013). According to the canonical RdDM pathway, which is best explained in Arabidopsis, generation of 24-nt small interfering RNAs (siRNAs) depends on Pol IV. Pol IV transcribes single strand RNA (ssRNA), which is converted to double strand RNA (dsRNA) by RDR2. Then dsRNA is diced by protein DCL3 resulting in siRNAs (siRNAs 24nt). These siRNAs are loaded in AGO4 protein and AGO4 interacts with RISC (RNA-induced silencing complex) and methylates DNA (Matzke and Mosher, 2014). Apart from these proteins, some other proteins also play roles in RdDM pathways for example, SHH1 (SAWADEE HOMEO-DOMAIN HOMOLOG 1), CLSY (CLASSY), NRPD (NUCLEAR RNA POLYMERASE D), DRM (DOMAINS REARRANGED METHYLASE), RDM (RNA- DIRECTED DNA METHYLATION), RRP6L (RRP6-like), DRD (DEFECTIVE IN RNA-DIRECTED DNA METHYLATION), DDM (DECREASE IN DNA METHYLATION), IDN (INVOLVED IN DE NOVO), HEN (HUA ENHANCER). Moreover, SHH, NRPD, NRPE (DNA-DIRECIED RNA polymerase V subunit 1), DRM, RDM1, RDM2, RDM3, RMD4, and RDR have SAWADEE, RNA_pol Rpb2, RNA_pol Rpb1, DNA methylase, RdDM_RDM1, RNA_pol Rpb4, Spt5-NGN, lwr1, and RdRP conserved domain, respectively (Wassenegger and Krczal, 2006; Ream et al., 2009; He et al., 2009a; He et al., 2009b; Henderson et al., 2010; Hirtreiter et al., 2010; Zhang et al., 2013; Sasaki et al., 2014). CLSY, DRD, and DDM have two conserved domains; Helicase_C and SNF2-rel_dom (Smith et al., 2007; Hu et al., 2013). RRP6L also has two domains (DNA_pol_A_exo1 and HRDC) (Lange et al., 2008). IDNs (zf-XS, XS and XH), HEN (Hen1_Lam_C, Methyltransf_31, and dsRBD2), and AGO (n-side PAZs, C-side Piwi, and ArgoMid) each has three conserved domains (Ausin et al., 2009; Huang et al., 2009; Zhao et al., 2014). Further, DCL has six conserved domains (DEAD, Helicase-C, Dicer_dimer, PAZ, Ribonuclease_3, and dsrm) (Margis et al., 2006). In most of the crops, AGO, DCL, and RDR gene families have been identified and characterized.

DNA methylation performs diverse functions in plant growth and development including stress tolerance (Zhang et al., 2018). RdDM-related genes have been well studied in various crops for their critical roles in stress tolerance. For example, RNA Pol V (NRPE), *RDR1/2/6*, *DCL2/4*, *CaDCLs*, and *RDR1* have been reported for their roles against different stresses in different crops i.e., Arabidopsis and pepper (Garcia-Ruiz et al., 2010; López et al., 2011; Qin et al., 2018). Further, *SlAGO4* showed a negative role in salt and drought stress tolerance in tomato (Huang et al., 2016). DNA methylation

may have potential roles in flowering, fruit setting, seed development, and fruit ripening. In Arabidopsis low levels of methylation caused delay in flowering (Jones and Sung, 2014). *AtRRP6L1* and *AtRRP6L2* regulated flowering in Arabidopsis *via* repressing the expression of *FLC* (*Flowering Locus C*) (Shin and Chekanova, 2014). On the other hand, high levels of DNA methylation are important for orange fruit development and ripening (Huang et al., 2019). In the abscission zone of the citrus sinensis fruit, key players of the RdDM pathway are down-regulated (Sabbione et al., 2019). Further, low DNA methylation levels in apples resulted in reduced fruit size (Daccord et al., 2017). In tomato a change in hypermethylation in *CNR* (*COLORLESS NON-RIPENING*) indicated abnormal fruit ripening (Manning et al., 2006). Cheng et al. (2018) has also reported the involvement of RdDM pathway in strawberry fruit ripening. During the ripening process in apples, DNA methylation of the *MdMYB10* promoter affected gene expression and fruit color (El-Sharkawy et al., 2015).

The role of RdDM in seed development and germination has been studied (Xiao et al., 2006; Zhong et al., 2020). Furthermore, the data suggested that DNA methylation can play a critical role in seed dormancy (Kawakatsu et al., 2017). The association of DNA methylation with seed development has been reported in *Arabidopsis thaliana* and *Brassica rapa* (Grover et al., 2018). For example, the involvement of *AtHEN1* in ovule development (Wei et al., 2020). Various phases of tissue have different degrees of DNA methylation. For example, the shoot apical meristem of early peach seedlings exhibited a higher amount of DNA methylation than adults (Bitonti et al., 2002). Most probably, the diverse DNA methylation levels in different tissues at different growth stages are due to different DNA methylation pathways (Bartels et al., 2018) and gene expression levels.

Grapevine (*Vitis vinifera* L.) is among the top fruit crops grown all over the world. The grapes are consumed in different ways including fresh fruit (table), wine, juices, and raisins (dried) (This et al., 2006). RdDM pathway genes have been identified in different plant species including Arabidopsis (Kurihara et al., 2008), *Brassica rapa* (Cao et al., 2016), Soybean (An et al., 2017), Citrus (Mosharaf et al., 2020), Apple (Li et al., 2019), Rice (Ahmad et al., 2019), Maize (Qian et al., 2011), Tomato (Bai et al., 2012), and Ginkgo (Gao et al., 2020). However, in most of the species scientists have reported only three RdDM (DCL, AGO, and RDR) gene families. There is a lack of information about *RdDM* genes in grapevine. The involvement of *RdDM* genes in different mechanisms of plant growth, development, and response to different stresses (biotic and abiotic) and antiviral defense, justifies the need of detailed bioinformatics studies of *RdDM* genes in grapevine. Here, we elaborated comprehensive genome-wide identification of *RdDM* grapevine genes, including phylogenetic analysis, chromosomal positions, intron-exon distribution, motif analysis, evolutionary history, selection pressure, and cis-elements analysis. Various transcriptome analyses under biotic stress, abiotic stress, hormone treatments, and at different stages of growth were performed. The expression of selected genes was further investigated during progressive phases of seed development in seedless and seeded cultivars. Moreover, the integration of all these studies can provide inklings about *VvRdDM* pathway genes function and their involvement in multiple biological processes. Our study provides basic information about *RdDM* grapevine genes and will provide new candidate genes for functional studies.

Materials and methods

Identification of RNA-directed DNA methylation genes in grapes

To identify RdDM pathway genes in grapes, the grape genome was retrieved from Grape Genome CRIBI Biotechnology (http://genomes.cribi.unipd.it/) and Grape Genome Database (http://www.genoscope.cn.fr). Protein sequences involved in the RdDM pathway of the Arabidopsis were downloaded from the TAIR database (http://www.arabidopsis.org/) and were used as a query in Pfam database (http://pfam.xfam.org/search#tabview=tab1) to obtained the Hidden Markov Model (HMM) profile of the RdDM conserved domain. SPDE software was used with the hmmer search function (HMMER 3.0) to obtain candidate genes (Xu et al., 2021). Arabidopsis RdDM pathway protein sequences were used as a query in National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov/), Grape Genome Database, and Grape Genome CRIBI Biotechnology using BLAST-P program (e-value threshold of $1e^{-10}$) (Altschul et al., 1997) to search for sequences of *VvRdDM* homologous genes. The genes obtained from HMM and BLAST-P were combined. To further determine the reliability of RdDM pathways genes in grape, search results of the protein sequences were checked with SMART (http://smart.embl-heidelberg.de), Pfam database and proteins with incomplete domains were removed.

Cis-acting regulatory elements prediction

2 Kb upstream region (Xi et al., 2017) of the *VvRdDM* genes was retrieved from the latest *Vitis vinifera* genome assembly and annotations were downloaded from Phytozomev12 (http://www.phytozome.net) using the SPDE. PlantCARE database was used for conserved cis-elements prediction (Lescot et al., 2002).

Phylogenetic analysis and nomenclature

The full length amino acid sequences of RdDM proteins belonging to *A. thaliana*, *Oryza sativa*, *Solanum lycopersicum*,

and *V. vinifera* were downloaded from TAIR and JGI Data Portal (https://data.jgi.doe.gov/). Multiple alignment of RdDM protein sequences was performed using ClustalW program with default settings. MEGA7.0 software was used for phylogenetic tree construction (Kumar et al., 2018). The phylogenetic trees were generated with the following parameters; Neighbor-Joining (NJ) method, p-distance, complete deletion, and 1000 bootstrap replicates (Ahmad et al., 2020). The candidate grapevine genes were named based on phylogenetic relationships and sequence homologies with corresponding Arabidopsis homologs (Grimplet et al., 2014).

Exon-intron distribution and conserved motifs analysis

The conserved motif distribution patterns of *VvRdDM* genes were assessed using Expectation Maximization for Motif Elicitation (MEME) online software (version 5.4.1) (https://meme-suite.org/meme/tools/meme) (Bailey and Elkan, 1994) with the following parameters: default settings and maximum number of motifs were designated to identify 20 motifs. SMART program and Pfam database were used to annotate the MEME motifs. TBtools were used to visualize the results of conserved motif analysis (Chen et al., 2020). The full genomic sequences and respective coding sequences of RdDM grape genes were obtained from Grape Genome Database and online Gene Structure Display Server 2.0 (http://gsds.gao-lab.org/index.php) was used for exon-intron analysis.

Evolutionary history analysis and estimation of selection pressure

Tandem duplication was determined using the Houlb-described criterion (Holub, 2001). The Plant Genome Duplication Database (http://chibba.agtec.uga.edu/duplication/) was used to retrieve syntenic pairs between Arabidopsis and grapes and among grape genes (Lee et al., 2012). The Circos diagrams were created using TBtools (Chen et al., 2020). The synonymous substitution rate (Ks) and non-synonymous substitution rate (Ka) of duplicated genes were calculated using an online software (http://services.cbu.uib.no/tools/kaks). For estimating selection pressure, the Ka/Ks ratio was calculated (Siltberg and Liberles, 2002).

Plant materials

In this experiment, 2 seeded ('Zuijinxiang' and 'Kyoho') and 2 seedless ('Crimson seedless' and 'Flame seedless') cultivars were used. All the plants were grown in grapevine germplasm orchard of Hebei Agricultural University, Baoding, Hebei, China (38°51′N, 115°29′E) under field conditions, where the average annual temperature is 13.4°C, (Below - 4.3°C in winter and above 26.4°C in summer), the average annual sunshine hours are 2511, and the mean annual precipitation is 498.9 mm. Seed samples were taken from berries using tweezers at 4 different growth stages, 20, 30, 40, and 50 days after full bloom (DAF). All the samples were quickly dipped in liquid nitrogen and preserved at -80°C for further studies. Three biological replicates were obtained at each time point.

Expression analysis of RdDM pathway genes in grapes

Expression analysis of the *VvRdDM* pathway genes were carried out using transcriptome data of NCBI SRA database (https://www.ncbi.nlm.nih.gov/sra/?term=), the corresponding accession IDs are as followings: for biotic stress treatment: downy mildew infection (SRP013835), botrytis cinerea infection (SRP120480), powdery mildew infection (SRP253455); For abiotic stress treatment: cold (SRP202053), heat (SRP091989); For hormone treatment: gibberellic acid (GA_3) (SRP045605), methyl jasmonate (MeJA) and salicylic acid (SA) (SRP378285), abscisic acid (ABA) (SRP098802); For organ development at different stages: seeds (SRP081137), buds (SRP159132), inflorescences (SRP045605), berries (SRP265116). Then the raw transcriptomic data was checked using FastQC (v0.11.9) software, and low-quality reads and adapters were removed using Trimmomatic v0.39. All clean reads were mapped to the *V. vinifera* reference genome (PN40024.v4) by Hisat2 software. Gene expression quantification [Transcripts Per kilobase of exon model per Million mapped reads (TPM)] were estimated using StringTie (v2.2.1) (Pertea et al., 2016). R package DeSeq2 (Love et al., 2014) with the principle adjusted p-value (padj) < 0.05 and the |log2 fold change (log2 FC) | ≥ 1 was used for the identification of significantly differently expressed genes (DEGs). Heat maps were generated using TBtools.

RNA extraction and qRT-expression profiling

The OmniPlant RNA Kit (Dnase I, Comwin Biotech, Beijing, China) was utilized for the extraction of total RNA. For checking the quality and OD value, Agarose gel electrophoresis and nano drop spectrophotometer (Thermo Fisher Scientific, Yokohama, Japan) were used, respectively. Reverse transcription was carried out after gDNA was eliminated using the Prime ScriptRTase (Trans Gen Biotech, Beijing, China). gDNA Eraser (contains Dnase) eliminated genome DNA. Additionally, a qPCR assay on crude RNA was carried out to determine the level of gDNA contamination. Then cDNA was diluted to 200 ng/μl by using distilled water. Primer Premier 7 software was used for designing

specific primers (Supplementary Table S1). Each reaction was carried out in 20 µl volume comprising of 1 µl cDNA,1 µl of each primer (1.0 µM), 10 µl of 2 X Fast Super EvaGreen® qPCR Mastermix (US Everbright Inc., Suzhou, China), and 7 µl sterile distilled H_2O. Every reaction was performed three times. With three biological and three technical replicates, the quantitative real time-PCR (qRT-PCR) was carried out using the SYBR Green (Applied Biosystem) in the LightCycler®96 System (Applied Biosystems, Roche, Shanghai, China). We used the grape *ACTIN* gene (GenBank Accession number NC 012010) as an internal reference, and relative gene expression was computed using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001).

Statistical analysis

SPSS Statistics 22.0 software was used for data analysis. One way analysis of variance (ANOVA) on ranks followed by tukey *post-hoc* analysis was used to evaluate the overall significance of the data (Tukey's test; $p < 0.05$) (Wang et al., 2020). The graphs were created using sigma Plot 14.

Results

VvRdDM genes identification

There were 60 RdDM genes found in grapes. The obtained genes were structurally analyzed and all genes were divided into different gene families according to conserved domains. There were 4, 2, 3, 2, 1, 9, 2, 10, 1, 3, 1, 4, 5, and 13 genes in SHH, CLSY, NRPD, NRPE, DRM, RDM, RRP6L, DRD, DDM, IDN, HEN, DCL, RDR, AGO sub-family, respectively. The protein prediction lengths of SHH, RDM, DRD, DCL, RDR, and AGO sub-families ranged from 248- 697, 174-1034, 729-2266, 1340-1689, 919-1128, and 879-1039 amino acids, respectively. The complete information about *VvRdDM* genes including gene locus ID, chromosomal position, coding sequence length, and protein sequence length are shown in Table 1. All the genes were named as per earlier criteria i.e., based on protein sequence similarity with Arabidopsis RdDM genes.

Phylogenetic and structural analysis

To study the evolutionary relationship of the grape RdDM genes. The phylogenetic trees were constructed among Arabidopsis, grape, tomato, and rice RdDM subfamilies i.e., AGO, DCL, and RDR. According to Figure 1A, AGO proteins can be divided into 4 subgroups including AGO4, ZIPPY/AGO7, AGO1, and MEL/AGO5, and the protein affinities among dicots are closer. Further, AGO4 subgroup was more conserved with respect to proteins number. Except tomato all other have same number of proteins. The same trend of distribution in different subgroups was noted in all families of RdDM genes of grapes except RDM and DRD. Four VvDCLs proteins were divided in four subgroups: DCL1, DCL2, DCL3, and DCL4 (Figure 1D). Likewise, VvRDRs protein were distributed in RDR1, RDR2, RDR6, and RDR3/4/5 (Figure 1B). The VvRDMs and VvDRDs proteins of grapes were grouped into three sub-classes: RDM3, RDM2/4, RDM1 (Figure 1C) and I, II, III (Figure 1E), respectively. The more differences between grapes and Arabidopsis RDM and DRD protein numbers indicate that the RDM and DRD have observed more changes during the evolutionary process.

For the better understanding of gene evolutionary process, the exon-intron structure of the grape RdDM genes was studied. The number of exons ranged from 2 (*VvRDM1*) to 32 (*VvDRD1j*) (Figure 2). As the AGO subfamily is concerned, there were almost the same number of exons within the same subgroup except AGO7 subgroup. This indicates that AGO7 is less conserved as compared to other groups. In addition, the DRD, RDM, RDR, SHH, DCL, NRPD, and IDN sub-family contained 8-32, 4-22, 4-18, 6-10, 21-25, 10-25, and 6-7 exons, respectively. Except, DCL and IDN sub-families, there were differences in the exon-intron distribution in the subfamilies and even within same subgroups, reflecting the diversity emerged during evolution. However, four duplicated gene pairs showed conserved (*VvAGO10a-VvAGO10b*, *VvIDN2a-VvIDN2b*, *VvAGO5a-VvAGO5b*, and *VvSHH1a-VvSHH1b*) exon-intron number.

Most of the subfamilies including AGO, DRD, RDR, SHH, DCL, IDN, RRP6L, and CLSY displayed conserved motif distribution patterns (Figure 3). Further, DRD, DCL, CLSY, and DDM subfamilies have a common structural domain (Helicase_C), but with different position and number. However, the members of RDM sub-family (supergene family) have different domains and protein sequences. Four duplicated gene pairs having same exon-intron distribution (as mentioned above) also showed similar motif distribution patterns.

Promoter analysis of *VvRdDM* genes

The distribution patterns of different types of cis-regulatory elements may provide inklings about gene expressions and functions. A total of 18 cis-elements were identified and can be divided into 4 functional types; (i) Hormone responsive (MeJA, ABA, Auxin, SA, GA), (ii) Light responsive, (iii) Stress responsive (Defense and stress, Anaerobic, Drought, Cold), (iv) Participate in plant development (Endosperm, Meristem, Seed specific, Circadian control).

Except four (*VvNRPD1*, *VvNRPD2a*, *VvDRD1i*, *VvAGO3a*), all other genes have hormone responsive elements including MeJA responsive element (CGTCA-motif, TGACG-motif), the ABA responsive element (ABRE), the Auxin responsive element

TABLE 1 List of *VvRdDM* genes.

Gene Name	Accession No.	VCost.v3 ID	Gene Locus ID	Chr.	Start	End	CDS (bp)	ORF (aa)
VvSHH1a	XP_002283948.1	Vitvi09g00305	GSVIVT01017001001	9	3362926	3368211	744	248
VvSHH1b	XP_002277697.2	Vitvi04g00007	GSVIVT01035290001	4	81666	94722	738	246
VvSHH2a	XP_010644158.1	Vitvi19g00157	GSVIVT01014258001	19	2065687	2074580	1260	419
VvSHH2b	XP_010654302.1	Vitvi08g02188	GSVIVT01034198001	8	14493878	14509826	2082	694
VvCLSY1	XP_019081447.1	Vitvi02g00600	GSVIVT01013277001	2	5723377	5725314	1665	555
VvCLSY2	XP_010658217.1	Vitvi13g00115	GSVIVT01032746001	13	1072521	1084172	4908	1635
VvNRPD1	XP_010661365.1	Vitvi02g00414	GSVIVT01019869001	2	4035665	4081274	4080	1359
VvNRPD2a	XP_002283296.1	Vitvi04g01927	GSVIVT01035963001	4	6537828	6551330	3663	1221
VvNRPD2b	XP_002274051.1	Vitvi12g00313	GSVIVT01020543001	12	4643379	4650715	3582	1193
VvNRPE1	XP_002265533.1	Vitvi13g01420	GSVIVT01013491001	13	1672655	1730512	5673	1891
VvNRPE3B	XP_002285322.1	Vitvi03g00543	GSVIVT01003174001	3	5981436	5989264	621	206
VvDRM1	XP_019079260.1	Vitvi12g02119	GSVIVT01023152001	12	21903254	21909960	1383	460
VvRDM1	XP_019075614.1	Vitvi05g00537	GSVIVT01018017001	5	5505841	5509295	525	174
VvRDM2a	XP_010646234.1	Vitvi10g01687	GSVIVT01003467001	Un	10539130	10543945	720	240
VvRDM2b	XP_002282413.1	Vitvi12g02216	GSVIVT01020782001	12	2049955	2053435	417	139
VvRDM3a	CBI24807.3	Vitvi13g01507	GSVIVT01000315001	Un	2999773	3004686	684	227
VvRDM3b	XP_010658348.1	Vitvi13g00674	GSVIVT01016203001	13	6641768	6659876	2028	676
VvRDM3c	CBI33588.3	Vitvi14g01042	GSVIVT01030928001	14	19211208	19212047	375	125
VvRDM3d	RVW12274.1	Vitvi19g01303	GSVIVT01038034001	19	16557000	16569957	756	252
VvRDM3e	XP_002265283.2	Vitvi18g02832	GSVIVT01010022001	18	13526074	13535704	3105	1034
VvRDM4	XP_010650840.1	Vitvi06g01773	GSVIVT01024608001	6	8458427	8471269	1068	355
VvRRP6L1a	XP_002269553.2	Vitvi08g00120	GSVIVT01029953001	8	2060775	2111425	2706	901
VvRRP6L1b	XP_010657201.1	Vitvi12g00367	GSVIVT01030425001	12	5464697	5472560	2808	936
VvDRD1a	XP_002273814.2	Vitvi03g00003	GSVIVT01024225001	3	24426	40615	2571	876
VvDRD1b	XP_010660172.1	Vitvi14g00743	GSVIVT01023393001	14	12730208	12834845	4434	1477
VvDRD1c	XP_010656983.1	Vitvi12g00289	GSVIVT01020568001	12	4227744	4234100	2625	874
VvDRD1d	CBI37137.3	Vitvi07g00624	GSVIVT01028359001	7	6562398	6570091	3351	1117
VvDRD1e	XP_002264260.1	Vitvi04g01734	GSVIVT01026450001	4	23417241	23454255	3132	1044
VvDRD1f	XP_010653739.1	Vitvi08g01952	GSVIVT01033231001	8	22355168	22372711	4674	1557
VvDRD1g	XP_003631348.1	Vitvi01g01488	GSVIVT01010282001	1	18955322	18986890	2187	729
VvDRD1h	XP_002275100.1	Vitvi06g01262	GSVIVT01037235001	6	17025273	17066889	4620	1540
VvDRD1i	XP_010649878.1	Vitvi05g00453	GSVIVT01017921001	5	4705770	4728522	3333	1110
VvDRD1j	XP_010649796.1	Vitvi05g01863	GSVIVT01017820001	5	3724867	3754247	6801	2266
VvDDM1	XP_010649157.1	Vitvi04g01275	GSVIVT01018979001	4	18137893	18145402	2403	801
VvIDN2a	XP_002278500.1	Vitvi13g01987	GSVIVT01016528001	13	3094114	3102829	1926	642
VvIDN2b	XP_010651012.1	Vitvi06g00477	GSVIVT01024947001	6	5783783	5800541	1881	627
VvIDN2c	XP_002267670.1	Vitvi11g01117	GSVIVT01010892001	11	16213276	16217511	1740	579
VvHEN1	XP_002264328.3	Vitvi10g00954	GSVIVT01021670001	10	9034552	9050667	2796	932
VvDCL1	XP_010661522.1	Vitvi15g00864	GSVIVT01027462001	15	16551973	16563829	4425	1475
VvDCL2	XP_010649214.1	Vitvi04g01202	GSVIVT01019052001	4	17331795	17344656	4023	1340
VvDCL3	XP_010648400.1	Vitvi04g00186	GSVIVT01035494001	4	1758028	1786430	5067	1689
VvDCL4	XP_010656556.1	Vitvi11g00618	GSVIVT01001045001	11	6922239	7007216	4872	1624
VvRDR1a	XP_002284914.1	Vitvi01g00505	GSVIVT01011643001	1	5650814	5659360	2760	919
VvRDR1b	XP_002281315.1	Vitvi01g00503	GSVIVT01011645001	1	5627438	5633661	2832	943
VvRDR2	XP_002280099.1	Vitvi17g00776	GSVIVT01007792001	17	8977900	8986870	3384	1128
VvRDR3	XP_010656269.1	Vitvi11g00273	GSVIVT01015313001	11	2582031	2599309	2790	930
VvRDR6	XP_010648660.1	Vitvi04g00477	GSVIVT01035851001	4	4881183	4886034	3108	1035
VvAGO1	XP_002271225.1	Vitvi17g01218	GSVIVT01029383001	17	15920269	15928175	3117	1039
VvAGO2	XP_010655928.1	Vitvi10g01339	GSVIVT01026268001	10	15190286	15195417	2985	994

(Continued)

TABLE 1 Continued

Gene Name	Accession No.	VCost.v3 ID	Gene Locus ID	Chr.	Start	End	CDS (bp)	ORF (aa)
VvAGO3a	XP_003633060.1	Vitvi10g01346	GSVIVT01026261001	10	15100708	15106946	2952	983
VvAGO3b	XP_002274149.1	Vitvi10g01342	GSVIVT01026264001	10	15146900	15151225	2937	979
VvAGO4a	XP_002275928.1	Vitvi06g01020	GSVIVT01037488001	6	12830288	12839311	2742	913
VvAGO4b	XP_010653696.1	Vitvi08g00884	GSVIVT01025868001	8	10978235	10995752	2640	879
VvAGO5a	XP_010651834.1	Vitvi06g01378	GSVIVT01031430001	6	18665350	18674795	3099	1032
VvAGO5b	XP_010654011.1	Vitvi08g01559	GSVIVT01033726001	8	18308666	18314201	2556	852
VvAGO6a	XP_010657243.1	Vitvi12g00448	GSVIVT01030512001	12	6352396	6365926	2703	900
VvAGO6b	XP_019079568.1	Vitvi13g01085	GSVIVT01001941001	13	15074590	15079381	2646	882
VvAGO7	XP_002267746.1	Vitvi01g01134	GSVIVT01012490001	1	13829611	13833472	2670	889
VvAGO10a	XP_002279408.1	Vitvi05g00574	GSVIVT01018054001	5	5849436	5858977	2862	954
VvAGO10b	XP_010656388.1	Vitvi11g00408	GSVIVT01015464001	11	3938426	3949275	2718	906

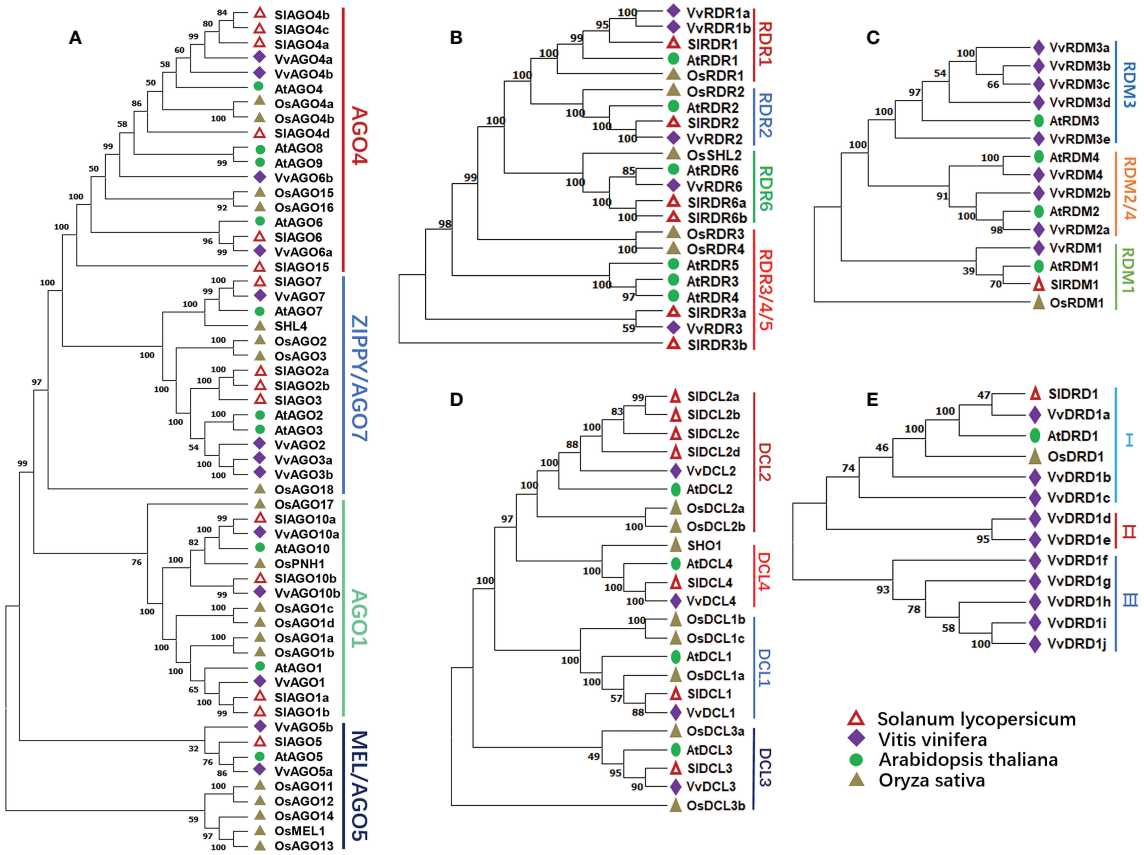


FIGURE 1
Phylogenetic analysis of AGO, RDR, RDM, DCL, and DRD proteins from grape, Arabidopsis, tomato, and rice. Hollow triangle, diamonds, circles, and triangle represent tomato, grapes, Arabidopsis, rice proteins, respectively. Numbers near the branches denote bootstrap values. **(A)** Phylogenetic tree of AGO proteins. AGO4, AGO7, AGO1, and AGO5 represent different subgroups. **(B)** Phylogenetic tree of RDR proteins. RDR1, 2, 3, 4, and 5 represent different subgroups. **(C)** Phylogenetic tree of RDM proteins. RDM1, 2, 3, and 4 represent different subgroups. **(D)** Phylogenetic tree of DCL proteins. DCL1, 2, 3, and 4 represent different subgroups. **(E)** Phylogenetic tree of DRD proteins. I, II, and III represent different subgroups.

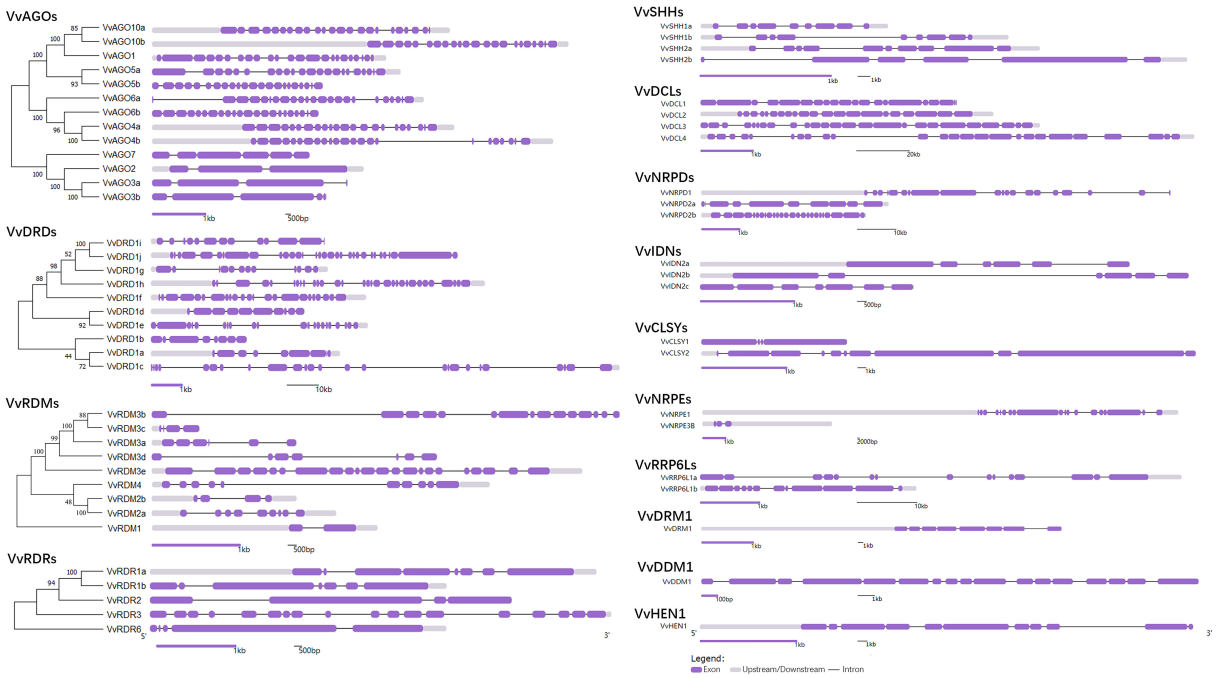


FIGURE 2
Phylogenetic tree and exon-intron distribution of *VvRdDM* genes. Purple boxes and black lines denote exons and intron, respectively.

(TGA-element, AuxRR-core), the SA responsive element (TCA-element) and GA responsive element (P-box,TATC-box) (Figure 4A). However, the ABRE were more abundant accounting up to 31% share in hormone responsive elements (Figure 4B).

Stress response cis-elements were present in most of the genes i.e., ARE (involvement in anaerobic induction), LTR (Low temperature-responsive elements), and TC-rich repeats (defense and stress response). Further, the presence of plant growth and development (CAT-box, circadian, GCN4-motif, and RY-elements) related elements suggests the involvement of *VvRdDM* genes in grapevine growth and development.

Expansion patterns and chromosomal arrangements of *VvRdDM* genes

VvRdDM genes were mapped on grape chromosomes by utilizing the available grape genome annotation information. All genes were unequally present on 19 chromosomes and there was no gene on chromosome 16. Chromosome 4 contained the maximum number of genes (7) (Figure 5). Further, total 8 genes undergone segmental duplication in the form of 4 pairs (*VvAGO10b-VvAGO10a*, *VvIDN2a-VvIDN2b*, *VvAGO5a-VvAGO5b*, and *VvSHH1a-VvSHH1b*) and two events of tandem duplication (*VvRDR1a-VvRDR1b* and *VvAGO2-VvAGO3a-VvAGO3b*) were observed among five genes. Interesting, AGO family genes have undergone both duplication events and more than 50% genes (7 out of 13) are the result of duplication events. These results suggest that segmental and tandem duplication both have a role in the expansion of some *VvRdDM* sub-families. Further, Ka/Ks ratio was calculated for the estimation of selection pressure. According to the results all the duplicated genes have purifying mode of selection i.e., Ka/KS < 1 (Supplementary Table S2).

Evolutionary relationships between arabidopsis and grapes

To get further information about evolution of *VvRdDM* genes, the syntenic relationships between grape and Arabidopsis RdDM genes were found. There were 19 syntenic blocks among 17 *VvRdDM* and 19 *AtRdDM* genes (Figure 6; Supplementary Table S3). In each syntenic pair, both members belonged to the same subgroup. These syntenic relationships among grape and Arabidopsis suggest that these homologous genes have common ancestors before the speciation. Interestingly, *VvCLSY2* paired with three different Arabidopsis genes (*VvCLSY2-AT5G05410/AT3G11020/AT2G40340*). Further, *AT5G59390* also paired with more than one grapevine genes (*AT5G59390-VvIDN2a/*

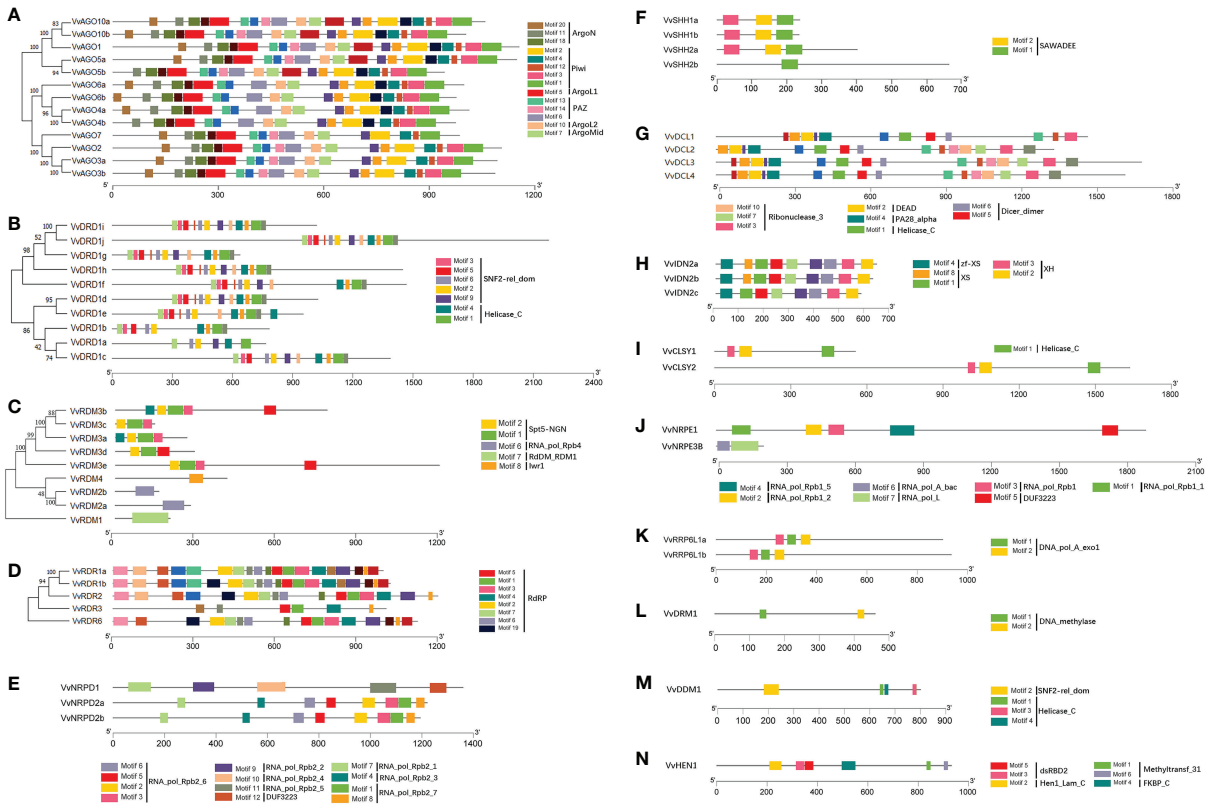


FIGURE 3
Conserved motifs analysis of *VvRdDM* genes. Different color boxes represent different motifs The letters **(A–N)** represent the different subfamilies.

VvIDN2b). These results suggest that *VvCLSY2* and *AT5G59390* have gained more changes during the evolutionary process.

Transcriptomic analysis of *VvRdDM* genes under biotic stress, abiotic stress and hormone treatment

To comprehensively explore potential functions of *VvRdDM* genes, expression patterns were investigated with a variety of transcriptome data. *VvRdDM* genes showed significantly differential expression against botrytis cinerea and downy mildew infection. For example, following *B. cinerea* infection, five genes (*VvRDR6*, *VvAGO4a*, *VvDCL3*, *VvRDM3b*, *VvNRPD1*) showed significant down-regulation and one gene (*VvDRM1*) displayed significant up-regulation expression at berry ripened stage. However, some of the *VvRdDM* genes (*VvRDR1a*, *VvAGO1*, *VvAGO4a*, *VvAGO5a*, etc) showed time specific expression under botrytis cinerea treatment i.e., higher expression level at berry hard green stage than ripened stage. Further, expressions of *VvDCL2*, *VvAGO3a*, and *VvAGO3b* were significantly down-regulated with downy mildew infection. However, none of the *VvRdDM* genes showed significantly different expression after infection with powdery mildew (Figure 7). After heat stress (35°C, 40°C, 45°C) treatments the number of significantly expressed genes was highest (13) at the 40°C. Interestingly, *VvAGO3a* showed a significant up-regulation against all temperature treatments.

Moreover, following cold treatment at 4°C, *VvIDN2c* and *VvAGO10b* were significantly up-regulated, and it is worth noting that the significantly down-regulated gene *VvDCL2* also had a high change in expression (108.58 to 27.81). *VvRdDM* genes showed better response to heat stress (Figure 7) i.e., *VvNRPD2b*, *VvRDM2b*, *VvIDN2a*, *VvAGO1*, and *VvAGO4a* showed decrease in expression with rise in temperature.

Under hormones treatment (Figure 8), most *VvRdDM* genes showed significantly different expression at 24 h post treatment. For example, *VvDDM1* and *VvAGO6a* were significantly up-regulated and *VvDCL2* was significantly down-regulated after GA_3 treatment. Following MeJA treatment *VvRDR3* and *VvAGO7* were significantly up-regulated and *VvRDR1b* was significantly down-regulated. After ABA treatment some of the genes were significantly down-regulated i.e., *VvRDR1b*, *VvRDR1a*, *VvDCL2* and *VvAGO5a*.

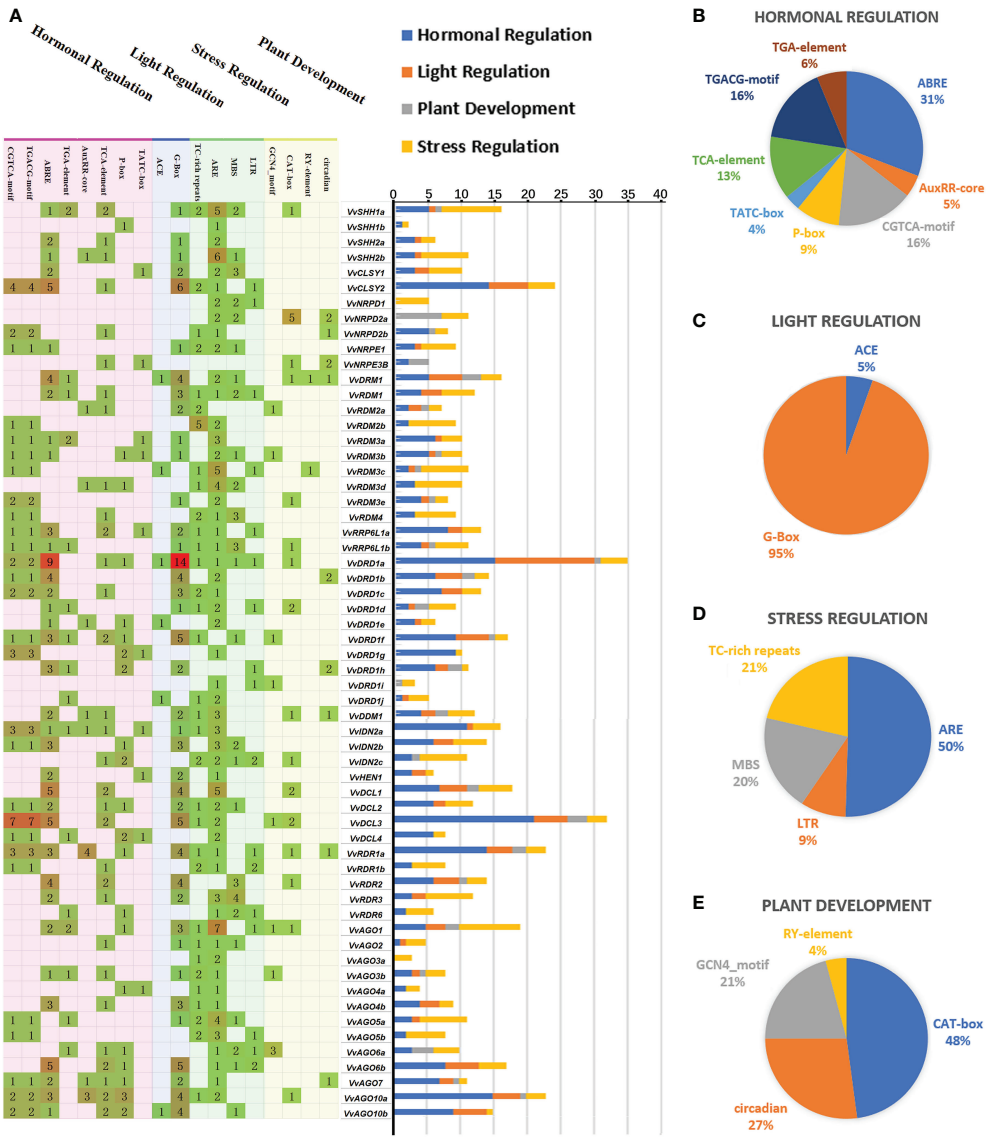


FIGURE 4
Cis-acting element analysis of *VvRdDM* genes. **(A)** The numbers and type of cis-elements in the VvRdDM promoter regions. The cis-elements belonging to different categorizes are denoted with different colors. **(B)** Relative percentage of hormonal regulation cis-elements. **(C)** Relative percentage of light regulating cis-elements. **(D)** Relative percentage of stress regulating cis-elements. **(E)** Relative percentage of plant development related.

Expression profiling of *VvRdDM* genes in various organs and developmental stage

Transcriptomic study of RdDM grapevine genes was undertaken to acquire some hints regarding the probable roles of RdDM grapevine genes in different organs (buds, inflorescences, berries, seeds). The results showed (Figure 9) that the *VvNRPD2b*, *VvNRPE3B*, *VvRDM2b*, *VvRDM3e*, *VvRRP6L1a*, *VvDRD1f*, *VvDRD1g*, *VvAGO1*, and *VvAGO4a* gene showed a broad spectrum of high expression in different organs at various stages, especially *VvRDM2b* and *VvAGO1*. During the process of bud development, *VvRdDM* genes (*VvNRPD2b*, *VvRDM2b*, *VvRRP6L1a*, *VvDCL1*, etc.) initially showed an increase in expression followed by a subsequent decrease. During April most of the genes showed a significant decrease in expression. However, some of the genes like *VvRDM2a*, *VvRDR1a*, and *VvAGO4a* showed a trend of initial decrease followed by increase in expression. Further, some genes (*VvRDM2b*, *VvDCL2*, *VvAGO5a*, etc.) also showed significant changes at the berry ripening stage, and this trend (change of

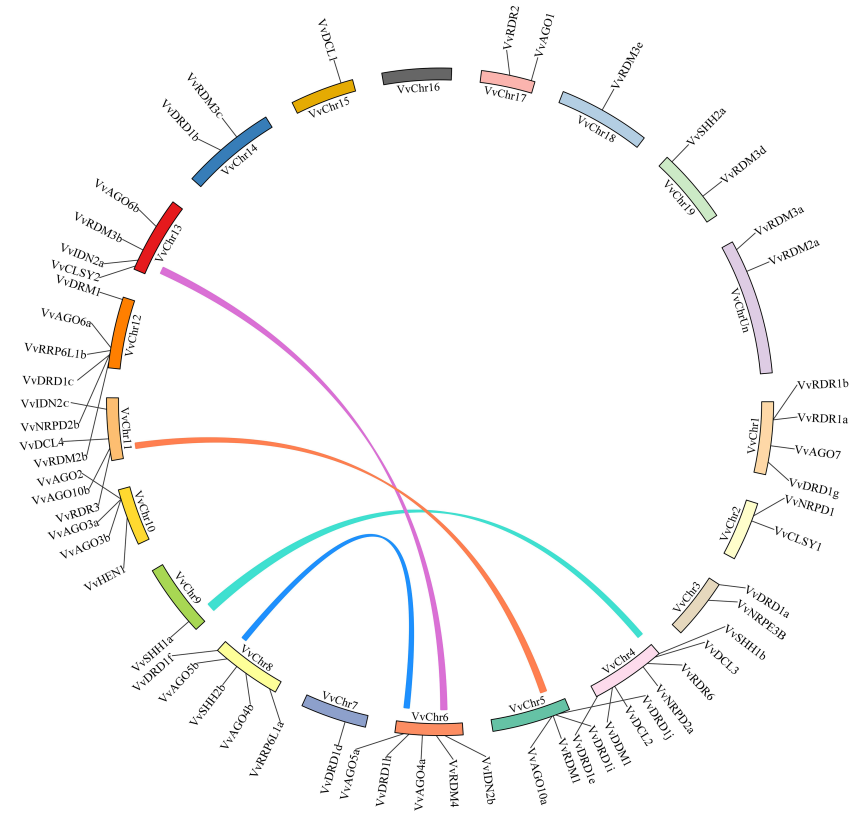


FIGURE 5
Chromosomal distribution and Synteny analysis. The corresponding genes on two chromosomes were regarded as segmental duplications and colored bars connecting two chromosomal regions indicate syntenic regions. Chr: chromosomes.

expression) was also linked with the growth and development of grape organs.

The transcriptomic expressions of genes during seed development, are presented in the form of a ratio between seedless and seeded (Seedless/Seeded) cultivars at specific time points (initial stage, stage with the highest weight, and stage with the lowest weight during seedless progeny seed development) (Wang et al., 2016) (Figure 10A). The expression of *VvRDR1a*, *VvRDR1b*, *VvRDR3*, *VvRDR6*, and *VvRRP6L1b* was significantly higher in seedless progenies than in seeded progenies. Surprisingly, no gene showed significantly higher expression during all stages of seed development in seeded cultivar as compared to seedless. However, *VvAGO5b*, *VvAGO10a*, *VvCLSY1*, and *VvDRD1a* displayed significantly higher expression in seeded progeny only at one or at two times points. Overall, transcriptomic analysis suggests that some of the *VvRdDM* genes might have some roles in seed development or ovule abortion.

Based on transcriptomic analysis, 15 (*VvSHH1a*, *VvCLSY2*, *VvNRDP2b*, *VvRDM4*, *VvRRP6L1b*, *VvDRD1a*, *VvDRD1e*, *VvIDN2a*, *VvHEN1*, *VvDCL4*, *VvRDR1a*, *VvRDR6*, *VvAGO3a*, *VvAGO3b*, and *VvAGO5a*) genes were selected for further qRT-PCR analysis. The qRT-PCR analysis of selected genes was performed to check whether the trend of expression (differential) is common in different cultivars of grapes or it is variety specific. As shown in Figure 10B, *VvSHH1a*, *VvDRD1a*, and *VvIDN2a*, were significantly highly expressed during different stages of seed development in seeded progenies. Further all three genes showed relatively higher change ratio between seedless and seeded cultivars at 40 and 50 DAF as compared to earlier stages. Based on the findings, we postulate that these genes might have a significant role in seed development. On other hand, few genes were highly expressed in specific stages of seed development in seedless cultivars, for example, *VvRRP6L1b*, *VvDCL4*, *VvRDR1a*, and *VvRDR6*. However, *VvRDR1a* and *VvRDR6* showed significant higher expression during all phases of seed development in seedless cultivars than seeded ones. These findings imply that these genes may contribute to ovule abortion. The outcomes of qRT-PCR and transcriptome were generally in agreement.

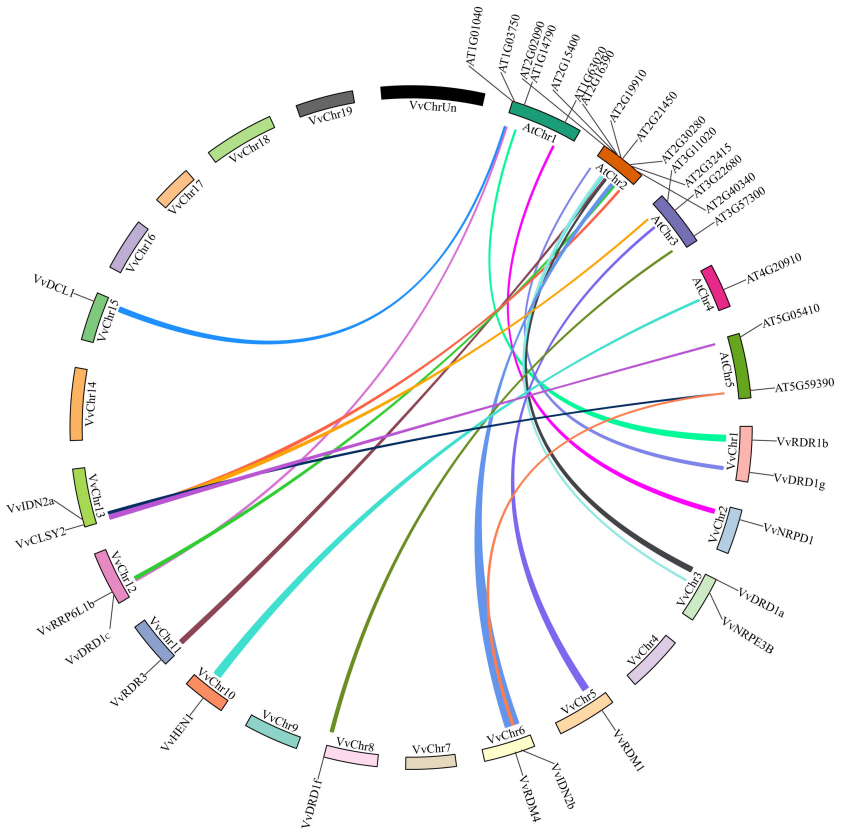


FIGURE 6
Synteny analysis of *VvRdDM* genes between grapes and Arabidopsis. Relative positions were estimated according to the grape and Arabidopsis chromosomes; colored lines denote syntenic regions.

Discussion

In plants, RdDM pathway genes control the epigenetic states *via De novo* methylation and play important roles in multiple biological processes (Matzke and Mosher, 2014; Zhao and Chen, 2014; Matzke et al., 2015). Due to recent advancements in whole-genome mapping and next-generation sequencing, plant breeders are paying more attention in epigenetic research and its use in breeding programmes. The RdDM pathway genes have been characterized in many plants i.e., Arabidopsis, Citrus, Tomato, etc (Kurihara et al., 2008; Qian et al., 2011; Bai et al., 2012; Cao et al., 2016; An et al., 2017; Ahmad et al., 2019; Li et al., 2019; Gao et al., 2020; Mosharaf et al., 2020). However, there have been limited reports for RdDM pathway genes in grapes. Previously, Zhao et al. (2014) identified only three important RdDM pathway-related gene families in grapes i.e., AGO, DCL, and RDR. Here by using the latest high-quality grape reference genome data, we identified 60 RdDM grapevine genes belonging to 14 different subfamilies. Moreover, integrated bioinformatics, transcriptomics, and expression profiling to identify their potential functional roles.

Proteins sequences of one monocot (rice) and three dicots (Arabidopsis, grapes, and tomato) species were used to study the phylogenetic relationships of different RdDM families. Based on phylogenetic studies, VvAGO, VvDCL, and VvRDR proteins were more closely related with SlAGO, SlDCL, and SlRDR proteins, respectively. However, VvRDM proteins were closely related to AtRDM proteins. These results suggest that AGO, DCL, and RDR protein families of grapes and tomato have similarities in evolution and ancestral history. Likewise, RDM proteins of grapes and Arabidopsis might have common ancestors. The pattern of exon-intron distribution within a gene family can provide inklings about evolutionary history. Some of the *VvRdDM* genes families (AGO, DCL, and IDN) showed similar exon-intron numbers within the same sub-families, proposing that these subfamilies are functionally conserved. Other families showed differences in exon-intron number and distribution patterns even within the subfamilies. These results suggest that during the evolutionary process these subfamilies have either gained or lost exon/introns. Changes in exon number during the evolutionary process play important functions in the evolution of new gene families (Xu et al., 2012).

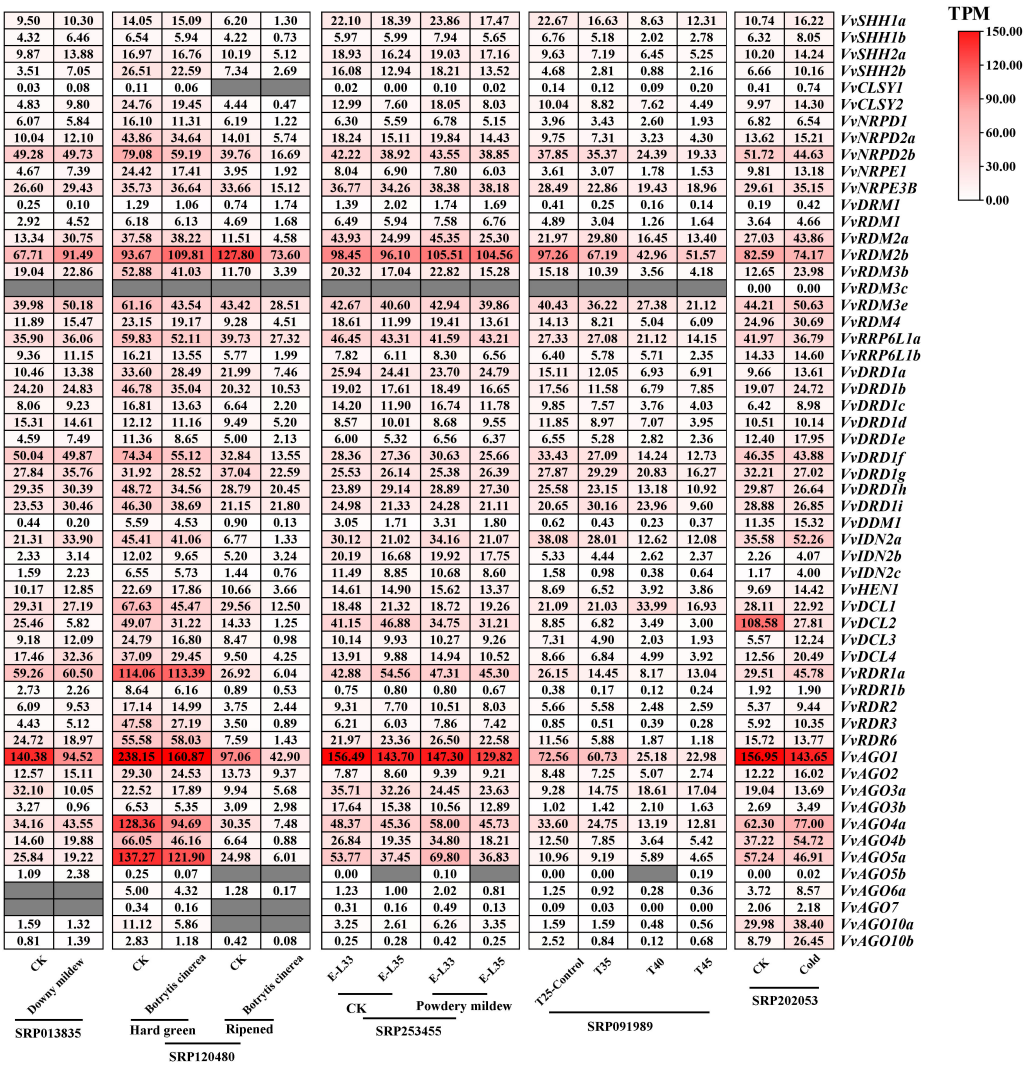


FIGURE 7
Transcriptomic expression of *VvRdDM* genes following biotic and abiotic stresses. Transcriptome data were downloaded from the NCBI SRA datasets (SRP013835, SRP120480, SRP253455, SRP091989 and SRP202053). For transcriptomic analysis plant samples following different stress were collected; For downy mildew infected leaves; Botrytis cinerea infected fruits at 2 stages (hard green and ripened stage); Powdery mildew treated fruits at 2 growth stages (EL33 and EL35). For high temperature stress, plants grown at 25°C were taken as control, and 35°C, 40°C, and 45°C were the treatments. Low temperature treatment is 4°C, plants grown at 25°C were taken as control. The white and red color scale denotes low and high expression levels, respectively. No gene expression is denoted with grey color. Numbers denote TPM of *VvRdDMs*.

Further, the newly evolved genes are initially redundant and gain new functions with the passage of time (Dias et al., 2003). Like intron-exon arrangement, most of the families (DRD, DCL, CLSY, and DDM) showed similar motif distribution patterns within the subfamilies suggesting these subfamilies are conserved during the evolutionary process. More interestingly, DRD, DCL, CLSY, and DDM all have a common domain (Helicase_C), which consists of ATP-dependent proteins which participate in epigenetic mechanisms *via* chromatin remodeling by using the energy provided by ATP (Flaus et al., 2006). This indicates this domain has some key roles in the RdDM pathways. However, RDM members showed differences in motif arrangements.

Gene duplications and exon-intron patterns have played a key role in the expansion of gene families during the evolutionary process (Cannon et al., 2004). In case of duplicated gene pairs previous trend was observed i.e., members of four duplicated gene pairs (*VvAGO10a-VvAGO10b*, *VvIDN2a-VvIDN2b*, *VvAGO5a-VvAGO5b*, and *VvSHH1a-VvSHH1b*) showed a similar trend of exon-intron and motif distribution patterns. Moreover, some duplicated gene pairs (*VvAGO10a-VvAGO10b*, *VvAGO3a-VvAGO3b*) showed

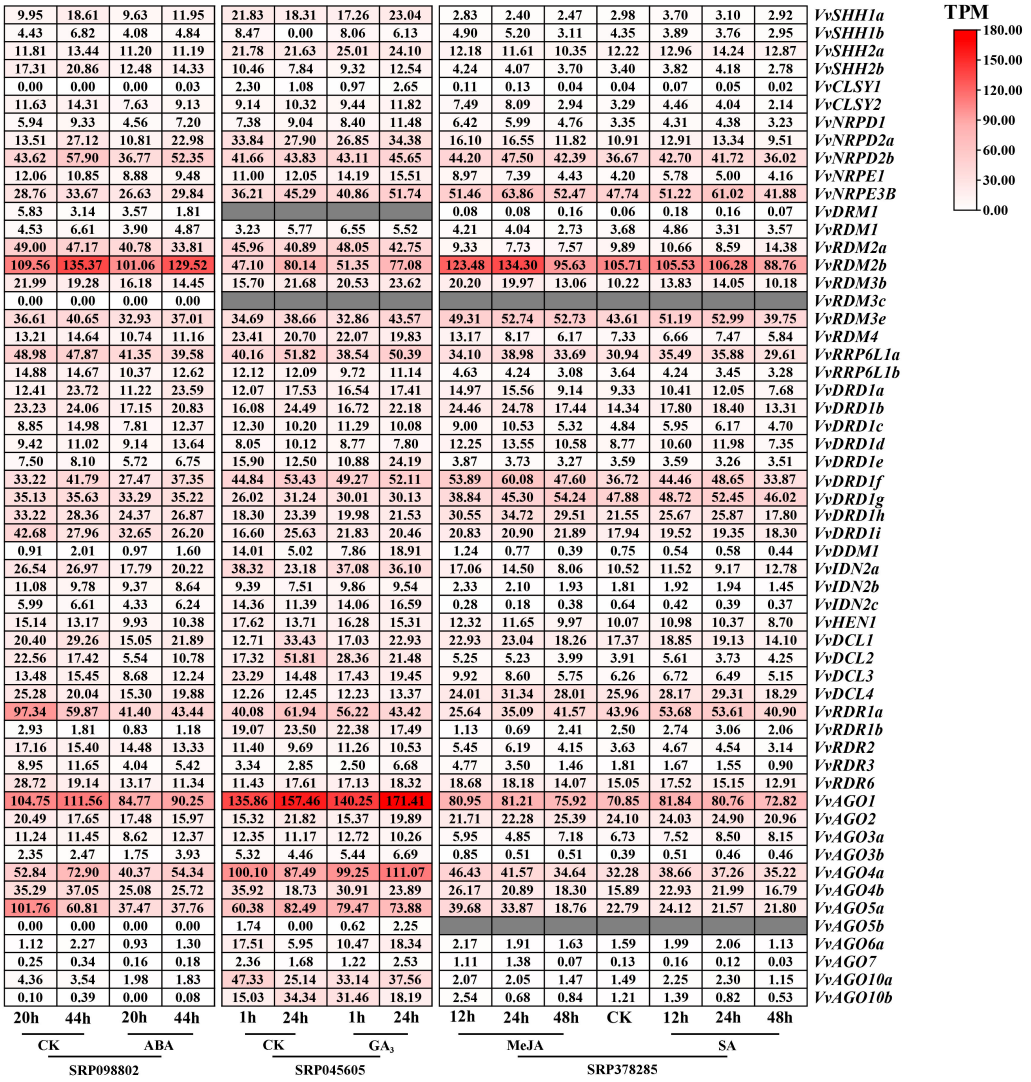
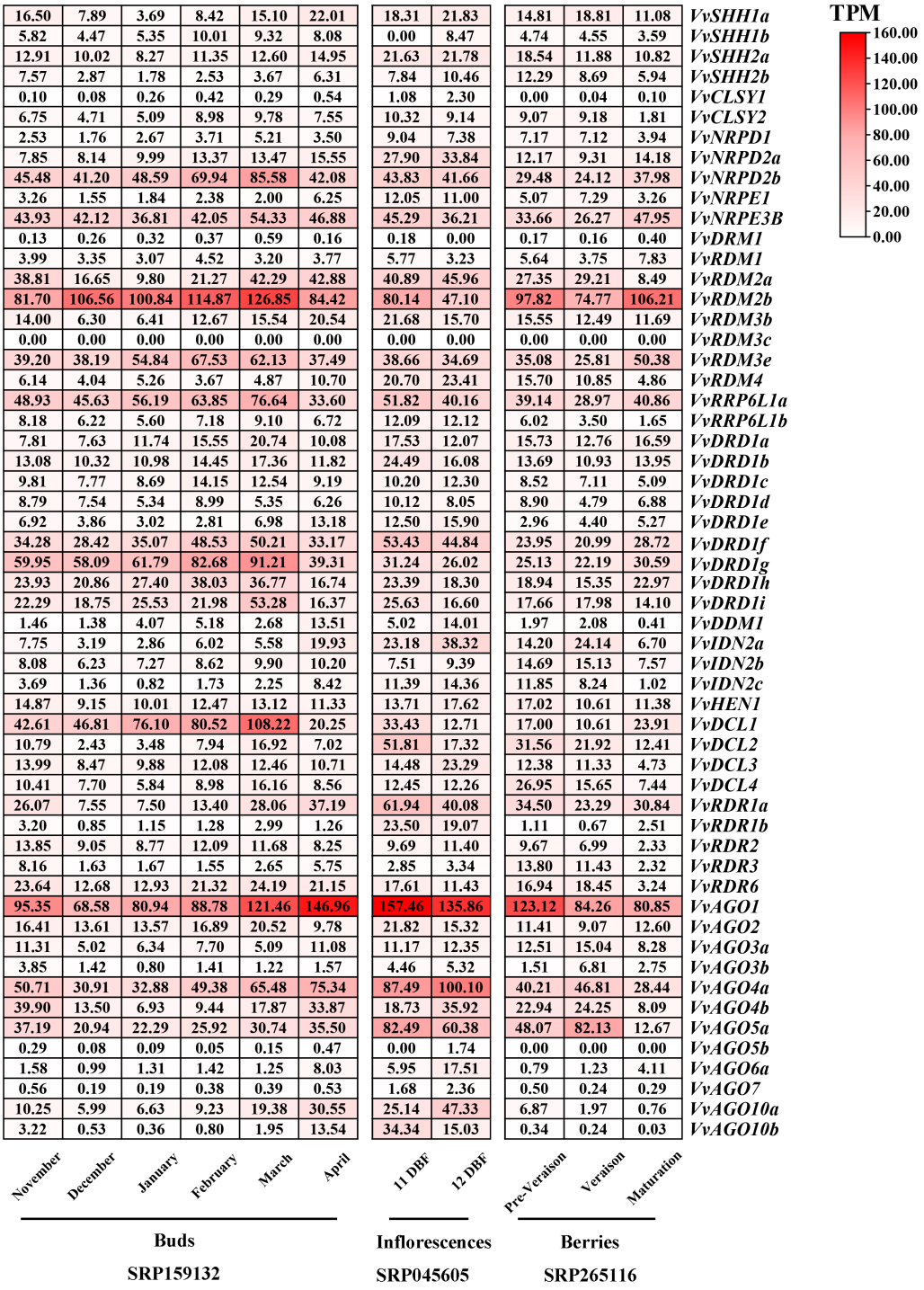


FIGURE 8
Transcriptomic expression of *VvRdDM* genes following different hormone treatments. Transcriptome data were retrieved from the NCBI SRA datasets (SRP09880, SRP045605 and SRP378285). The white and red color scale denotes low and high expression levels, respectively. No gene expression is denoted with grey color. Numbers denote TPM of *VvRdDMs*.

similar expression patterns under high temperature treatment (Figure 11). Only thirteen genes (8 segmental and 5 tandem) belonging to 4 families (AGO, IDN, SHH, and RDR) are the result of duplication events. All duplicated gene pairs have a purifying mode of selection. These results suggest that these four families are diversified and all others are conserved. Further, the AGO family is more diversified because 50% of members of this family have undergone duplication events. Overall, in the expansion of *VvRdDM* gene families, we noticed the previous trend that segmental duplication occurs more often than tandem duplication (Cannon et al., 2004). Synteny analysis can provide inklings about the gene functions. We identified 19 syntenic pairs between grape and Arabidopsis RdDM genes. These findings propose that these homologous genes have common ancestors before the speciation. However, relying solely on syntenic linkages to explain evolutionary relationships is difficult. Interestingly, two genes *VvCLSY2* and *AT5G59390* paired with more than one gene, implying that these genes have gained more changes during the evolutionary process. Based on the stated functions of orthologous genes in one species (Arabidopsis) functions can be predicted in other species (grapes).

According to promoter analysis, most of the *VvRdDM* genes have hormone-responsive cis-elements. However, ABA response element (ABRE) was a more abundant hormone-responsive cis-element, and genes *VvDCL2*, *VvRDR1a* (contained ABRE



16.50	7.89	3.69	8.42	15.10	22.01	18.31	21.83	14.81	18.81	11.08	*VvSHH1a*
5.82	4.47	5.35	10.01	9.32	8.08	0.00	8.47	4.74	4.55	3.59	*VvSHH1b*
12.91	10.02	8.27	11.35	12.60	14.95	21.63	21.78	18.54	11.88	10.82	*VvSHH2a*
7.57	2.87	1.78	2.53	3.67	6.31	7.84	10.46	12.29	8.69	5.94	*VvSHH2b*
0.10	0.08	0.26	0.42	0.29	0.54	1.08	2.30	0.00	0.04	0.10	*VvCLSY1*
6.75	4.71	5.09	8.98	9.78	7.55	10.32	9.14	9.07	9.18	1.81	*VvCLSY2*
2.53	1.76	2.67	3.71	5.21	3.50	9.04	7.38	7.17	7.12	3.94	*VvNRPD1*
7.85	8.14	9.99	13.37	13.47	15.55	27.90	33.84	12.17	9.31	14.18	*VvNRPD2a*
45.48	41.20	48.59	69.94	85.58	42.08	43.83	41.66	29.48	24.12	37.98	*VvNRPD2b*
3.26	1.55	1.84	2.38	2.00	6.25	12.05	11.00	5.07	7.29	3.26	*VvNRPE1*
43.93	42.12	36.81	42.05	54.33	46.88	45.29	36.21	33.66	26.27	47.95	*VvNRPE3B*
0.13	0.26	0.32	0.37	0.59	0.16	0.18	0.00	0.17	0.16	0.40	*VvDRM1*
3.99	3.35	3.07	4.52	3.20	3.77	5.77	3.23	5.64	3.75	7.83	*VvRDM1*
38.81	16.65	9.80	21.27	42.29	42.88	40.89	45.96	27.35	29.21	8.49	*VvRDM2a*
81.70	106.56	100.84	114.87	126.85	84.42	80.14	47.10	97.82	74.77	106.21	*VvRDM2b*
14.00	6.30	6.41	12.67	15.54	20.54	21.68	15.70	15.55	12.49	11.69	*VvRDM3b*
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	*VvRDM3c*
39.20	38.19	54.84	67.53	62.13	37.49	38.66	34.69	35.08	25.81	50.38	*VvRDM3e*
6.14	4.04	5.26	3.67	4.87	10.70	20.70	23.41	15.70	10.85	4.86	*VvRDM4*
48.93	45.63	56.19	63.85	76.64	33.60	51.82	40.16	39.14	28.97	40.86	*VvRRP6L1a*
8.18	6.22	5.60	7.18	9.10	6.72	12.09	12.12	6.02	3.50	1.65	*VvRRP6L1b*
7.81	7.63	11.74	15.55	20.74	10.08	17.53	12.07	15.73	12.76	16.59	*VvDRD1a*
13.08	10.32	10.98	14.45	17.36	11.82	24.49	16.08	13.69	10.93	13.95	*VvDRD1b*
9.81	7.77	8.69	14.15	12.54	9.19	10.20	12.30	8.52	7.11	5.09	*VvDRD1c*
8.79	7.54	5.34	8.99	5.35	6.26	10.12	8.05	8.90	4.79	6.88	*VvDRD1d*
6.92	3.86	3.02	2.81	6.98	13.18	12.50	15.90	2.96	4.40	5.27	*VvDRD1e*
34.28	28.42	35.07	48.53	50.21	33.17	53.43	44.84	23.95	20.99	28.72	*VvDRD1f*
59.95	58.09	61.79	82.68	91.21	39.31	31.24	26.02	25.13	22.19	30.59	*VvDRD1g*
23.93	20.86	27.40	38.03	36.77	16.74	23.39	18.30	18.94	15.35	22.97	*VvDRD1h*
22.29	18.75	25.53	21.98	53.28	16.37	25.63	16.60	17.66	17.98	14.10	*VvDRD1i*
1.46	1.38	4.07	5.18	2.68	13.51	5.02	14.01	1.97	2.08	0.41	*VvDDM1*
7.75	3.19	2.86	6.02	5.58	19.93	23.18	38.32	14.20	24.14	6.70	*VvIDN2a*
8.08	6.23	7.27	8.62	9.90	10.20	7.51	9.39	14.69	15.13	7.57	*VvIDN2b*
3.69	1.36	0.82	1.73	2.25	8.42	11.39	14.36	11.85	8.24	1.02	*VvIDN2c*
14.87	9.15	10.01	12.47	13.12	11.33	13.71	17.62	17.02	10.61	11.38	*VvHEN1*
42.61	46.81	76.10	80.52	108.22	20.25	33.43	12.71	17.00	10.61	23.91	*VvDCL1*
10.79	2.43	3.48	7.94	16.92	7.02	51.81	17.32	31.56	21.92	12.41	*VvDCL2*
13.99	8.47	9.88	12.08	12.46	10.71	14.48	23.29	12.38	11.33	4.73	*VvDCL3*
10.41	7.70	5.84	8.98	16.16	8.56	12.45	12.26	26.95	15.65	7.44	*VvDCL4*
26.07	7.55	7.50	13.40	28.06	37.19	61.94	40.08	34.50	23.29	30.84	*VvRDR1a*
3.20	0.85	1.15	1.28	2.99	1.26	23.50	19.07	1.11	0.67	2.51	*VvRDR1b*
13.85	9.05	8.77	12.09	11.68	8.25	9.69	11.40	9.67	6.99	2.33	*VvRDR2*
8.16	1.63	1.67	1.55	2.65	5.75	2.85	3.34	13.80	11.43	2.32	*VvRDR3*
23.64	12.68	12.93	21.32	24.19	21.15	17.61	11.43	16.94	18.45	3.24	*VvRDR6*
95.35	68.58	80.94	88.78	121.46	146.96	157.46	135.86	123.12	84.26	80.85	*VvAGO1*
16.41	13.61	13.57	16.89	20.52	9.78	21.82	15.32	11.41	9.07	12.60	*VvAGO2*
11.31	5.02	6.34	7.70	5.09	11.08	11.17	12.35	12.51	15.04	8.28	*VvAGO3a*
3.85	1.42	0.80	1.41	1.22	1.57	4.46	5.32	1.51	6.81	2.75	*VvAGO3b*
50.71	30.91	32.88	49.38	65.48	75.34	87.49	100.10	40.21	46.81	28.44	*VvAGO4a*
39.90	13.50	6.93	9.44	17.87	33.87	18.73	35.92	22.94	24.25	8.09	*VvAGO4b*
37.19	20.94	22.29	25.92	30.74	35.50	82.49	60.38	48.07	82.13	12.67	*VvAGO5a*
0.29	0.08	0.09	0.05	0.15	0.47	0.00	1.74	0.00	0.00	0.00	*VvAGO5b*
1.58	0.99	1.31	1.42	1.25	8.03	5.95	17.51	0.79	1.23	4.11	*VvAGO6a*
0.56	0.19	0.19	0.38	0.39	0.53	1.68	2.36	0.50	0.24	0.29	*VvAGO7*
10.25	5.99	6.63	9.23	19.38	30.55	25.14	47.33	6.87	1.97	0.76	*VvAGO10a*
3.22	0.53	0.36	0.80	1.95	13.54	34.34	15.03	0.34	0.24	0.03	*VvAGO10b*

FIGURE 9
Transcriptomic expression of *VvRdDM* genes in different organs at various periods of time. The bud samples were collected at every month from November to April. The inflorescence samples were taken at 11 and 12 days before flowering (DBF). The fruit samples were collected at three developmental stages (pre-veraison, veraison, maturation). Transcriptome data were downloaded from the NCBI SRA datasets (SRP159132, SRP045605 and SRP265116). The white and red color scale denotes low and high expression levels, respectively. No gene expression is denoted with grey color. Numbers denote TPM of *VvRdDMs*.

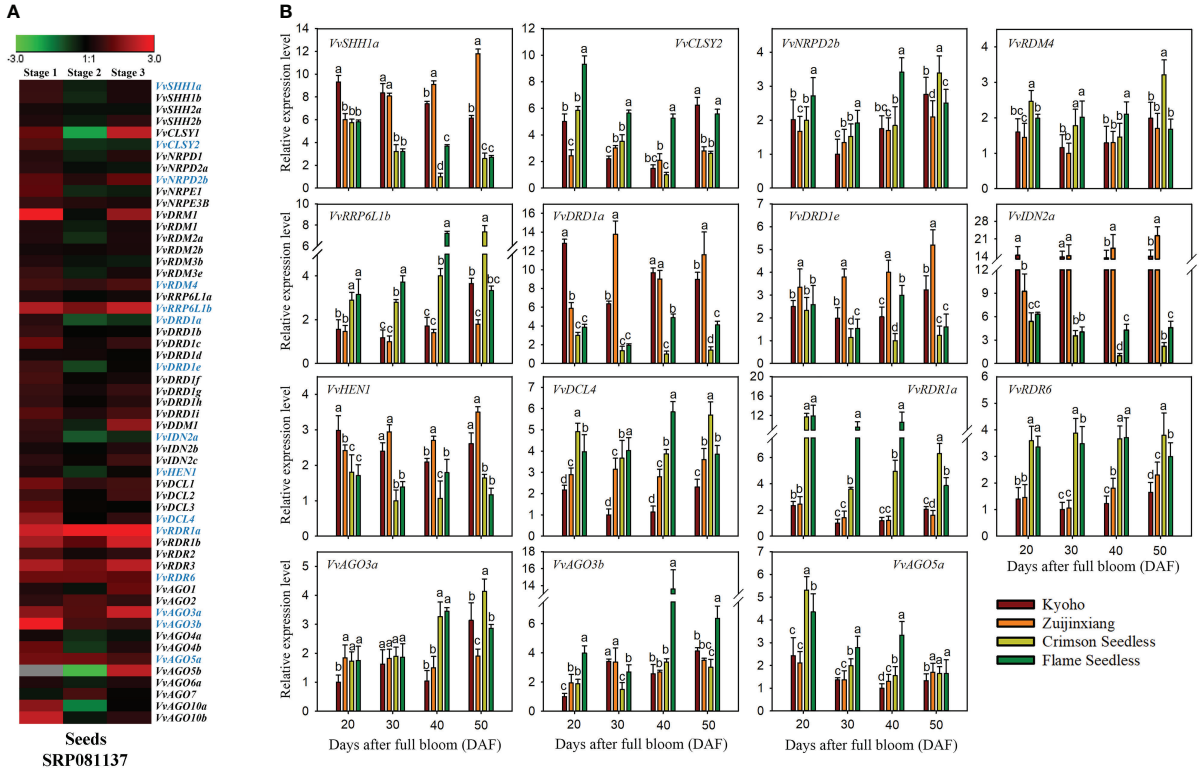


FIGURE 10
Transcriptomic and expression analysis of grapevine *VvRdDM* genes during seed development. **(A)** Heat map of grapevine *VvRdDM* genes in seeded and seedless progenies during seed development, transcriptome data were downloaded from the NCBI SRA datasets (SRP081137). Data were normalized by $\log_2$ the ratio of TPM (seedless/seeded) of each gene during the all three stages. The red and green boxes indicate the relatively high and low expression levels of each gene, respectively. **(B)** qRT-PCR analysis of *VvRdDM* genes in four cultivars during seed development. Different cultivars are denoted with different colors. Different letters denote statistically significant differences (ANOVA with a Tukey *post-hoc* analysis, 5% level). The mean ± S.D. of three biological replicates are mentioned.

elements) significantly responded ABA treatment, moreover, the DEGs after MeJA treatment (*VvRDR1b*, *VvAGO7*-CGTCA-motif, TGACG-motif) and GA_3 treatment (*VvDCL2*, *VvAGO6a*, *VvAGO7*-P-box) also contained related elements, suggesting the involvement of *VvRdDM* genes in hormone signaling pathways. Previously, Mosharaf et al. (2020) has also reported the presence of more hormone-responsive cis-elements in *CsRdDM* genes. Apart from hormone-responsive elements, stress and growth-related elements were also identified, signifying the role of RdDM genes in disease resistance and grapevine growth signaling pathways. These findings predict that regulation of RdDM genes is a complex process and rationalizes the need for further research.

Transcriptome data from SRA can provide some information for further research. In this study, previously published transcriptome data was used to evaluate the response and expression of *VvRdDM* genes against different treatments. Following different stress treatments or in different organs, the AGO family showed high expression or significant expression change as compared to control, suggesting it's involvement in multiple biological processes (Figure 11). Previously, AGO family members have displayed diverse functions in different species. For example, in *A. thaliana* AGO mutant showed enhanced susceptibility to fungi (Ellendorff et al., 2009) and in another study *AtAGO4* mutants were associated with resistance to bacterial pathogens (Le et al., 2014). *SlAGO4A* has been reported for its role in salt and drought stress tolerance in tomato (Huang et al., 2016) and *OsAGO17* regulated seed size in rice (Zhong et al., 2020). AGO proteins are core components of the RNA-induced silencing complex (RISC) (Hutvagner and Simard, 2008), and this may be the potential explanation to its diverse roles. *VvRDM2b* and *VvAGO1* (Figure 7–9) gene showed a broad spectrum of high expression in different organs at various stages even under different treatment. *VvRdDM* genes not only showed diverse expression but also showed spatio-temporal expression patterns. For example, *VvRDR1a*, *VvIDN2a*, and *VvNRPD2b*, were significantly more expressed at the hard green stage than at

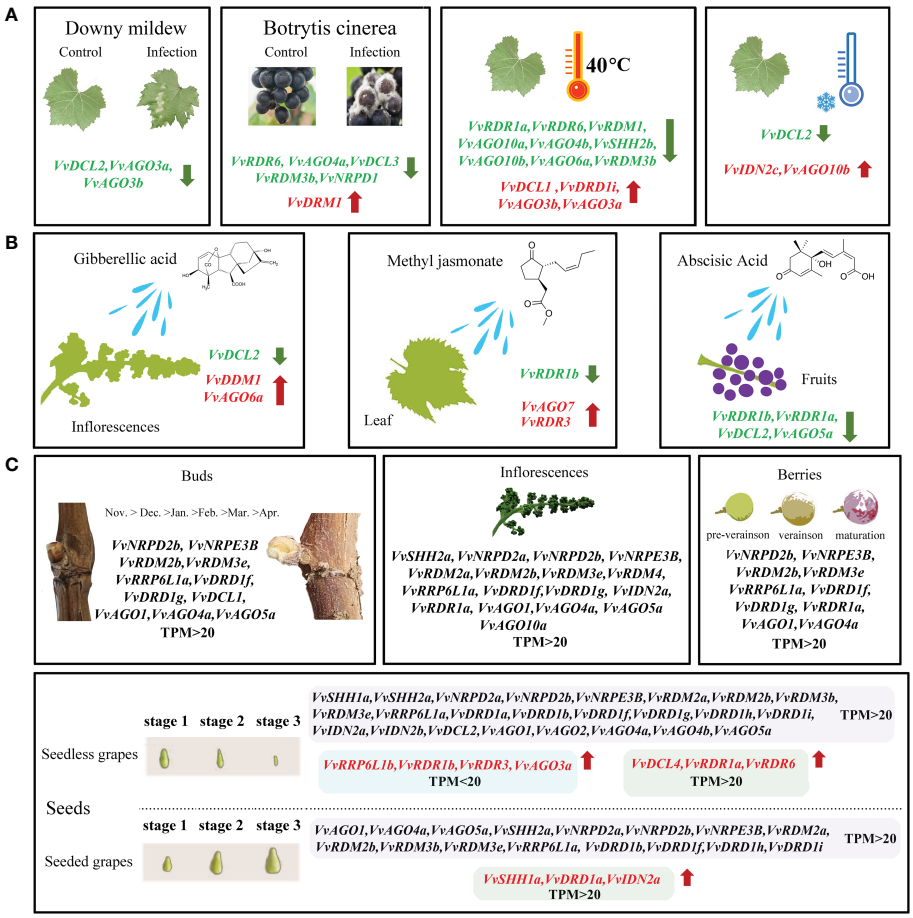


FIGURE 11
Transcriptomic and expression analysis of grapevine *VvRdDM* genes. The green and red font represents significant down and up-regulation, respectively. **(A)** The differential expression of *VvRdDM* genes following different biotic (downy mildew, botrytis cinerea), and abiotic (heat, cold) treatments. **(B)** The differential expression of *VvRdDM* genes under hormone treatments (GA_3, MeJA, ABA). **(C)** The expression of *VvRdDM* genes in different organs: buds, inflorescence, seeds, and berries. The genes having TPM>20 are listed. In the seeds, the red font indicates that the expression was significantly up-regulated during the seed development in the seedless cultivars as compared with the seeded cultivars and vice versa.

the ripened stage. *VvAGO5a* was significantly differentially expressed at all stages of fruit ripening. Most of the *VvRdDM* genes showed significant down-regulation following different treatments (Figure 11). Surprisingly, *VvDCL2* displayed same expression trend under the conditions of downy mildew, low temperature, GA_3 and ABA treatment. Further, the expression level of *VvDCL2* was significantly down-regulated following low temperature treatment (108 to 27) suggesting some role in the adaptation of grape to low temperatures. Our results are also supported by the earlier findings i.e., the activity of *DCL2* in *A. thaliana* was related with temperature (Zhang et al., 2012). *Prupe DCL* under drought treatment was differentially expressed in peach (Belal et al., 2022). These findings suggest, that the functions of DCL2 are diverse, and further research is needed for exploring the exact functions of *VvDCL2*. The *VvRdDM* genes showed differential expressions in different organs, indicating their involvement in the growth and development process. Our results are in line with the previous findings, as *AtRRP6L1* and *AtRRP6L2* (homologous genes of *VvRRP6L1a*) regulated flowering (Shin and Chekanova, 2014), *DDM1* (homolog *VvDDM1*) mutant effected seed development process (Xiao et al., 2006) in *A. thaliana*. Silencing of *FvAGO4* (homolog *VvAGO4a*) during fruit development resulted in early ripening of strawberry fruit (Cheng et al., 2018). Based on these findings, we speculate that *VvRdDM* genes can play critical role in the growth and development of grapes.

Expression patterns of fifteen genes during seed development were compared among four different (2 seeded and 2 seedless) cultivars of grapes. Interestingly, three genes (*VvSHH1a*, *VvDRD1a*, and *VvIDN2a*) showed higher expression in seeded cultivars especially at 40 and 50 DAF, signifying their key roles in grapevine seed development pathways. Previous

results support our findings; for example, in *Hypericum perforatum* L, a gene with high similarity to *AtIDN2* demonstrated differential expression in the ovule of apomictic *H. perforatum*. Furthermore, in *A. thaliana*, silencing the IDN2 gene led to decreased seed set and changes in seed size (Basso et al., 2019). However, in seedless cultivars, *VvRDR1a* and *VvRDR6* showed significantly higher expression throughout all phases of seed formation. Similar trend of higher expression was noted during transcriptomic analysis. Relying on the findings, we can predict that these two genes might have roles in ovule abortion in grapes. Our results are in concordant with the previous findings of Grover et al. (2018) who noticed the abortion of *brassica rapa* seeds due to maternal alterations in the Pol IV-dependent small RNA pathway. They reported that RdDM pathway genes have critical roles in *brassica rapa* seed development. These findings rationalize the need for a future study about the functional characterization of these (*VvRDR1a* and *VvRDR6*) genes.

Conclusion

A total of 60 RdDM genes in grapevine were identified, and comprehensive bioinformatic and expression analysis were performed. VvRdDM proteins were evaluated in terms of chromosomal positions, exon-intron distribution, phylogenetic, and evolutionary history. Expression analysis including transcriptomic analysis of *VvRdDM* genes following different stresses and hormonal treatments and qRT-PCR analysis of selected *VvRdDM* genes during different stages of seed development suggested their roles in multiple biological processes. *VvRDR1a* and *VvRDR6* gene were proposed as candidate gene resources for further functional characterization. Our study provides new resources for grape molecular breeding and inklings about the role of epigenetics in grapevine growth and development.

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

LW and RX designed the research. LW and GD supervised the experiments. RX performed most of the experiments. BA, LY and CL provided technical assistance. RX and XS analyzed the data. BA, RX and LW wrote the article with contributions of all the authors. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Natural Science Youth Foundation of Hebei, China (C2021204146), the Scientific Research Program of Hebei Educational Commission (QN2020232), the Strat-up Funds of Hebei Agricultural University to LW (YJ201806), the Earmarked Fund for Hebei Agriculture Research System (13000022P00E90410004M), and the Earmarked Fund for Sci-Tech Innovation in Modern Breeding Industry (21326310D).

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1089392/full#supplementary-material

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

Ahmad, B., Yao, J., Zhang, S., Li, X., Zhang, X., Yadav, V., et al. (2020). Genome-wide characterization and expression profiling of GASA genes during different stages of seed development in grapevine (*Vitis vinifera* l.) predict their involvement in seed development. *Int. J. Mol. Sci.* 21 (3), 1088. doi: 10.3390/ijms21031088

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25 (17), 3389–3402. doi: 10.1093/nar/25.17.3389

An, Y. C., Goettel, W., Han, Q., Bartels, A., Liu, Z., and Xiao, W. (2017). Dynamic changes of genome-wide DNA methylation during soybean seed development. *Sci. Rep.* 7 (1), 12263. doi: 10.1038/s41598-017-12510-4

Ausin, I., Mockler, T. C., Chory, J., and Jacobsen, S. E. (2009). IDN1 and IDN2 are required for *de novo* DNA methylation in *Arabidopsis thaliana*. *Nat. Struct. Mol. Biol.* 16 (12), 1325–1327. doi: 10.1038/nsmb.1690

Bailey, T. L., and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in bipolymers. *Proc. Int. Conf Intell. Syst. Mol. Biol.* 2, 28–36.

Bai, M., Yang, G. S., Chen, W. T., Mao, Z. C., Kang, H. X., Chen, G. H., et al. (2012). Genome-wide identification of dicer-like, argonaute and RNA-dependent RNA polymerase gene families and their expression analyses in response to viral infection and abiotic stresses in *Solanum lycopersicum*. *Gene.* 501 (1), 52–62. doi: 10.1016/j.gene.2012.02.009

Bartels, A., Han, Q., Nair, P., Stacey, L., Gaynier, H., Mosley, M., et al. (2018). Dynamic DNA methylation in plant growth and development. *Int. J. Mol. Sci.* 19 (7), 2144. doi: 10.3390/ijms19072144

Barter, M., Bui, C., and Young, D. (2012). Epigenetic mechanisms in cartilage and osteoarthritis: DNA methylation, histone modifications and microRNAs. *Osteoarthritis Cartilage.* 20 (5), 339–349. doi: 10.1016/j.joca.2011.12.012

Basso, A., Barcaccia, G., and Galla, G. (2019). Annotation and expression of IDN2-like and FDM-like genes in sexual and aposporous *Hypericum perforatum* l. accessions. *Plants.* 8 (6), 158. doi: 10.3390/plants8060158

Belal, M. A., Ezzat, M., Zhang, Y., Xu, Z., Cao, Y., and Han, Y. (2022). Integrative analysis of the *DICER-like* (DCL) genes from peach (*Prunus persica*): A critical role in response to drought stress. *Front. Ecol. Evol.* 10. doi: 10.3389/fevo.2022.923166

Bewick, A. J., and Schmitz, R. J. (2017). Gene body DNA methylation in plants. *Curr. Opin. Plant Biol.* 36, 103–110. doi: 10.1016/j.pbi.2016.12.007

Bitonti, M. B., Cozza, R., Chiappetta, A., Giannino, D., Castiglione, M. R., Dewitte, W., et al. (2002). Distinct nuclear organization, DNA methylation pattern and cytokinin distribution mark juvenile, juvenile-like and adult vegetative apical meristems in peach (*Prunus persica* (L.) batsch). *J. Exp. Botany.* 53 (371), 1047–1054. doi: 10.1093/jexbot/53.371.1047

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 (1), 1–21. doi: 10.1186/1471-2229-4-10

Cao, J. Y., Xu, Y. P., Li, W., Li, S. S., Rahman, H., and Cai, X. Z. (2016). Genome-wide identification of dicer-like, argonaute, and RNA-dependent RNA polymerase gene families in brassica species and functional analyses of their arabidopsis homologs in resistance to *Sclerotinia sclerotiorum*. *Front. Plant Sci.* 7. doi: 10.3389/fpls.2016.01614

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 (8), 1194–1202. doi: 10.1016/j.molp.2020.06.009

Cheng, J., Niu, Q., Zhang, B., Chen, K., Yang, R., Zhu, J.-K., et al. (2018). Downregulation of RdDM during strawberry fruit ripening. *Genome Biol.* 19 (1), 212. doi: 10.1186/s13059-018-1587-x

Daccord, N., Celton, J.-M., Linsmith, G., Becker, C., Choisne, N., Schijlen, E., et al. (2017). High-quality *de novo* assembly of the apple genome and methylome dynamics of early fruit development. *Nat. Genet.* 49 (7), 1099–1106. doi: 10.1038/ng.3886

Dias, A. P., Braun, E. L., McMullen, M. D., and Grotewold, E. (2003). Recently duplicated maize R2R3 myb genes provide evidence for distinct mechanisms of evolutionary divergence after duplication. *Plant Physiol.* 131 (2), 610–620. doi: 10.1104/pp.012047

Dinh, T. T., O'Leary, M., Won, S. Y., Li, S., Arroyo, L., Liu, X., et al. (2013). Generation of a luciferase-based reporter for CHH and CG DNA methylation in *Arabidopsis thaliana*. *Silence.* 4 (1), 1–11. doi: 10.1186/1758-907X-4-1

Ellendorff, U., Fradin, E. F., de Jonge, R., and Thomma, B. P. (2009). RNA Silencing is required for *Arabidopsis* defence against *Verticillium* wilt disease. *J. Exp. Bot.* 60 (2), 591–602. doi: 10.1093/jxb/ern306

El-Sharkawy, I., Liang, D., and Xu, K. (2015). Transcriptome analysis of an apple (Malus× domestica) yellow fruit somatic mutation identifies a gene network module highly associated with anthocyanin and epigenetic regulation. *J. Exp. Botany.* 66 (22), 7359–7376. doi: 10.1093/jxb/erv433

Erdmann, R. M., and Picard, C. L. (2020). RNA-Directed DNA methylation. *PloS Genet.* 16 (10), e1009034. doi: 10.1371/journal.pgen.1009034

Flaus, A., Martin, D. M., Barton, G. J., and Owen-Hughes, T. (2006). Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic Acids Res.* 34 (10), 2887–2905. doi: 10.1093/nar/gkl295

Gao, C., Deng, M., Yang, X., Yu, W., Cai, J., Shi, Y., et al. (2020). Genome-wide identification and coexpression network analysis of DNA methylation pathway genes and their differentiated functions in *Ginkgo biloba* l. *Forests.* 11 (10), 1076. doi: 10.3390/f11101076

Garcia-Ruiz, H., Takeda, A., Chapman, E. J., Sullivan, C. M., Fahlgren, N., Brempelis, K. J., et al. (2010). *Arabidopsis* RNA-dependent RNA polymerases and dicer-like proteins in antiviral defense and small interfering RNA biogenesis during *Turnip mosaic virus* infection. *Plant Cell.* 22 (2), 481–496. doi: 10.1105/tpc.109.073056

Grimplet, J., Adam-Blondon, A.-F., Bert, P.-F., Bitz, O., Cantu, D., Davies, C., et al. (2014). The grapevine gene nomenclature system. *BMC Genomics* 15 (1), 1–14. doi: 10.1186/1471-2164-15-1077

Grover, J. W., Kendall, T., Baten, A., Burgess, D., Freeling, M., King, G. J., et al. (2018). Maternal components of RNA-directed DNA methylation are required for seed development in *Brassica rapa*. *Plant J.* 94 (4), 575–582. doi: 10.1111/tpj.13910

He, X.-J., Hsu, Y.-F., Pontes, O., Zhu, J., Lu, J., Bressan, R. A., et al. (2009a). NRPD4, a protein related to the RPB4 subunit of RNA polymerase II, is a component of RNA polymerases IV and V and is required for RNA-directed DNA methylation. *Genes Dev.* 23 (3), 318–330. doi: 10.1101/gad.1765209

He, X.-J., Hsu, Y.-F., Zhu, S., Liu, H.-L., Pontes, O., Zhu, J., et al. (2009b). A conserved transcriptional regulator is required for RNA-directed DNA methylation and plant development. *Genes Dev.* 23 (23), 2717–2722. doi: 10.1101/gad.1851809

Henderson, I. R., Deleris, A., Wong, W., Zhong, X., Chin, H. G., Horwitz, G. A., et al. (2010). The *de novo* cytosine methyltransferase DRM2 requires intact UBA domains and a catalytically mutated paralog DRM3 during RNA–directed DNA methylation in *Arabidopsis thaliana*. *PloS Genet.* 6 (10), e1001182. doi: 10.1371/journal.pgen.1001182

Hirtreiter, A., Damsma, G. E., Cheung, A. C. M., Klose, D., Grohmann, D., Vojnic, E., et al. (2010). Spt4/5 stimulates transcription elongation through the RNA polymerase clamp coiled-coil motif. *Nucleic Acids Res.* 38 (12), 4040–4051. doi: 10.1093/nar/gkq135

Holliday, R. (2006). Epigenetics: a historical overview. *Epigenetics.* 1 (2), 76–80. doi: 10.4161/epi.1.2.2762

Holub, E. B. (2001). The arms race is ancient history in *Arabidopsis*, the wildflower. *Nat. Rev. Genet.* 2 (7), 516–527. doi: 10.1038/35080508

Huang, Y., Ji, L., Huang, Q., Vassylyev, D. G., Chen, X., and Ma, J.-B. (2009). Structural insights into mechanisms of the small RNA methyltransferase HEN1. *Nature.* 461 (7265), 823–827. doi: 10.1038/nature08433

Huang, H., Liu, R., Niu, Q., Tang, K., Zhang, B., Zhang, H., et al. (2019). Global increase in DNA methylation during orange fruit development and ripening. *Proc. Natl. Acad. Sci.* 116 (4), 1430–1436. doi: 10.1073/pnas.1815441116

Huang, W., Xian, Z., Hu, G., and Li, Z. (2016). SlAGO4A, a core factor of RNA-directed DNA methylation (RdDM) pathway, plays an important role under salt and drought stress in tomato. *Mol. Breeding.* 36 (3), 1–3. doi: 10.1007/s11032-016-0439-1

Hutvagner, G., and Simard, M. J. (2008). Argonaute proteins: key players in RNA silencing. *Nat. Rev. Mol. Cell Biol.* 9 (1), 22–32. doi: 10.1038/nrm2321

Hu, Y., Zhu, N., Wang, X., Yi, Q., Zhu, D., Lai, Y., et al. (2013). Analysis of rice Snf2 family proteins and their potential roles in epigenetic regulation. *Plant Physiol. Biochem.* 70, 33–42. doi: 10.1016/j.plaphy.2013.05.001

Jones, A. L., and Sung, S. (2014). Mechanisms underlying epigenetic regulation in *Arabidopsis thaliana*. *Integr. Comp. Biol.* 54 (1), 61–67. doi: 10.1093/icb/icu030

Kawakatsu, T., Nery, J. R., Castanon, R., and Ecker, J. R. (2017). Dynamic DNA methylation reconfiguration during seed development and germination. *Genome Biol.* 18 (1), 1–12. doi: 10.1186/s13059-017-1251-x

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evolution.* 35 (6), 1547. doi: 10.1093/molbev/msy096

Kurihara, Y., Matsui, A., Kawashima, M., Kaminuma, E., Ishida, J., Morosawa, T., et al. (2008). Identification of the candidate genes regulated by RNA-directed DNA methylation in *Arzabidopsis*. *Biochem. Biophys. Res. Commun.* 376 (3), 553–557. doi: 10.1016/j.bbrc.2008.09.046

Lange, H., Holec, S., Cognat, V., Pieuchot, L., Le Ret, M., Canaday, J., et al. (2008). Degradation of a polyadenylated rRNA maturation by-product involves one of the three RRP6-like proteins in *Arabidopsis thaliana*. *Mol. Cell. Biol.* 28 (9), 3038–3044. doi: 10.1128/MCB.02064-07

Lee, T.-H., Tang, H., Wang, X., and Paterson, A. H. (2012). PGDD: a database of gene and genome duplication in plants. *Nucleic Acids Res.* 41 (D1), D1152–D1158. doi: 10.1093/nar/gks1104

Le, T. N., Schumann, U., Smith, N. A., Tiwari, S., 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 (9), 458. doi: 10.1186/s13059-014-0458-3

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 (1), 325–327. doi: 10.1093/nar/30.1.325

Li, W.-F., Ning, G.-X., Mao, J., Guo, Z.-G., Zhou, Q., and Chen, B.-H. (2019). Whole-genome DNA methylation patterns and complex associations with gene expression associated with anthocyanin biosynthesis in apple fruit skin. *Planta*. 250 (6), 1833–1847. doi: 10.1007/s00425-019-03266-4

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods*. 25 (4), 402–408. doi: 10.1006/meth.2001.1262

López, A., Ramírez, V., García-Andrade, J., Flors, V., and Vera, P. (2011). The RNA silencing enzyme RNA polymerase v is required for plant immunity. *PloS Genet*. 7 (12), e1002434. doi: 10.1371/journal.pgen.1002434

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15 (12), 550. doi: 10.1186/s13059-014-0550-8

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 (8), 948–952. doi: 10.1038/ng1841

Margis, R., Fusaro, A. F., Smith, N. A., Curtin, S. J., Watson, J. M., Finnegan, E. J., et al. (2006). The evolution and diversification of dicers in plants. *FEBS Letters*. 580 (10), 2442–2450. doi: 10.1016/j.febslet.2006.03.072

Matzke, M. A., Kanno, T., and Matzke, A. J. (2015). RNA-Directed DNA methylation: The evolution of a complex epigenetic pathway in flowering plants. *Annu. Rev. Plant Biol.* 66, 243–267. doi: 10.1146/annurev-arplant-043014-114633

Matzke, M. A., and Mosher, R. A. (2014). RNA-Directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat. Rev. Genet.* 15 (6), 394–408. doi: 10.1038/nrg3683

Mosharaf, M. P., Rahman, H., Ahsan, M. A., Akond, Z., Ahmed, F. F., Islam, M. M., et al. (2020). In silico identification and characterization of AGO, DCL and RDR gene families and their associated regulatory elements in sweet orange (*Citrus sinensis* l.). *PloS One* 15 (12), e0228233. doi: 10.1371/journal.pone.0228233

Pertea, M., Kim, D., Pertea, G. M., Leek, J. T., and Salzberg, S. L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and ballgown. *Nat. Protoc.* 11 (9), 1650–1667. doi: 10.1038/nprot.2016.095

Qian, Y., Cheng, Y., Cheng, X., Jiang, H., Zhu, S., and Cheng, B. (2011). Identification and characterization of dicer-like, argonaute and RNA-dependent RNA polymerase gene families in maize. *Plant Cell Rep.* 30 (7), 1347–1363. doi: 10.1007/s00299-011-1046-6

Qin, L., Mo, N., Muhammad, T., and Liang, Y. (2018). Genome-wide analysis of DCL, AGO, and RDR gene families in pepper (*Capsicum annuum* l.). *Int. J. Mol. Sci.* 19 (4), 1038. doi: 10.3390/ijms19041038

Ream, T. S., Haag, J. R., Wierzbicki, A. T., Nicora, C. D., Norbeck, A. D., Zhu, J.-K., et al. (2009). Subunit compositions of the RNA-silencing enzymes pol IV and pol V reveal their origins as specialized forms of RNA polymerase II. *Mol. Cell.* 33 (2), 192–203. doi: 10.1016/j.molcel.2008.12.015

Sabbione, A., Daurelio, L., Vegetti, A., Talon, M., Tadeo, F., and Dotto, M. (2019). Genome-wide analysis of AGO, DCL and RDR gene families reveals RNA-directed DNA methylation is involved in fruit abscission in *Citrus sinensis*. *BMC Plant Biol.* 19 (1), 401. doi: 10.1186/s12870-019-1998-1

Sasaki, T., Lorković, Z. J., Liang, S.-C., Matzke, A. J., and Matzke, M. (2014). The ability to form homodimers is essential for RDM1 to function in RNA-directed DNA methylation. *PloS One* 9 (2), e88190. doi: 10.1371/journal.pone.0088190

Shin, J. H., and Chekanova, J. A. (2014). *Arabidopsis* RRP6L1 and RRP6L2 function in *FLOWERING LOCUS c* silencing *via* regulation of antisense RNA synthesis. *PloS Genet.* 10 (9), e1004612. doi: 10.1371/journal.pgen.1004612

Siltberg, J., and Liberles, D. (2002). A simple covarion-based approach to analyse nucleotide substitution rates. *J. Evolutionary Biol.* 15 (4), 588–594. doi: 10.1046/j.1420-9101.2002.00416.x

Smith, L. M., Pontes, O., Searle, I., Yelina, N., Yousafzai, F. K., Herr, A. J., et al. (2007). An SNF2 protein associated with nuclear RNA silencing and the spread of a silencing signal between cells in *Arabidopsis*. *Plant Cell.* 19 (5), 1507–1521. doi: 10.1105/tpc.107.051540

This, P., Lacombe, T., and Thomas, M. R. (2006). Historical origins and genetic diversity of wine grapes. *Trends Genet.* 22 (9), 511–519. doi: 10.1016/j.tig.2006.07.008

Wang, L., Ahmad, B., Liang, C., Shi, X., Sun, R., Zhang, S., et al. (2020). Bioinformatics and expression analysis of histone modification genes in grapevinepredict their involvement in seed development, powdery mildew resistance, andhormonal signaling. *BMC Plant Biology* 20 (1), 1–20. doi: 10.1186/s12870-020-02618-7

Wang, L., Hu, X., Jiao, C., Li, Z., Fei, Z., Yan, X., et al. (2016). Transcriptomeanalyses of seed development in grape hybrids reveals a possible mechanisminfluencing seed size. *BMC Genomics* 17 (1), 1–15. doi: 10.1186/s12864-016-3193-1

Wassenegger, M., and Krczal, G. (2006). Nomenclature and functions of RNA-directed RNA polymerases. *Trends Plant Science*. 11 (3), 142–151. doi: 10.1016/j.tplants.2006.01.003

Wei, S. J., Chai, S., Zhu, R. M., Duan, C. Y., Zhang, Y., and Li, S. (2020). *HUA ENHANCER1* mediates ovule development. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.00397

Wierzbicki, A. T., Haag, J. R., and Pikaard, C. S. (2008). Noncoding transcription by RNA polymerase pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. *Cell*. 135 (4), 635–648. doi: 10.1016/j.cell.2008.09.035

Xiao, W., Brown, R. C., Lemmon, B. E., Harada, J. J., Goldberg, R. B., and Fischer, R. L. (2006). Regulation of seed size by hypomethylation of maternal and paternal genomes. *Plant Physiol.* 142 (3), 1160–1168. doi: 10.1104/pp.106.088849

Xi, Y., Liu, J., Dong, C., and Cheng, Z.-M. (2017). The CBL and CIPK gene family in grapevine (*Vitis vinifera*): genome-wide analysis and expression profiles in response to various abiotic stresses. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.00978

Xu, G., Guo, C., Shan, H., and Kong, H. (2012). Divergence of duplicate genes in exon–intron structure. *Proc. Natl. Acad. Sci.* 109 (4), 1187–1192. doi: 10.1073/pnas.1109047109

Xu, D., Lu, Z., Jin, K., Qiu, W., Qiao, G., Han, X., et al. (2021). SPDE: A multi-functional software for sequence processing and data extraction. *Bioinformatics*. 37 (20), 3686–3687. doi: 10.1093/bioinformatics/btab235

Zhang, H., Lang, Z., and Zhu, J. K. (2018). Dynamics and function of DNA methylation in plants. *Nat. Rev. Mol. Cell Biol.* 19 (8), 489–506. doi: 10.1038/s41580-018-0016-z

Zhang, H., Ma, Z.-Y., Zeng, L., Tanaka, K., Zhang, C.-J., Ma, J., et al. (2013). DTF1 is a core component of RNA-directed DNA methylation and may assist in the recruitment of pol IV. *Proc. Natl. Acad. Sci.* 110 (20), 8290–8295. doi: 10.1073/pnas.1300585110

Zhang, X., Zhang, X., Singh, J., Li, D., and Qu, F. (2012). Temperature-dependent survival of *Turnip crinkle virus*-infected *arabidopsis* plants relies on an RNA silencing-based defense that requires dcl2, AGO2, and HEN1. *J. Virol.* 86 (12), 6847–6854. doi: 10.1128/jvi.00497-12

Zhao, Y., and Chen, X. (2014). Non-coding RNAs and DNA methylation in plants. *Natl. Sci. Review*. 1 (2), 219–229. doi: 10.1093/nsr/nwu003

Zhao, H., Zhao, K., Wang, J., Chen, X., Chen, Z., Cai, R., et al. (2014). Comprehensive analysis of dicer-like, argonaute, and RNA-dependent RNA polymerase gene families in grapevine (*Vitis vinifera*). *J. Plant Growth Regulation*. 34 (1), 108–121. doi: 10.1007/s00344-014-9448-7

Zhong, J., He, W., Peng, Z., Zhang, H., Li, F., and Yao, J. (2020). A putative AGO protein, OsAGO17, positively regulates grain size and grain weight through OsmiR397b in rice. *Plant Biotechnol. J.* 18 (4), 916–928. doi: 10.1111/pbi.13256

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 25 January 2023
DOI 10.3389/fpls.2023.1124046

OPEN ACCESS

EDITED BY
Chenxia Cheng,
Qingdao Agricultural University, China

REVIEWED BY
Keji Yu,
Beijing Forestry University, China
Chengpeng Wang,
Shandong Academy of Agricultural Sciences, China

*CORRESPONDENCE
Sihong Zhou
bear824@126.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Functional and Applied Plant Genomics, a section of the journal Frontiers in Plant Science

RECEIVED 14 December 2022
ACCEPTED 13 January 2023
PUBLISHED 25 January 2023

CITATION
Cheng G, Wu D, Guo R, Li H, Wei R, Zhang J, Wei Z, Meng X, Yu H, Xie L, Lin L, Yao N and Zhou S (2023) Chromosome-scale genomics, metabolomics, and transcriptomics provide insight into the synthesis and regulation of phenols in *Vitis adenoclada* grapes.
Front. Plant Sci. 14:1124046.
doi: 10.3389/fpls.2023.1124046

Chromosome-scale genomics, metabolomics, and transcriptomics provide insight into the synthesis and regulation of phenols in *Vitis adenoclada* grapes

Guo Cheng[1†], Daidong Wu[1†], Rongrong Guo[1†], Hongyan Li[1], Rongfu Wei[1], Jin Zhang[1], Zhiyong Wei[2], Xian Meng[2], Huan Yu[1], Linjun Xie[1], Ling Lin[1], Ning Yao[3] and Sihong Zhou[1*]

[1]Grape and Wine Research Institute, Guangxi Academy of Agricultural Sciences, Nanning, China, [2]Bureau of Agriculture and Rural Affairs of Luocheng Mulao Autonomous County, Hechi, China, [3]Guangxi Luocheng Maoputao Experimental Station, Hechi, China

Vitis adenoclada is a wild grape unique to China. It exhibits well resistance to heat, humidity, fungal disease, drought, and soil infertility. Here, we report the high-quality, chromosome-level genome assembly of GH6 (*V. adenoclada*). The 498.27 Mb genome contained 221.78 Mb of transposable elements, 28,660 protein-coding genes, and 481.44 Mb of sequences associated with 19 chromosomes. GH6 shares a common ancestor with PN40024 (*Vitis vinifera*) from approximately 4.26–9.01 million years ago, whose divergence occurred later than *Vitis rotundifolia* and *Vitis riparia*. Widely-targeted metabolome and transcriptome analysis revealed that the profiles and metabolism of phenolic compounds in *V. adenoclada* varieties significantly were differed from other grape varieties. Specifically, *V. adenoclada* varieties were rich in phenolic acids and flavonols, whereas the flavan-3-ol and anthocyanin content was lower compared with other varieties that have *V. vinifera* consanguinity in this study. In addition, ferulic acid and stilbenes content were associated with higher expressions of COMT and STSs in *V. adenoclada* varieties. Furthermore, MYB2, MYB73-1, and MYB73-2 were presumably responsible for the high expression level of COMT in *V. adenoclada* berries. MYB12 (MYBF1) was positively correlated with PAL, CHS, FLS and UFGT.Meanwhile, MYB4 and MYBC2-L1 may inhibit the synthesis of flavan-3-ols and anthocyanins in two *V. adenoclada* varieties (YN2 and GH6). The publication of the *V. adenoclada* grape genome provides a molecular foundation for further revealing its flavor and quality characteristics, is also important for identifying favorable genes of the East Asian species for future breeding.

KEYWORDS

***Vitis adenoclada*, phenols, genomics, metabolomics, transcriptomics, metabolic regulation**

1 Introduction

Vitis adenoclada is a wild grape species native to China that is commonly distributed in Hunan, Fujian, Guangxi, and other provinces south of the Yangtze River. It belongs to the East Asian population of *Vitis* spp.–Maoputao group (Niu and He, 1996). Of note, *V. adenoclada* is easily confused with *Vitis heyneana*. Actually, *V. adenoclada* possesses unique purplish brown glandular hairs on the new shoots and old mature vines (Wang, 1988) while *V. heyneana* does not, which is the biggest biological difference between them (Kong, 2004; Liu C. H., 2012). There was also study indicated that *V. adenoclada* should be downgraded into a variety of *V. heyneana* (Xie et al., 2021). According to field observations, the presence and density of glandular hairs varies with variety, habitat, nutrition, and other factors in *V. adenoclada* (Supplementary Figure 1). Guangxi Province is located in Southern of China, and it is one of the original locations of the East Asian population of *Vitis* spp. The results of the Third National Crop Germplasm Resources Survey in China revealed that *V. heyneana*, *V. adenoclada*, *Vitis davidii*, and *Vitis pseudoreticulata* are widely distributed in the region (Chen et al., 2020). There is a long history of viniculture by using wild grape species in Guangxi, and *V. adenoclada* grape is an important raw material for wine making in local area (Cheng et al., 2018). Since 2011, research has been carried out to breed the *V. adenoclada* varieties. Thus far, a series of excellent *V. adenoclada* varieties have been established and designated the "Guiheizhenzhu" series (Wu et al., 2020). They adapt to the climate and environmental conditions of the south tropical and subtropical regions, and exhibit resistance to heat and humidity, fungal diseases and pests, drought, and soil infertility (Wu et al., 2020).

Phenolics are important compounds that affect the sensory quality of grapes, and subsequently wine. The composition and content of phenolic compounds are determined by genotype (species/variety), ecological conditions, viticulture practice and other factors (Sun et al., 2019). In addition to the reported structural genes, the synthesis of these compounds is also regulated by many transcription factors, such as MYB, bHLH, WRKY, AP2/EREBP, C2C2, NAC, and C2H2 (Sun et al., 2017). In recent years, more and more studies have been conducted on the regulatory mechanism of phenolic metabolism in different grape varieties using multi-omics methods (Sun et al., 2019; Ju et al., 2020; Yang et al., 2020). Since 2007, a highly homozygous genotype, the inbred Pinot noir line, PN40024 (*Vitis vinifera*), has been sequenced, marking the beginning of the genomics era in grape research (Jaillon et al., 2007). Subsequently, genomes sequencing of different varieties have been carried out to better serve for research in *Vitis* genus, such as *V. vinifera* of Sultanina (Di Genova et al., 2014), Cabernet Sauvignon (Chin et al., 2016) and Chardonnay (Zhou et al., 2019), *Vitis riparia* (Patel et al., 2020), *Vitis arizonica* (Massonnet et al., 2020), *Vitis rotundifolia* (Cochetel et al., 2021; Park et al., 2022), *Vitis amurensis* (Wang et al., 2021). However, there are few reports about wild grape resources, especially focused on the quality characteristics and metabolic mechanism of related varieties using these self-testing genomes combined with transcriptome and metabolome.

In this study, we established a high-quality *de novo* genome assembly of GH6 (*V. adenoclada*). We created a chromosome-level assembly with an overall scaffold length of 498.27 Mb that included 28,660 annotated genes using a combination of Illumina and Oxford Nanopore Technologies (ONT) sequencing data and high-throughput chromosome conformation Capture (Hi-C) mapping. Notably, we also used this self-testing genome to examine the metabolism of phenolics among the different grape varieties. Through a combination of phenolic-associated metabolic studies and transcriptome analysis, we constructed a regulatory network of biosynthesis of resveratrol, phenolic acid, flavonol, flavan-3-ol, and anthocyanin. Furthermore, we identified key transcription factors that modulate phenolic metabolism using transcription factor prediction and co-expression network analyses. Overall, the established genome sequence is not only important for understanding the quality characteristics of the *V. adenoclada*, but it will also contribute to the further development and utilization of East Asian grape resources.

2 Materials and methods

2.1 Plant material, berry sampling, and physical chemical index analysis

The experimental location was in the vineyards of the Guangxi Academy of Agricultural Sciences' Grape and Wine Research Institute in Nanning, Guangxi Province. The materials included eight varieties (Supplementary Figure 2): Cabernet Sauvignon (*V. vinifera*, CS), Marselan (*V. vinifera*, Mar), Petit Verdot (*V. vinifera*, PV), NW196 (*V. heyneana* × *V. vinifera*), Yeniang No.2 (*V. adenoclada*, YN2), Guiheizhenzhu No.4 (*V. adenoclada*, GH4), Guiheizhenzhu No.5 (*V. adenoclada*, GH5), and Guiheizhenzhu No.6 (*V. adenoclada*, GH6). Furthermore, YN2, GH4 and GH6 are bisexual flower varieties, whereas GH5 is a unisexual flower variety.

For genome sequencing and assembly, the GH6 plant was employed. The young fresh GH6 leaves were collected and swiftly frozen in liquid nitrogen. In addition, other young fresh GH6 leaves were collected, sliced using sharp blades, and fixed in a 2% formaldehyde solution at room temperature for 90 minutes before the cross-linking reaction being stopped by the addition of 2.5 M glycine. The tissues were treated for Hi-C library creation after being frozen in liquid nitrogen.

At harvest, berries from eight varieties in three biological replicates were gathered. 120 berries were randomly selected from at least 30 clusters within 9 vines for each biological replicate. After being transported to the laboratory, a subsample of 50 berries from each biological replicate were measured for fresh weight, pH, total soluble solids (TSS), and titratable acidity (TA) content. The leftover berries were flash-frozen in liquid nitrogen and stored at -80°C for further metabolomic and transcriptomic analyses.

2.2 Genome survey and sequencing

High-quality genomic DNA from GH6 was extracted *via* a CTAB-based protocol. DNA libraries with fragment lengths of about 350 bp were created using the Illumina-provided standard protocols. The libraries were sequenced in paired-end mode on an Illumina Novaseq 6000 platform with read lengths of 150 bp. The

sequencing results were used to assess the genomic parameters of GH6 *via* K-mer analysis, such as genome size, GC content, heterozygosity, and the frequency of repeat sequences.

ONT's standard protocol was followed for genome sequencing. To summarize, genomic DNA was randomly disrupted, and large DNA fragments were collected *via* the BluePippin device. The SQK-LSK109 kit was used to generate DNA libraries. Fragmentation, end repair, ligation of sequencing adapters, and magnetic bead purification were all performed on DNA fragments. Following that, DNA sequencing was carried out on the PromethION platform. All genome sequencing procedures were conducted by the Biomarker Technologies Corporation (Beijing, China).

2.3 Genome assembly and assessment

Raw Nanopore data were formatted, sequencing adapters were removed, and low-quality or short-length (<2000 bp) reads were filtered. After corrected using Canu (Koren et al., 2017), WTDBG (https://github.com/ruanjue/wtdbg) was used to assemble nanopore readings into contigs. The assembled contigs were further calibrated using Racon (Vaser et al., 2017) with two iterations and then polished using four iterations of Pilon (Walker et al., 2014) with the Illumina sequencing reads. Assembly quality was assessed based on three ways: CEGMA (Parra et al., 2007) (v2.5) and BUSCO (Simão et al., 2015) (v2.0) were used to examine the fullness of the core genes; Illumina sequencing data were mapped to the assembled genome using BWA (Li and Durbin, 2009) to estimate the mapping rates.

2.4 Construction of a Hi-C library and chromosomal assembly

Hi-C fragment libraries with a 300-700 bp insert size were constructed following the protocols described by Rao et al. (2014), then sequenced with the Illumina platform. To summarize, raw read adapter sequences were trimmed, and low-quality PE reads were deleted to clean the data. The clean Hi-C reads were first trimmed at the putative Hi-C junctions, and the trimmed reads were then BWA (Li and Durbin, 2009) (v0.7.10-r789) aligned to the assembly results. Invalid read pairs containing dangling-ends, self-cycles, re-ligation, and dumped products were removed by HiC-Pro (Servant et al., 2015) (v2.8.1). Only uniquely mapped read pairs were retained for assembly using LACHESIS (Burton et al., 2013). Following this procedure, the placement and orientation abnormalities that indicated clear discrete chromatin interaction patterns were manually corrected.

2.5 Genome annotation

Firstly, ab initio prediction for the repeat sequences was performed by using RepeatModeler2 (Flynn et al., 2020) (v2.0.1) with the softwares of RECON (Bao and Eddy, 2002) (v1.0.8) and RepeatScout (Price et al., 2005) (v1.0.6), then RepeatClassifier (Flynn et al., 2020) with database of Dfam (Wheeler et al., 2013) (v3.5) was used to classify the results of the prediction. Secondly, long terminal repeats (LTRs) were predicted based on the ab initio principle by using LTR_retriever (Ou and Jiang, 2018) (2.9.0) with LTRharvest (Ellinghaus et al., 2008) (v1.5.10) and LTR FINDER (Xu and Wang, 2007) (v1.07). It was then merged with all above predicted outcomes as the final repeat sequence database. RepeatMasker (Tarailo-Graovac and Chen, 2009) (v4.1.2) was used to predict the transposable elements (TEs) of GH6 based on the constructed repeat sequence database.

The prediction of protein-coding genes of the GH6 genome was done *via* three different strategies namely: ab initio prediction, homologous prediction, and RNA-seq prediction. Ab initio prediction was performed using Genscan (Burge and Karlin, 1997), Augustus (Stanke and Waack, 2003) (v2.4), GlimmerHMM (Majoros et al., 2004) (v3.0.4), GeneID (Blanco et al., 2007) (v1.4), and SNAP (Korf, 2004) (v2006-07-28). The homologous prediction of protein-coding genes based on other species (*V. vinifera*, *Z. jujuba*, *A. thaliana*, *O. sativa*) was done using GeMoMa (Keilwagen et al., 2016; Keilwagen et al., 2018) (v1.3.1). HISAT (Kim et al., 2015) (v2.0.4) and Stringtie (Pertea et al., 2015) (v1.2.3) were employed for assembly based on RNA-seq data with reference transcripts, then gene prediction was performed with TransDecoder (v2.0) (http://transdecoder.github.io) and Genemarks-T (Tang et al., 2015) (v5.1). Meanwhile, the prediction of Unigene sequences through the unreferenced assembly of RNA-seq data was performed with PASA (Campbell et al., 2006) (v2.0.2). Lastly, the prediction results of the above three methods were amalgamated *via* EVM (Haas et al., 2008) (v1.1.1). The predicted gene sequences were labelled annotations with functional databases including NR (Marchler-Bauer et al., 2011), KOG (Tatusov et al., 2001), KEGG (Kanehisa and Goto, 2000), and TrEMBL (Boeckmann et al., 2003) by BLAST (Altschul et al., 1990) (v2.2.31). Functional annotation of GO (Dimmer et al., 2012) was performed with Blast2GO (Conesa et al., 2005).

2.6 Comparative genomic analyses

Using Orthofinder (Emms and Kelly, 2019) (v2.4), protein sequences from *V. adenoclada* and nine other representative species were obtained for gene family clustering. The resulting gene families were further annotated using the PANTHER (Mi et al., 2019) (v15) database. Using the maximum likelihood approach and IQ-TREE (Nguyen et al., 2015) (v1.6.11), single-copy protein sequences were utilized to build a phylogenetic tree for *V. adenoclada* and the other nine species. The root was set to *A. trichopoda*, and the number of bootstraps was set to 1000. Subsequently, the divergence times were estimated using MCMCTREE (Puttick, 2019) in the PAML (Yang, 1997) package (v4.9i) and calibrated using the TimeTree (Kumar et al., 2017) website (http://www.timetree.org/). Based on the phylogenetic tree with divergence times and gene family clustering, the gene family expansion and contraction analysis were performed by CAFE (Han et al., 2013) (v4.2). The gene family members from the ancestor of each branch were estimated using the birth mortality model, which was applied to infer the contraction and expansion of the gene families (p-values<0.05). PANTHER was used to annotate the expanded and contracted gene families identified in *V. adenoclada*, and ClusterProfile was used to perform GO enrichment analyses on these families (Yu et al., 2012).

MUMmer (Marcais et al., 2018) (v4.0.0rc1) was used to identify the collinear blocks of two species genomes. Subsequently, visualization of Genome collinearities between *V. adenoclada* and the other three grape species of *V. vinifera*, *V. riparia*, and *V. rotundifolia* was performed by NGenomeSyn (https://github.com/hewm2008/NGenomeSyn). Using Diamond (Buchfink et al., 2015) (v0.9.29.130), the gene sequences of two species were compared and comparable gene pairs were determined. *V. adenoclada* was compared with *V. vinifera*, *V. riparia*, and *V. rotundifolia*. Genomes of *V. adenoclada*, *V. vinifera*, *V. riparia*, *V. rotundifolia*, and *Z. jujuba* were used for WGD analyses. Based on the distribution of 4DTv rate, which was estimated using the HKY model (Hasegawa et al., 1985) and a Perl script (https://github.com/JinfengChen/Scripts), WGD events were determined.

2.7 Widely-targeted metabolomic analysis

Metabolites detection was performed with the help of Metware Biotechnology Co., Ltd. (Wuhan, China). A vacuum freeze-dryer was used to freeze-dry eight fruit samples with three biological replicates. Lyophilized powder (100 mg) was dissolved in 1.2 ml of 70% methanol solution, vortexed for 30 seconds every 30 minutes for a total of 6 times, and stored in a refrigerator overnight at 4°C. After centrifugation at 12,000 rpm for 10 minutes, the extracts were filtered (SCAA-104, 0.22 μm pore size; ANPEL, Shanghai, China), then analyzed by UPLC-ESI-MS/MS. The analytical conditions, raw data preprocessing, basis data analysis, KEGG annotation, and metabolic pathway analyses of differential metabolites all referred to the previous report (Wang et al., 2020). The mass spectrometry data was processed using Analyst 1.6.3 software (AB Sciex, Framingham, MA, USA). The identified metabolites were annotated based on the KEGG compound database (http://www.kegg.jp/kegg/compound/) and then mapped to the KEGG pathway database (http://www.kegg.jp/kegg/pathway.html).

2.8 RNA extraction, library construction, and sequencing

RNA extraction, library creation, and sequencing were done according to a previously reported method (Fu et al., 2021), and Illumina sequencing was carried out by the Gene Denovo Biotechnology Co. Ltd. (Guangzhou, China). To guarantee high-quality clean reads for further assembly and analysis, reads were filtered by fastp (Chen et al., 2018) (v0.18.0). The parameters were chosen to eliminate adapter-carrying reads, reads having more than 10% unknown nucleotides (N), and low-quality reads containing more than 50% low-quality (Q-value ≤20) bases. The short read alignment tool, Bowtie 2 (Langmead and Salzberg, 2012) (v2.2.8), was used for mapping reads to the ribosomal RNA (rRNA) database. The mapped rRNA readings were then deleted and the remaining clean reads were used for assembly and calculating gene abundance. HISAT (Kim et al., 2015) (v2.2.4) was used to map the clean reads to self-assembled genomes. StringTie (Pertea et al., 2015; Pertea et al., 2016) (v1.3.1) was used to assemble the mapped reads for each sample using a reference-based technique. Using RSEM (Li and Dewey, 2011), an FPKM (fragment per kilobase of transcript per million mapped reads) value was computed for each transcriptional domain to evaluate its expression abundance and variation. All transcripts were annotated using databases such as GO, KEGG, NR, and Swiss-Prot, and RNA differential expression was analyzed using DESeq2 (Love et al., 2014) between two groups. DEGs or transcripts were defined as genes or transcripts with an FDR less than 0.05 and an absolute fold-change ≥2.

2.9 Transcription factor (TF) analysis

Considering the important role of TFs in the synthesis of phenols, the TFs expressed in all samples were identified. All putative TFs were retrieved by the predicted protein sequences compared with the plant TF database (TFdb) using hmmscan. For structural genes of the same family, the highest expression amount of gene with similar expression pattern was analyzed by clustering screening. Furthermore, a co-expression analysis was done between the phenolic biosynthetic pathway genes and TF genes (correlation coefficient >0.85). Networks were visualized using Cytoscape (Shannon et al., 2003) v.3.7.1.

2.10 Weighted gene co-expression network analysis

The overlapping DEGs and DAPs (differentially accumulated phenolics) were selected for co-expression network analysis using the WGCNA (v1.47) package in R (Langfelder and Horvath, 2008). More than half of the samples with genes of low abundance (FPKM value < 0.8) were filtered out to decrease the interference in the network analysis. The co-expression modules were obtained using the automatic network construction function (blockwiseModules) with default parameters, apart from the soft threshold power of 10, TOM type was signed, merge CutHeight was 0.6, and the minModuleSize was 50. After the initial module division, we obtained the Dynamic Tree Cut of the preliminarily divided module. Because some modules are very similar, we also merged the modules with similar expression modes according to the similarity of module eigenvalues to obtain 9 merged modules. Furthermore, the correlation coefficients between the hub genes in the module and the DAPs were calculated using OmicShare tools.

3 Results

3.1 Sequencing, assembly, and quality assessment of genome

The GH6 plant, a variety of *V. adenoclada* (Figure 1A), was chosen for genome sequencing and assembly. A genome survey was performed to analyze the Illumina sequencing data. According to a K-mer analysis, the approximate genome size of GH6 was 524.23 Mb, with a heterozygosity of 0.62% and a repeat sequence proportion of 49.18% (Supplementary Table 1). The ONT platform generated a total of 103.24 Gb of raw data. After cleaning, 94.99 Gb clean reads were obtained with a mean read length of 25.87 kb. Nanopore sequencing

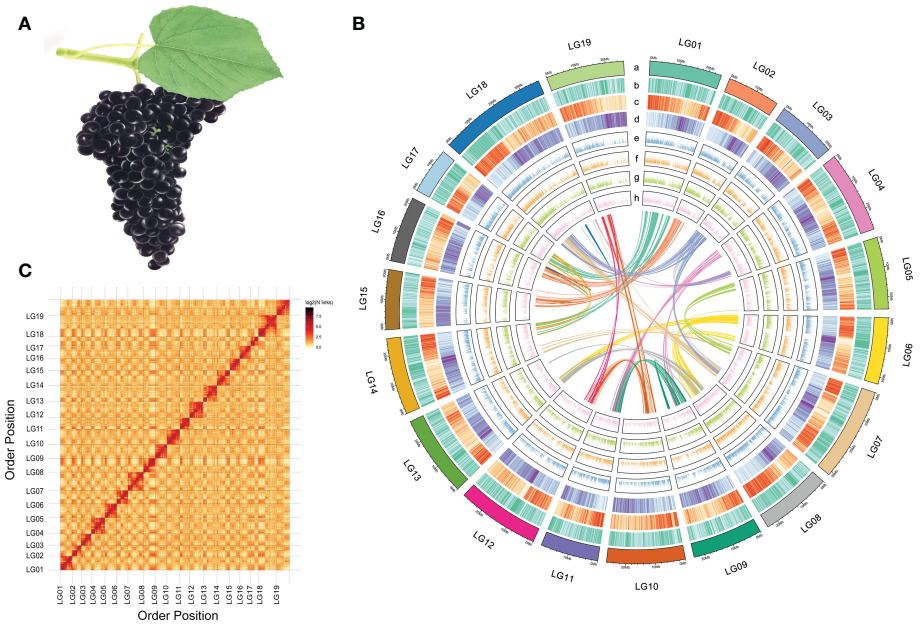


FIGURE 1
Morphology and genome information of *V. adenoclada*. **(A)** Drawing of *V. adenoclada* specimens. **(B)** Landscape of the *V. adenoclada* genome. a: chromosome ideograms; b: GC density; c: gene density; d: TE density; e~h: gene expression levels of YN2, GH4, GH5, and GH6, respectively; center: syntenic blocks within the genome. **(C)** Hi-C interaction heatmap of *V. adenoclada* genome.

clean reads were subjected to genome assembly, calibrating, and polishing. The draft genome assembly size for GH6 was 498.21 Mb with a contig N50 of 2.91 Mb (Supplementary Table 2).

A total of 54.56 Gb data sequenced using the Illumina platform were used to construct a chromosome-level genome assembly for GH6. After assessing and filtering the paired-end reads, valid interaction pairs were applied to facilitate the Hi-C assembly (Supplementary Table 3). As a result, 481.44 Mb sequences could be anchored to chromosomes, accounting for 96.63% of the contig genome assembly (Supplementary Table 4). The resulting contigs were clustered into 19 pseudochromosomes (Figure 1B), of which 462.46 Mb (96.06%) could be verified by order and direction (Supplementary Table 5). Ultimately, the final chromosome-scale genome assembly of GH6 was 498.27 Mb with a scaffold N50 of 25.26 Mb (Supplementary Table 4). A chromosomal interaction heatmap was created to demonstrate a pattern consistent with the Hi-C genome assembly and to confirm the pseudochromosome construction (Figure 1C).

The assembly quality and completeness of the GH6 genome was assessed by two methods, CEGMA and BUSCO. CEGMA analysis indicated that the assembled genome fully recalled 422 (92.14%) of the 458 core eukaryotic genes (CEGs) and 184 (74.19%) of the 248 extremely conserved CEGs. BUSCO analysis revealed that 1463 (90.64%) of the 1614 orthologs from the Embryophyta dataset were fully captured in the assembly (Supplementary Table 6). In addition, when Illumina sequencing data were mapped to the assembled genome, the mapping rate was 96.97% (the proper mapping rate was 92.15%) against the genome assembly.

3.2 Genome annotation

TE is one main type of the repetitive sequences. The fully assembled genome of GH6 contained 221.78 Mb (44.51%) of TEs which distributed in 19 chromosomes (Figure 1B). LTR retrotransposons, which included Gypsy repeats (10.05%) and Copia repeats (10.41%), were the most prominent class of repetitive sequences (Supplementary Table 7). There were 28,660 protein-coding genes predicted with a full length of 150.75 Mb that were randomly distributed across the 19 chromosomes. A total of 27,711 (96.69%) genes were labelled functional annotations by BLAST using the GO, KEGG, KOG, TrEMBL, and NR databases (Supplementary Table 8-9).

3.3 Comparative genomic analyses

The genome sequences of nine demonstrative plant species (Supplementary Table 10) were selected to perform a gene family cluster analysis with the genome sequence of GH6 (*V. adenoclada*) along with three grapes of PN40024 (*V. vinifera*), Riparia Gloire de Montpellier (*V. riparia*), and Trayshed (*V. rotundifolia*), common jujube (*Ziziphus jujuba*), apple (*Malus domestica*), kiwifruit (*Actinidia chinensis*), another dicots (*Arabidopsis thaliana*), a monocot (*Oryza sativa*), and one in the basal lineage of angiosperms (*Amborella trichopoda*). All genes from 10 selected plant species were clustered into 37,190 gene families. In GH6, a total of 19,775 gene families were identified, 178 of which (comprising 421 genes) were unique to the GH6 genome (Supplementary Figure 3). Moreover, in GH6, a total of 16,507 single-copy genes accounted for 57.6 percent of the predicted genes, equal to PN40024 (16,945/56.7%) but higher than Riparia Gloire de Montpellier (14,281/54.8%) and Trayshed (14,152/55.1%) (Figure 2A). The clustering of gene families in the four grapes indicated that GH6 harbors 475 specific gene families compared with the other three grapes (Figure 2B).

From the ten species, 417 conserved single-copy orthologs were identified and utilized to generate a phylogenetic tree with *A. trichopoda* as an outgroup. According to phylogenetic analysis, *V. adenoclada* is closely connected to *V. vinifera* and forms a clade with *V. riparia* and V. *rotundifolia*. Among four grapes, *V. rotundifolia* of *Muscadinia* and the other grapes of *Euvitis* diverged from their common ancestor at approximately 8.02–24.68 Mya. *V. riparia* of the North America population diverged at approximately 5.15–11.44 Mya before *V. adenoclada* of the East Asian population, which diverged approximately 4.26–9.01 Mya with *V. vinifera* of the European population (Figure 2C). These outcomes were consistent with those of prior researches (Liang et al., 2019). Based on the phylogenetic tree, 142 and 223 gene families were contracted and expanded in *V. adenoclada*, respectively (Figure 2C). GO functional analysis revealed that the contracted gene families of *V. adenoclada* were involved in lignin catabolic process, apoplast, ADP binding, etc., whereas the expanded gene families of *V. adenoclada* were involved in DNA integration, extracellular region, and ADP binding, etc. (Supplementary Figure 4). The gene families of expansion and contraction in the *V. adenoclada* genome relative to their most

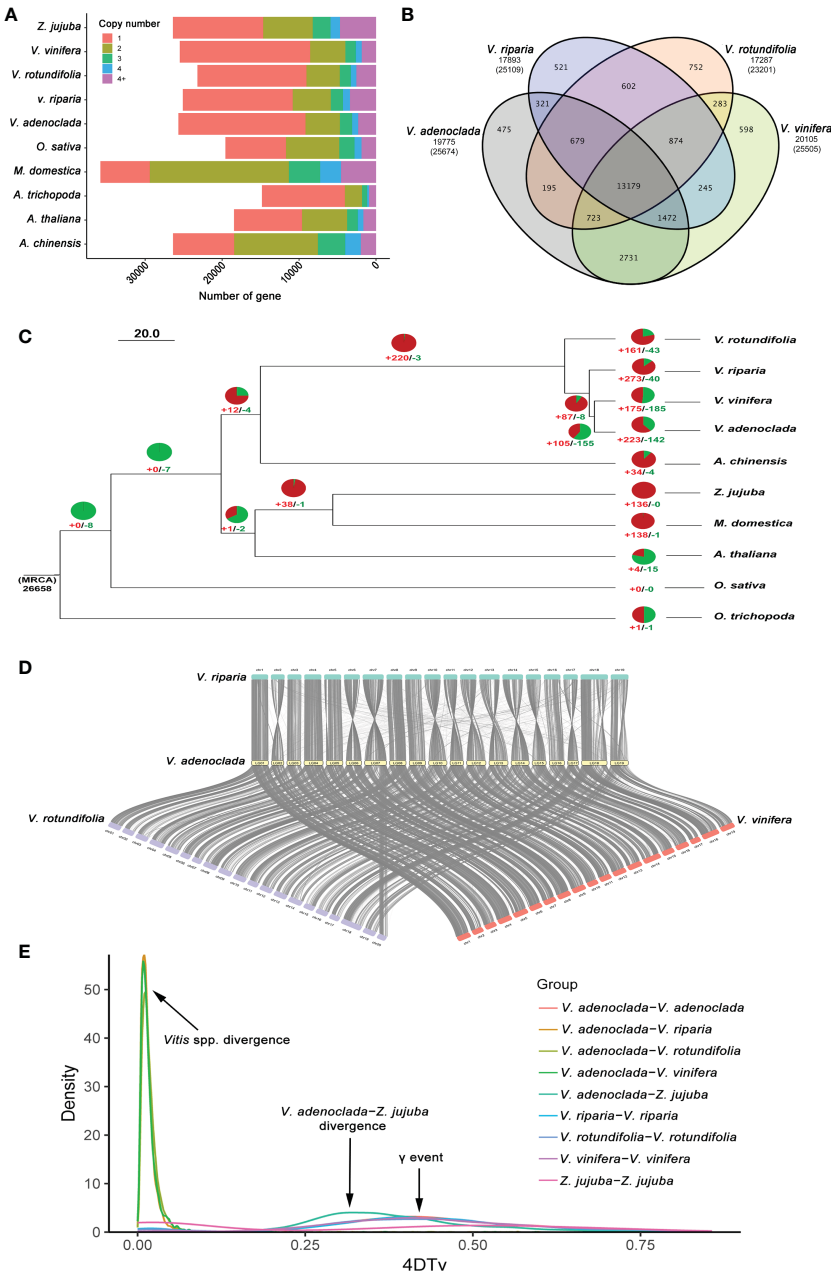


FIGURE 2
Comparative genomics analysis of *V. adenoclada* and other representative plant species. **(A)** Gene copy number distribution in *V. adenoclada* and nine other plant species. **(B)** Venn diagram of gene families in *V. adenoclada* and the other three grapes. **(C)** Phylogenetic analysis, gene family expansion/contraction analyses and branching time approximations. Green and red represent the number of gene family contraction and expansion occurrences, respectively. Branching times (Mya) are denoted by the numbers adjacent to the nodes. **(D)** Genome collinearity analyses between *V. adenoclada* and the other three grapes of *V. vinifera*, *V. riparia*, and *V. rotundifolia*. **(E)** Distribution of 4DTv among *V. adenoclada*, *V. vinifera*, *V. riparia*, *V. rotundifolia*, and *Z. jujuba* in intra- and intergenomic comparisons.

recent common ancestor (MRCA) are annotated in Supplementary Table 11-12.

Genome collinearity analyses between *V. adenoclada* and the other three grapes of *V. vinifera*, *V. riparia*, and *V. rotundifolia* are illustrated (Figure 2D). The findings suggested a high degree of gene order and locus conservation between *V. adenoclada* and other three grapes, and chromosome 7 of *V. adenoclada* was observed to be divided into chromosomes 7 and 20 in *V. rotundifolia*, similar as previously reported (Cochetel et al., 2021; Park et al., 2022). Five genomes of *V. adenoclada*, *V. vinifera*, *V. riparia*, *V. rotundifolia*, and *Z. jujuba* were used to calculate distribution of four-fold synonymous third-codon transversion (4DTv) rate (Figure 2E). The results showed that all four grapes underwent ancient whole-genome triplication (γ event) in all core eudicots of ~120 Mya (Jiao et al., 2011) before *V. adenoclada* diverged from *Z. jujuba*, whereas there were no recent whole-genome duplication (WGD) events that occurred in the genomes of the four grapes. The 4DTv rate distribution among the species suggested that *Vitis* spp. didn't occur divergence until very recent age.

3.4 Metabolic profiling differences

Metabolic profiling was done to identify the metabolite characteristics, especially the quality related products in *V. adenoclada* varieties. Fundamental physical and chemical indexes of the fruit are listed in Supplementary Table 13. Widely-targeted metabolomic analyses revealed 674 metabolites, which included organic acids, phenolic acids, tannins, flavonoids, and terpenes (Supplementary Table 14). A comparison of the results for all metabolites among the different varieties are presented (Figure 3A). These metabolites were divided into two groups by horizontal clustering. Group I included some amino acids and derivatives, most organic acids and phenolic acids, and most flavonols, which accumulated preferentially in *V. adenoclada* varieties. Group II included various anthocyanins, flavanols, tannins, and some lipids, which accumulate more in *V. vinifera* varieties.

PCA divided all eight varieties into three groups and the result was consistent with the cluster dendrogram (Figures 3B, C). The results indicated that the four *V. adenoclada* varieties may be grouped into one class and GH6 is much closer to YN2 compared with GH4 and GH5. Although NW196, is closer to *V. adenoclada* varieties compared with the three *V. vinifera* varieties, its metabolic profile showed different characteristics.

3.5 Differential accumulation of phenolic compounds among eight grape varieties

To more clearly analyze the metabolite accumulation characteristics of *V. adenoclada*, especially the sequenced variety of GH6, K-means clustering analysis was done. The 674 metabolites in the eight varieties were clustered into 12 subclasses based on metabolic profiling differences (Supplementary Figure 5). A total of 42 metabolites in subclasses 8 and 10 exhibited a higher abundance in *V. adenoclada* varieties, hybrid of *V. heyneana* and *V. vinifera* than those in *V. vinifera* varieties (Supplementary Table 15), whereas the content of 31 metabolites in subclasses 11 and 12 were higher in the *V. vinifera* varieties (Supplementary Table 16). Specifically, in subclasses 8 and 10, flavonoids accounted for 38%, followed by phenolic acids at 19% (Figure 4A). Flavonoids also exhibited the highest proportion in subclasses 11 and 12, followed by tannins, and phenolic acids (Figure 4B). Most phenolic acids were more abundant in *V. adenoclada* varieties, except chlorogenic acid (Figure 4C). The cinnamic acid content in YN2 was higher compared with that in other varieties, whereas the ferulic acid and salicylic acid content in GH5 and GH6 were higher, respectively. The resveratrol and piceid content in GH6 were the highest among all varieties. In general, the content of the three flavonols, quercetins, myricetins, and kaempferols in GH4 and GH6 were higher, followed by CS and

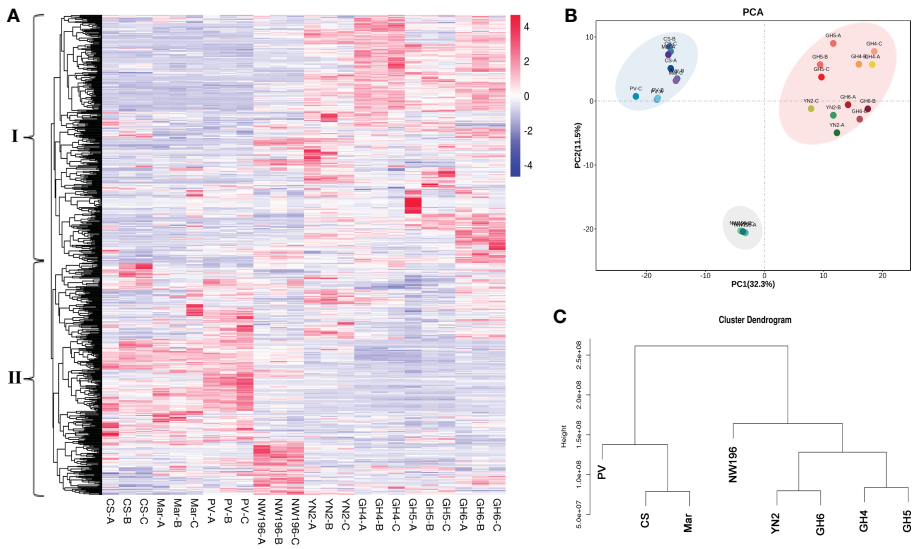


FIGURE 3
Widely targeted metabolomic analysis in eight grape varieties. **(A)** Overview of 674 metabolites in eight grape varieties containing biological duplications. **(B)** Principal component analysis (PCA). **(C)** Cluster dendrogram of metabolome data from eight grape varieties.

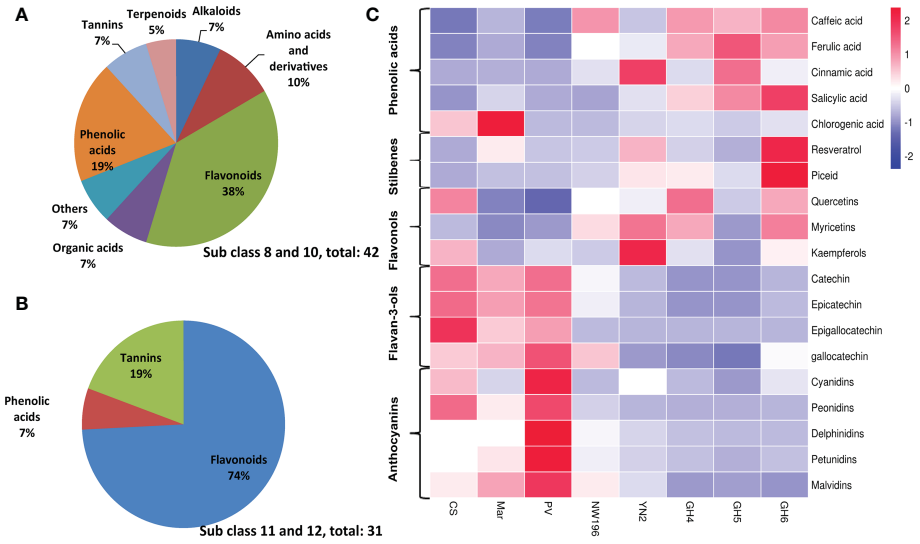


FIGURE 4
Differential accumulation of phenolic compounds in eight grape varieties. **(A)** The proportion of metabolites in subclasses 8 and 10. **(B)** The proportion of metabolites in subclasses 11 and 12. **(C)** The comparison of non-flavonoids and flavonols in different grape varieties.

NW196. For *V. adenoclada* varieties, the stilbene and flavonol content in GH5 was lower compared with that in GH4 and GH6. For the flavan-3-ols and anthocyanins, the content in *V. vinifera* varieties was higher compared with that in the other varieties, and PV was the highest.

3.6 Transcriptome sequencing, clustering, and enrichment analysis of degs in phenolic synthesis pathway

After alignment with the self-testing genome, we obtained 35.77-51.55 million total reads (Supplementary Table 17). The matching rate of these high-quality reads to the reference genome ranged from 84.05% to 92.50%. The GC content of the 24 samples ranged from 45.57% to 46.61%. The Q30 percentage of them was ≥93.46%, suggesting that the sequencing data was reliable and satisfied the threshold for downstream analysis.

In total, 30,257 genes were found to be expressed in 24 grape samples, which contained 1,597 novel genes. In addition, a total of 13,423 DEGs were identified *via* DESeq2 centered on |log2fold-change| ≥1 and a false discovery rate (FDR) < 0.05 in all samples (Supplementary Table 18). Based on the expression pattern among the different grape varieties, these DEGs were divided into two groups (Figure 5A). We used R language pheatmap package to calculate the euclidean distance between two samples by using expression information. The clustering in Figure 5A was achieved by euclidean distance. Group I genes were more abundant in *V. vinifera* varieties, whereas group II genes were more abundant in *V. adenoclada* varieties, a hybrid of *V. heyneana* and *V. vinifera*. The PCA and cluster dendrogram of the transcriptome were basically congruent with the results of the gene expression pattern classification (Figures 5B, C).

We employed RNA-seq to assess variations in gene expression at the transcript level in the berries of eight grape varieties in relation to the phenylpropanoid/flavonoid biosynthesis pathway-related structural genes (Figure 6). The findings suggested that the transcription of genes encoding enzymes upstream of the phenylpropane metabolic pathway, including *PAL* (EC:4.3.1.24), *C4H* (EC:1.14.13.11), and *STS* (EC 2.3.1.74), showed significantly higher expressions in berries of PV, NW196, YN2, and GH6. Expression of the *COMT* (EC: 2.1.1.68) gene in four *V. adenoclada* varieties higher than the other four varieties, whereas most of *4CL* (EC 6.2.1.12) presented significantly or moderately higher expressions in three *V. vinifera* varieties or NW196. Generally, *V. adenoclada* varieties exhibited many similarities when compared with the varieties that have *V. vinifera* consanguinity. However, there were also some differences between the four *V. adenoclada* varieties. Specifically, *CHS* (EC 2.3.1.74), *CHI* (CHI, EC 5.5.1.6), and *F3H* (EC:1.14.11.9) exhibited higher levels of expression in YN2. In general, almost all members of *FLS*, *LAR* (EC 1.17.1.3), *ANR* (EC 1.3.1.77), and *UFGT* (EC 2.4.1.115) presented higher expressions in four *V. adenoclada* varieties compared with the *V. vinifera* varieties. In addition, the gene encoding glutathione S-transferase (GST4, Vad04G007830) showed the highest expression in NW196.

3.7 Co-expression analysis between phenolic biosynthetic pathway genes and transcription factor (TF) genes

To analyze the expression of TFs related to phenolic metabolism, the predicted protein sequences were compared using the plant TF database by hmmscan (Figure 7). A total of 1442 TFs were predicted, and the top four transcription factors were MYB (130), bHLH (121), ERF (104), and C2H2 (91) (Supplementary Table 19). To predict the TFs and genes with high connectivity that regulate key structural genes in the phenolic biosynthetic pathway, a co-expression analysis was done between the structural and TF genes. Five phenolic synthesis genes, including genes encoding PALs, COMT, STSs,

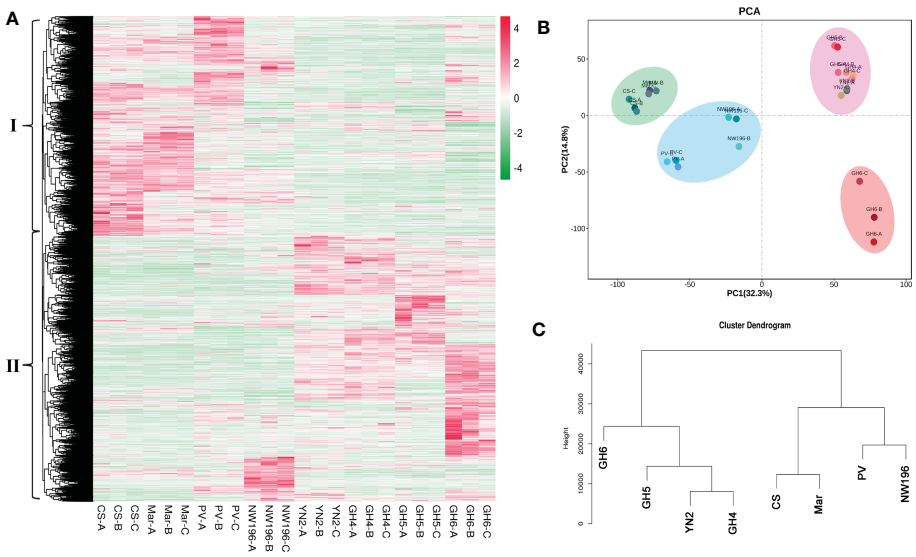


FIGURE 5
Transcriptome data analysis in eight grape varieties. **(A)** Overview of 30,257 genes in eight grape varieties containing biological duplications. **(B)** Principal component analysis (PCA). **(C)** Cluster dendrogram of transcriptome data from grape varieties.

CHSs, and UFGT, were selected as "target genes." The absolute value of the Pearson Correlation Coefficient and *p*-value between them and the expression levels of TF genes using RNA-seq data was calculated. We selected an absolute value for PPC greater than 0.85 and a *p*-value less than 0.01 as conditions. For structural genes with multiple transcripts, we selected the most representative ones through expression pattern clustering for analysis.

Among the identified co-expressed TFs, the most abundant positively correlated TFs were members of the MYB, ERF, bHLH, WRKY, GRAS, NAC, and C2H2 families. For phenylpropane metabolic pathway entry enzymes, four members (Vad08G006000, Vad13G005770, Vad16G000240 and Vad16G000260) of the PAL family were mainly positively regulated by MYB, ERF, bHLH, GRAS, NAC, WRKY and C2H2 families. On the other hand, a total of 218 TFs regulated the five members (Vad16G006240, Vad16G006270, Vad16G006280, Vad16G006300, Vad16G006450) of the STSs, mainly belonging to MYB, ERF, bHLH, WRKY, GRAS and NAC families. Among them, 16 TFs played a negative regulatory role. In addition, STS (Vad16G006300) was in the network with PAL (Vad16G000240), and they were regulated by 102 TFs. Of note, TALE (Vad11G006430) was the only negatively regulated TF among the 102 common TFs. However, STS (Vad16G006450) and PAL (Vad08G006000) were linked by only five TFs: GRAS (Vad11G004310), C3H (Vad12G012060), G2-like (Vad02G009070), MYB (Vad01G000670) and WRKY (Vad13G001870). In addition, GRAS (Vad11G004310) was the only negatively regulated TF. In the other network, PAL (Vad16G000260), STS (Vad16G006270) and CHS (Vad14G009010) had 16 common TFs. COMT (Vad02G002760) alone formed a network as a "target gene" with 48 TFs, and 21 of them exercised negative regulation. Moreover, 24 TFs had regulation effect on UFGT (Vad16G001310), and B3 (Vad06G006930), BBR-BPC (Vad17G001580), bHLH (Vad18G007780, Vad07G014270) presented negative regulation of it. bZIP (Vad06G008410) had negative regulatory effects on UFGT (Vad16G001310) and PAL (Vad13G005770). Interestingly, MYB (Vad07G004350) was the only TF that connected PAL (Vad13G005770), CHS (Vad05G007050) and UFGT (Vad16G001310) in the same network.

3.8 Weighted gene co-expression network analysis (WGCNA) of the DEGs

To gain further insight into the accumulation and variation of metabolites among the grapes, a WGCNA was performed to identify the co-expression networks of DEGs. Co-expression modules are defined as clusters of highly interconnected genes with high correlation coefficients. Genes with similar expression patterns were clustered into 9 distinct modules with gene numbers ranging from 64 (gray) to 6,952 (navajowhite) (Figure 8A). Then, the association between modules and specific phenolic compounds were analyzed (Figure 8B). Notably, the lavenderblush module consisted of DEGs that were significantly (p-value ≤ 0.01) positively correlated with caffeic acid, ferulic acid, and salicylic acid. For navajowhite and tan modules, all of the DEGs were significantly positively correlated with flavan-3-ols (catechin, epicatechin, gallocatechin, and epigallocatechin) and most anthocyanins (peonidins, delphinidins, petunidins, and malvidins). However, the cyan module consisted of DEGs that were significantly positively correlated with stilbenes (resveratrol and piceid) and myricetins. Of note, STSs were mainly found in darkslateblue module, and the correlation between different transcripts and TFs deserves further study in the future.

GO enrichment analysis revealed that the top 20 GO terms in the lavenderblush and navajowhite modules all belonged to biological processes and cellular components. GO: 0009536 (Plastid, 1,708 genes) and GO: 0005623 (cell, 7,143 genes) exhibited the largest number of background genes, with 280 and 2,293 DEGs, respectively (Figures 8C, D). In addition, the top 20 GO terms in the cyan module

10.3389/fpls.2023.1124046

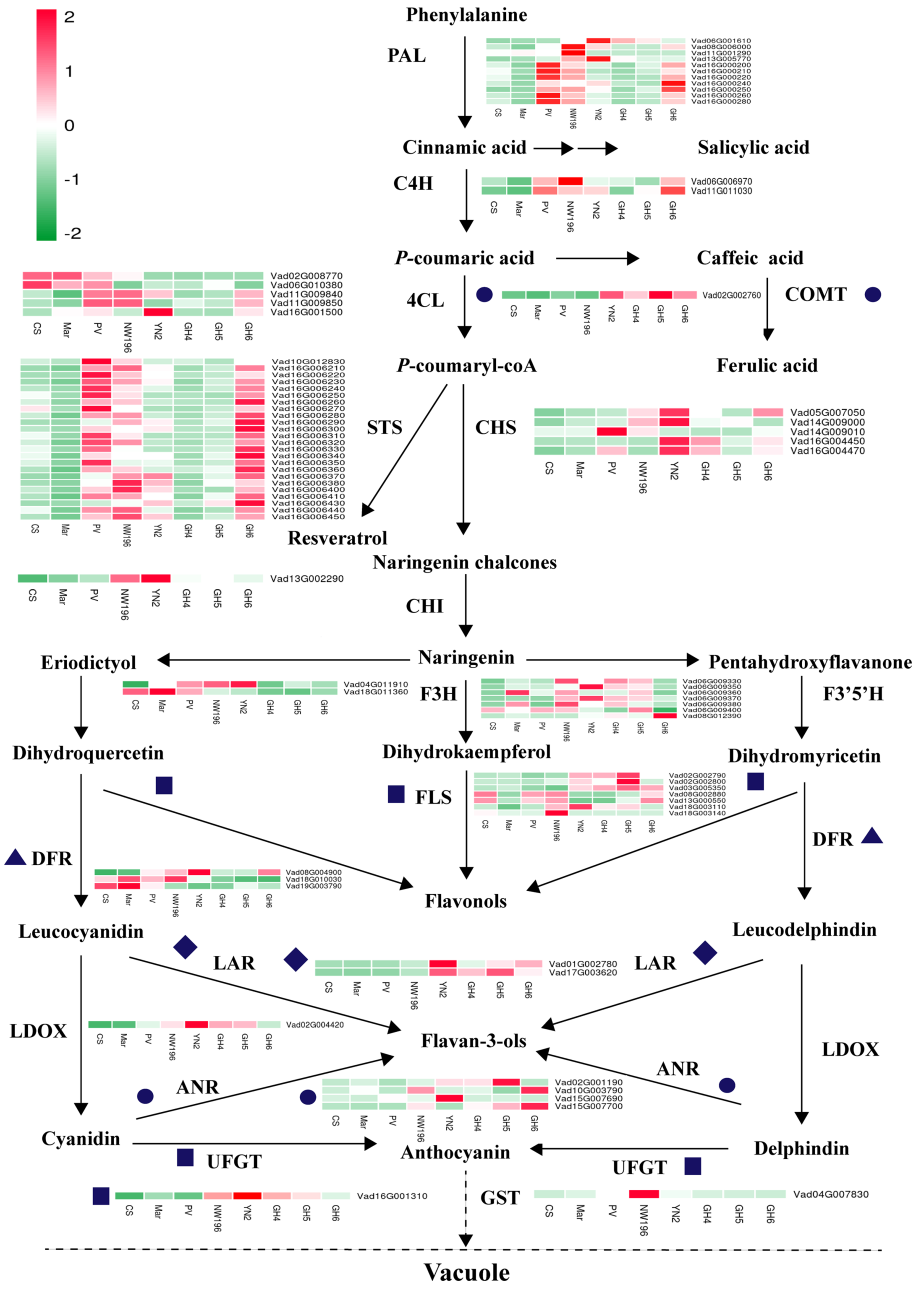


FIGURE 6
Transcriptomic analysis of structural DEGs implicated in phenolic biosynthesis within berries of eight grape varieties. 4CL, 4-coumarate: CoA ligase; ANR, anthocyanidin reductase; C4H, *trans*-cinnamate 4-monooxygenase; CHS, chalcone synthase; CHI, chalcone isomerase; COMT, caffeic acid 3-*O*-methyltransferase; DFR, dihydroflavonol reductase; F3H, flavonoid 3-hydroxylase; F3'5'H, flavonoid 3',5'-hydroxylase; FLS, flavonol synthase; GST, glutathione S-transferase; LAR, leucoanthocyanidin reductase; LDOX, leucoanthocyanidin dioxygenase; UFGT, UDP-glucose: flavonoid 3-*O*-glucosyltransferase; PAL, phenylalanine ammonia-lyase; STS, stilbene synthase. As highlighted in the legend, each square in the heatmap next to its gene names corresponds to the mean FPKM value of the gene in each sample.

all belonged to biological processes, molecular functions, and cellular components. GO: 1901576 (organic substance biosynthetic process, 4,505 genes) exhibited the largest number of background genes with 936 DEGs (Figure 8E).

There were 93 TFs annotated in the lavenderblush module, most of which were bHLH (11) and ERF (10) family members, and most connected with the module were SCL13, BBX19, ERF5, GLK1, ARF19, WRKY28, MYB73, ERF012, BLH9, and MYB5 (Supplementary Table 20). COMT was also located in this module and the TFs with a correlation coefficient >0.8 were: ERF5 (Vad16G002600), ARF19 (Vad11G000440), MYB2 (Vad11G001070), BLH9 (Vad08G008660), WRKY28 (Vad12G000470), GLK1 (Vad12G002660), MYB73-1 (Vad18G008660), and MYB73-2 (Vad03G006650) (Supplementary Table 21). These TFs may be involved in the regulation of ferulic acid biosynthesis.

There were 315 TFs annotated in the cyan module, most of which were bHLH (26) and ERF (27) family members, and most connected with the module were WRKY65, WRKY21, bHLH130, ASIL2, COL10, ARF9, STOP1, ERF118, PIL15, and SCL14 (Supplementary

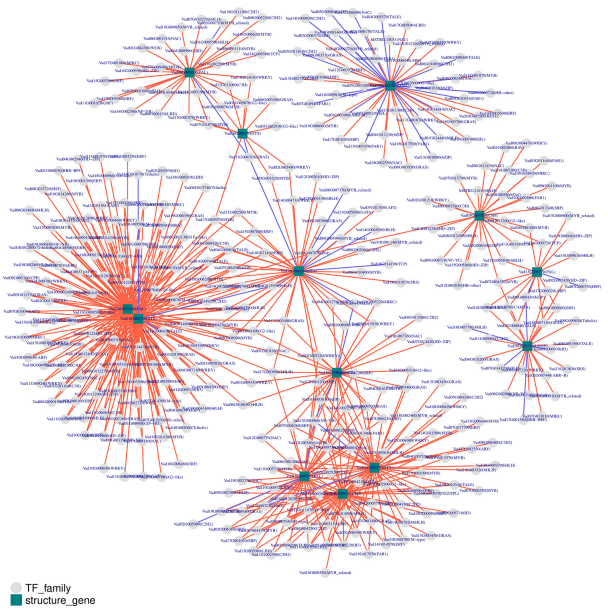


FIGURE 7
Co-expression analysis of key enzymes located in phenolic biosynthesis and transcription factors (TFs) in berries from eight grape varieties. Key enzymes are shown in rectangles and TFs in circles. CHS, chalcone synthase; COMT, caffeic acid 3-*O*-methyltransferase; PAL, phenylalanine ammonia-lyase; STS, Stilbene synthase; UFGT, UDP-glucose: flavonoid 3-*O*-glucosyltransferase. Positive and negative correlations are provided by red arrows and purple lines, respectively.

Table 22). FLS (Vad18G003110) was located in the cyan module, and the TFs with a correlation coefficient >0.8 were: ARF2A (Vad17G000320), BLH7 (Vad03G000040), HOX16 (Vad14G012380), LRP1 (Vad06G009660), and MYB12 (Vad07G004350). UFGT (Vad16G001310) was also located in the cyan module and the TFs with a correlation coefficient >0.8 were: ZFP2 (Vad11G009860), WRKY44 (Vad08G007220), MYB12 (Vad07G004350), ARF2A (Vad17G000320), HOX22 (Vad02G002420), and LRP1 (Vad06G009660) (Supplementary Table 23). Therefore, these TFs may play a regulatory role in the synthesis of flavonols and anthocyanins.

4 Discussion

Since the publication of the first grape genome (PN40024), studies on grapes have made a qualitative leap at the molecular level (Jaillon et al., 2007). In recent years, whole-genome sequencing or resequencing of wild *Vitis* species from East Asian populations has been reported, especially those that originate from China (Liang et al., 2019; Wang et al., 2021). This has provided rich data for illuminating the evolutionary biology of the *Vitis* species and has achieved a more accurate comparison with different species/varieties at the genomic level. However, little is known regarding the metabolism and accumulation of the East Asian grapevine using a self-testing genome.

We provide a high-quality, chromosome-level genome assembly of GH6, a *V. adenoclada* varieties, based on a collection of Illumina and ONT sequence data followed by Hi-C mapping. The resultant genome size of GH6 was 498.27 Mb with a lower level of gaps (0.059 Mb) compared with PN40024 (~15 Mb) (GenBank assembly accession: GCA_000003745.2), Trayshed (~1 Mb) (Cochetel et al., 2021), Shanputao (~4 Mb) (Wang et al., 2021), Riparia Gloire de Montpellier (~6 Mb) (GenBank assembly accession: GCA_004353265.1). A total of 481.44 Mb sequences were anchored to 19 chromosomes of GH6 *via* the Hi-C assembly, accounting for 96.62% of the final genome assembly. This rate was higher compared with that of PN40024 (87.7%) (GenBank assembly accession: GCA_000003745.2), Shanputao (82.6%) (Wang et al., 2021), Trayshed (92.5%) (Cochetel et al., 2021) and Riparia Gloire de Montpellier (94.2%) (GenBank assembly accession: GCA_004353265.1). The high-quality and chromosome-level genome of GH6 that was deciphered will be helpful for the utilization of the East Asian wild germplasm resources of *Vitis* spp. for future grape genetic improvement and evolutionary studies. We also found that quite a number of gene family expansions in the genome of GH6 associated with defense response including RPP13-like proteins, TMV resistance proteins, RPM1s, STSs, and so on, which suggests that the expanded gene families of GH6 may contribute to resistance against diseases.

Chinese wild grapes generally have the advantages of high yield, strong resistance to stress, and a rich content of phenolics. The most studied and reported species include *V. amurensis*, *Vitis pseudoreticulata*, *V. heyneana* and *V. davidii* (Ma et al., 2018; Ju et al., 2020; Wang et al., 2021; Yan et al., 2021). In the current work, phenols are crucial in differentiating species and varieties using a metabolomics approach. Most phenolic acids were found in higher concentrations in the fruits of *V. adenoclada* and the hybrid with East Asian lineage than in *V. vinifera*. Phenolic acids are non-flavonoid phenolic compounds synthesized by the phenylpropane metabolic pathway (Garrido and Borges, 2013). Here, we determined that the content of caffeic acid, cinnamic acid, ferulic acid, and salicylic acid in *V. adenoclada* berries were higher compared with those of *V. vinifera*.

Resveratrol plays an important role in resistance to biotic and abiotic stresses, especially in pathogen resistance (Yan et al., 2021). In the current investigation, among all varieties, we observed that GH6 has the highest concentration of resveratrol and its glycosides (piceid). It should be noted that wine grapes cultivated in the south subtropical region are more likely to suffer from the threat of pathogens owing to hot and humid condition, which has a devastating impact on the quality and yield (Bai et al., 2008; Liu K. Y., 2012). Therefore, effective utilization of the *V. adenoclada* grape will facilitate the breeding of new grapevine cultivars during the breeding process.

Flavonoids in grapes mainly include flavonols, flavanols, and anthocyanins (Wu et al., 2020). In this study, the flavonol content in GH4 and GH6 was the highest among all varieties. However, the flavan-3-ols and anthocyanins of *V. vinifera* varieties were most abundant, and NW196 contained higher levels than *V. adenoclada* varieties. The previous study demonstrated that the content of flavonoids in *V. heyneana* and *V. davidii* was higher compared with that of CS, whereas *V. adenoclada* was similar (Jiang, 2016). From the trend analysis, the higher content of metabolites in European or East Asian grape species were associated with flavonoids. Therefore, it is necessary to further explore the phenolics synthesis in different grape varieties by RNA-seq.

Currently, studies on the characteristics and regulation of grape fruit quality revealed by metabolomics and transcriptomics have been

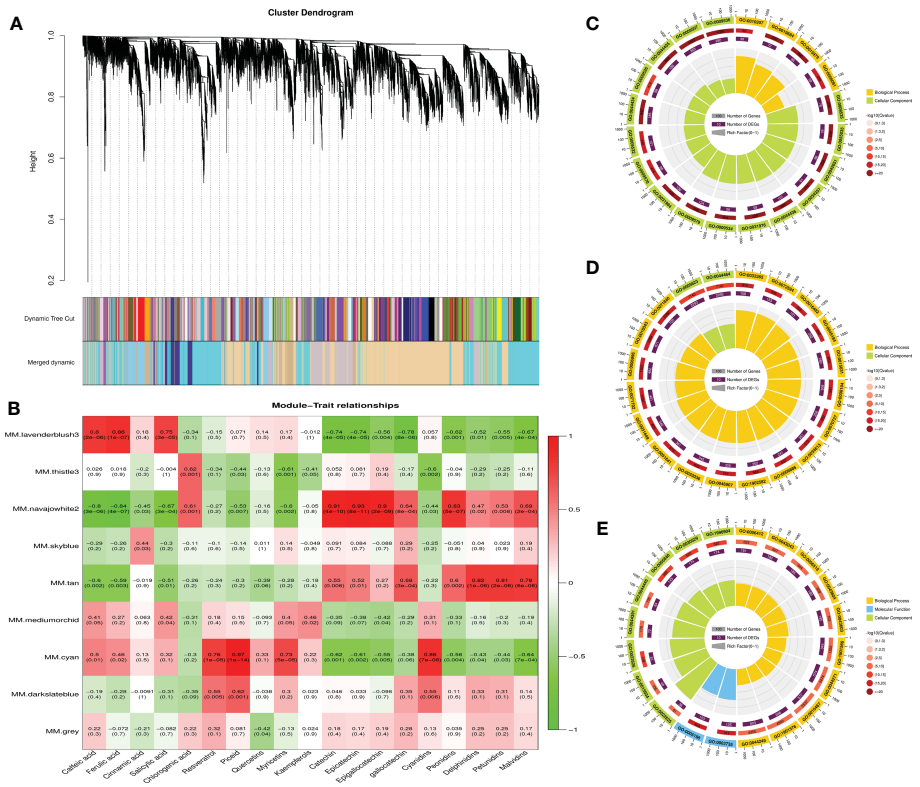


FIGURE 8
Weighted gene co-expression network analysis (WGCNA) of DEGs among different grape varieties and GO enrichment analyses of DEGs associated with the accumulation of phenolic compounds. **(A)** Cluster dendrogram showing co-expression modules identified by WGCNA. The major tree branches constitute 9 modules labeled by different colors. **(B)** Module–trait relationships. Each row corresponds to a module. Each column corresponds to a group of phenolic compounds. The color and data of each cell indicates the correlation coefficient between the module and the phenolic compound. A high degree of correlation between a specific module and the phenolic compound is indicated by red (positive) or green (negative). Red to green represents a positive to negative correlation between them. **(C)** GO enrichment analysis of DEGs in the lavenderblush module. **(D)** GO enrichment analysis of DEGs in the navajowhite module. **(E)** GO enrichment analysis of DEGs in the cyan module.

primarily focused on *V. vinifera* (Sun et al., 2019; Yang et al., 2020). For unpublished species and variety genomes, most transcriptome studies have selected the genome assembly of PN40024 as a reference genome (Cheng et al., 2019; Zhang et al., 2020). In the present study, the mapped rates of *V. adenoclada* varieties using a self-testing genome increased 0.86%–1.46% compared with that of the PN40024 genome (Supplementary Table 24). In contrast, the mapped rates of *V. vinifera* varieties and NW196 using the genome of GH6 decreased by 6.35%–8.84%. This indicates that selecting genomes with a closer genetic background as a reference can reduce error and improve the accuracy of the comparison.

In the present study, the *V. adenoclada* varieties were clearly distinguished from the *V. vinifera* varieties and NW196. There were significant variations in the expression of 15 structural genes in the phenylpropane metabolic pathway across the eight varieties, ranging from early *PAL* to late *GST*. PAL as an entry enzyme in the phenylpropane pathway (Sun et al., 2019). In this study, most PAL members were highly expressed in PV, whereas they were highly expressed in GH6 among the *V. adenoclada* varieties. In addition, the *C4H* and *COMT* were closely related to phenolic acid synthesis, whereas 4CL catalyzes *p*-coumaric acid and continues to lead the flavonoid synthesis pathway (Cheng et al., 2019). The expression of C4H was higher in PV, NW196, and GH6, whereas those of COMT and 4CL were higher in *V. adenoclada* and *V. vinifera* varieties, respectively. This could explain why the phenolic acid concentration of *V. adenoclada* berries was higher than that of *V. vinifera* berries. STS is a key enzyme involved in the biosynthesis of resveratrol and has been linked to plant resistance to fungal diseases (Shi et al., 2014). In the current study, STSs exhibited higher expression level in GH6, which may explain the high content of resveratrol and piceid in the results of the widely-targeted metabolome. From our previous analysis, the content of flavanols and anthocyanins in PV was the highest among the eight grape varieties. However, *CHS*, *LAR*, *ANR*, *UFGT*, and *GST4* were generally highly expressed in *V. adenoclada* varieties or NW196. These results indicate that the synthesis of flavonoids was more affected by the upstream genes, *PAL* or *4CL*. The previous study regarding the anthocyanins in *V. vinifera* and the hybrid of *V. vinifera* and *V. amurensis* arrived at similar estimates (Liu et al., 2021).

To better understand the regulatory process of phenolic metabolism in different grape varieties, we performed a TF prediction and genome-wide co-expression network analysis. Previous studies on the regulation of transcription factors on flavonoid pathway genes have primarily focused on *V. vinifera*, but there have been few reports on the wild resources of the East Asian species (Pilati et al., 2007; Terrier et al., 2009; Sun et al., 2019; Cheng et al., 2021). In the present study, a total of 1,442 transcription factors were predicted and the MYB family was the most abundant. To date, multiple TFs belonging to the MYB, WRKY, ERF, and bHLH families have been shown to regulate flavonoid biosynthesis,

and the MYB family is the most well-studied (Wei et al., 2021). Many key MYB TFs, including MYBA1, MYBA2, MYB5A, MYB5B, MYBPA1 MYBPA2, MYB4, and MYB86 have been identified that promote or inhibit flavonoid biosynthesis in the grapevine (Sun et al., 2019; Cheng et al., 2021). From our study, these TFs not only regulate flavonoid synthesis, but also play an important role in the synthesis of non-flavonoid phenols. Co-expression analysis revealed that the PALs and STSs family members were positively regulated by MYB, ERF, bHLH, GRAS, NAC and WRKY family TFs. Previous studies have demonstrated that *MYBF1* (also known as *MYB12*) was specifically responsible for flavonol biosynthesis in grapes (Sun et al., 2019). In our study, MYB12 (Vad07G004350) positively regulated PAL (Vad13G005770), CHS (Vad05G007050), FLS (Vad18G003110) and UFGT (Vad16G001310). WGCNA analysis further indicated that COMT is highly correlated with the MYB family genes, such as MYB2 (Vad11G001070), MYB73-1 (Vad18G008660), and MYB73-2 (Vad03G006650). Presumably, these TFs are involved in the regulation of ferulic acid biosynthesis. Target genes of *MYBPA1*, *MYBPA2*, and *MYBPAR* in the grape are *LAR* and *ANR*, which regulate the synthesis of proanthocyanidins (Bogs et al., 2007; Terrier et al., 2009; Koyama et al., 2014). In the present study, MYB5b (Vad06G000560) and MYBPAR (Vad11G001070) were highly expressed in four *V. adenoclada* varieties, which corresponded with the higher expressions of LAR and ANR in these varieties. However, MYBPA1 (Vad15G006940) was highly expressed in CS and Mar. In recent years, MYB4 and MYBC2-L1 have been shown to inhibit proanthocyanidin and anthocyanin synthesis (Cavallini et al., 2015; Pérez-Díaz et al., 2016; Zhu et al., 2019). From the results of our study, MYB4 (Vad05G005030) and MYBC2-L1 (Vad01G004530) were highly expressed in YN2 and GH6. This may inhibit the synthesis of flavan-3-ols and anthocyanins in both varieties to some extent. From the present study, transcriptome analysis with self-testing genome will provide insight into the synthesis and regulation of phenols in *V. adenoclada* grapes.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://ngdc.cncb.ac.cn/, PRJCA009537, https://ngdc.cncb.ac.cn/, PRJCA009575.

Author contributions

GC, DW, RG and SZ conceptualized the study and contributed to original draft and funding acquisition. GC and SZ contributed to reviewing and editing of the manuscript. HL, RW, HY and LL participated in experimental processing and collecting material. GC, DW, ZW, XM and NY contributed to resources. GC, JZ, LX and SZ analyzed the data. SZ supervised the project. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Natural Science Foundation of China (31860529), the Natural Science Foundation of Guangxi Province (2021GXNSFAA196016), the Guangxi Key Research and Development Program (GuikeAB18294032, GuikeAB18126005, GuikeAB18294003), the Project for Science and Technology Development Fund of Guangxi Academy of Agricultural Sciences (Guinongke2021JM26, Guinongke2021JM27), and the Guangxi Luocheng Maoputao Experimental Station (GuiTS201418).

Acknowledgments

We thank Dr. Jianfu Jiang at Zhengzhou Fruit Research Institute in Chinese Academy of Agricultural Sciences for his support of classification and identification of *Vitis* spp., Dr. Yuan Liu at Xishuangbanna Tropical Botanical Garden in Chinese Academy of Sciences for her support of grape specimens drawing, and Dr. Muming Cao at Grape and Wine Research Institute in Guangxi Academy of Agricultural Sciences for her support of NW196 cultivation.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1124046/full#supplementary-material

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Bai, X. J., Li, Y. R., Huang, J. L., Liu, J. B., Peng, H. X., Xie, T. L., et al. (2008). One-year-two-harvest cultural technique system for kyoho grape in southern region of guangxi. *Southwest. China J. Agric. Sci.* 21, 953–955. doi: 10.16213/j.cnki.scjas.2008.04.041

Bao, Z., and Eddy, S. R. (2002). Automated *de novo* identification of repeat sequence families in sequenced genomes. *Genome Res.* 12, 1269–1276. doi: 10.1101/gr.88502

Blanco, E., Parra, G., and Guigó, R. (2007). Using geneid to identify genes. *Curr. Protoc. Bioinf.* 18, 4.3.1–4.3.28. doi: 10.1002/0471250953.bi0403s18

Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M. C., Estreicher, A., Gasteiger, E., et al. (2003). The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res.* 31, 365–370. doi: 10.1093/nar/gkg095

Bogs, J., Jaffe, F. W., Takos, A. M., Walker, A. R., and Robinson, S. P. (2007). The grapevine transcription factor VvMYBPA1 regulates proanthocyanidin synthesis during fruit development. *Plant Physiol.* 143, 1347–1361. doi: 10.1104/pp.106.093203

Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59–60. doi: 10.1038/nmeth.3176

Burge, C., and Karlin, S. (1997). Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* 268, 78–94. doi: 10.1006/jmbi.1997.0951

Burton, J. N., Adey, A., Patwardhan, R. P., Qiu, R., Kitzman, J. O., and Shendure, J. (2013). Chromosome-scale scaffolding of *de novo* genome assemblies based on chromatin interactions. *Nat. Biotechnol.* 31, 1119–1125. doi: 10.1038/nbt.2727

Campbell, M. A., Haas, B. J., Hamilton, J. P., Mount, S. M., and Buell, C. R. (2006). Comprehensive analysis of alternative splicing in rice and comparative analyses with arabidopsis. *BMC Genomics* 7, 327. doi: 10.1186/1471-2164-7-327

Cavallini, E., Matus, J. T., Finezzo, L., Zenoni, S., Loyola, R., Guzzo, F., et al. (2015). The phenylpropanoid pathway is controlled at different branches by a set of R2R3-MYB C2 repressors in grapevine. *Plant Physiol.* 167, 1448–1470. doi: 10.1104/pp.114.256172

Chen, D. K., Deng, T. J., Yao, J. Y., Li, D. B., Deng, B., Wang, X. M., et al. (2020). *Germplasm resources of crops in guangxi-fruit tree volume. 1st ed.* (Beijing, China: Science Press).

Cheng, J., Yu, K., Shi, Y., Wang, J., and Duan, C. (2021). Transcription factor VviMYB86 oppositely regulates proanthocyanidin and anthocyanin biosynthesis in grape berries. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.613677

Cheng, G., Zhou, S.-H., Wen, R.-D., Xie, T.-L., Huang, Y., Yang, Y., et al. (2018). Anthocyanin characteristics of wines in *Vitis* germplasms cultivated in southern China. *Food Sci. Technol.* 38, 513–521. doi: 10.1590/1678-457x.37516

Cheng, G., Zhou, S., Zhang, J., Huang, X., Bai, X., Xie, T., et al. (2019). Comparison of transcriptional expression patterns of phenols and carotenoids in 'Kyoho' grapes under a two-crop-a-year cultivation system. *PloS One* 14, e0210322. doi: 10.1371/journal.pone.0210322

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. doi: 10.1093/bioinformatics/bty560

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

Cochetel, N., Minio, A., Massonnet, M., Vondras, A. M., Figueroa-Balderas, R., and Cantu, D. (2021). Diploid chromosome-scale assembly of the *Muscadinia rotundifolia* genome supports chromosome fusion and disease resistance gene expansion during *Vitis* and *Muscadinia* divergence. *G3* 11, jkab033. doi: 10.1093/g3journal/jkab033

Conesa, A., Gotz, S., Garcia-Gomez, J. M., Terol, J., Talon, M., and Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674–3676. doi: 10.1093/bioinformatics/bti610

Di Genova, A., Almeida, A. M., Muñoz-Espinoza, C., Vizoso, P., Travisany, D., Moraga, C., et al. (2014). Whole genome comparison between table and wine grapes reveals a comprehensive catalog of structural variants. *BMC Plant Biol.* 14, 7. doi: 10.1186/1471-2229-14-7

Dimmer, E. C., Huntley, R. P., Alam-Faruque, Y., Sawford, T., O'Donovan, C., Martin, M. J., et al. (2012). The UniProt-GO annotation database in 2011. *Nucleic Acids Res.* 40, D565–D570. doi: 10.1093/nar/gkr1048

Ellinghaus, D., Kurtz, S., and Willhoeft, U. (2008). LTRharvest, an efficient and flexible software for *de novo* detection of LTR retrotransposons. *BMC Bioinf.* 9, 18. doi: 10.1186/1471-2105-9-18

Emms, D. M., and Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* 20, 238. doi: 10.1186/s13059-019-1832-y

Flynn, J. M., Hubley, R., Goubert, C., Rosen, J., Clark, A. G., Feschotte, C., et al. (2020). RepeatModeler2 for automated genomic discovery of transposable element families. *Proc. Natl. Acad. Sci. U.S.A.* 117, 9451–9457. doi: 10.1073/pnas.1921046117

Fu, M., Yang, X., Zheng, J., Wang, L., Yang, X., Tu, Y., et al. (2021). Unraveling the regulatory mechanism of color diversity in *Camellia japonica* petals by integrative transcriptome and metabolome analysis. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.685136

Garrido, J., and Borges, F. (2013). Wine and grape polyphenols — a chemical perspective. *Food Res. Int.* 54, 1844–1858. doi: 10.1016/j.foodres.2013.08.002

Haas, B. J., Salzberg, S. L., Zhu, W., Pertea, M., Allen, J. E., Orvis, J., et al. (2008). Automated eukaryotic gene structure annotation using EVidenceModeler and the program to assemble spliced alignments. *Genome Biol.* 9, R7. doi: 10.1186/gb-2008-9-1-r7

Han, M. V., Thomas, G. W., Lugo-Martinez, J., and Hahn, M. W. (2013). Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. *Mol. Biol. Evol.* 30, 1987–1997. doi: 10.1093/molbev/mst100

Hasegawa, M., Kishino, H., and Yano, T.-a. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* 22, 160–174. doi: 10.1007/BF02101694

Jaillon, O., Aury, J. M., Noel, B., Policriti, A., Clepet, C., Casagrande, A., et al. (2007). The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* 449, 463–467. doi: 10.1038/nature06148

Jiang, Y. (2016). *The quality evaluation and main components of Chinese wild grape berry* (Yangling, China: Northwest A & F University).

Jiao, Y., Wickett, N. J., 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

Ju, Y. L., Yue, X. F., Cao, X. Y., and Fang, Y. L. (2020). Targeted metabolomic and transcript level analysis reveals quality characteristic of Chinese wild grapes (*Vitis davidii* foex). *Foods* 9, 1387. doi: 10.3390/foods9101387

Kanehisa, M., and Goto, S. (2000). KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30. doi: 10.1093/nar/28.1.27

Keilwagen, J., Hartung, F., Paulini, M., Twardziok, S. O., and Grau, J. (2018). Combining RNA-seq data and homology-based gene prediction for plants, animals and fungi. *BMC Bioinf.* 19, 189. doi: 10.1186/s12859-018-2203-5

Keilwagen, J., Wenk, M., Erickson, J. L., Schattat, M. H., Grau, J., and Hartung, F. (2016). Using intron position conservation for homology-based gene prediction. *Nucleic Acids Res.* 44, e89. doi: 10.1093/nar/gkw092

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317

Kong, Q. (2004). *Chinese Ampelography. 1st ed.* (Beijing, China: Chinese Agricultural Science and Technology Press).

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., and Phillippy, A. M. (2017). Canu: scalable and accurate long-read assembly *via* adaptive k-mer weighting and repeat separation. *Genome Res.* 27, 722–736. doi: 10.1101/gr.215087.116

Korf, I. (2004). Gene finding in novel genomes. *BMC Bioinf.* 5, 59. doi: 10.1186/1471-2105-5-59

Koyama, K., Numata, M., Nakajima, I., Goto-Yamamoto, N., Matsumura, H., and Tanaka, N. (2014). Functional characterization of a new grapevine MYB transcription factor and regulation of proanthocyanidin biosynthesis in grapes. *J. Exp. Bot.* 65, 4433–4449. doi: 10.1093/jxb/eru213

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

Langfelder, P., and Horvath, S. (2008). WGCNA: an r package for weighted correlation network analysis. *BMC Bioinf.* 9, 559. doi: 10.1186/1471-2105-9-559

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Liang, Z., Duan, S., Sheng, J., Zhu, S., Ni, X., Shao, J., et al. (2019). Whole-genome resequencing of 472 *Vitis* accessions for grapevine diversity and demographic history analyses. *Nat. Commun.* 10, 1190. doi: 10.1038/s41467-019-09135-8

Li, B., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. *BMC Bioinf.* 12, 323. doi: 10.1186/1471-2105-12-323

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

Liu, C. H. (2012). *Studies on taxonomy and geographical distribution of Chinese wild grape species* (Zhengzhou, China: Henan Agricultural University).

Liu, K. Y. (2012). *Identification, evaluation and genetic diversity of vitis adenoclada hand.-mazz* (Changsha, China: Hunan Agricultural University).

Liu, S. F., Liu, Z. Y., Li, P., Wu, L., Huang, H. R., and Shan, S. M. (2021). Comparison of the content of phenolic compounds and the expression levels of their biosynthesis related genes in ripening stage of *Vitis vinifera* × *Vitis amurensis* red grape. *J. Hunan. Agric. Univ. (Natural. Sciences).* 47, 30–34. doi: 10.13331/j.cnki.jhau.2021.01.005

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. doi: 10.1186/s13059-014-0550-8

Majoros, W. H., Pertea, M., and Salzberg, S. L. (2004). TigrScan and GlimmerHMM: two open source ab initio eukaryotic gene-finders. *Bioinformatics* 20, 2878–2879. doi: 10.1093/bioinformatics/bth315

Marcais, G., Delcher, A. L., Phillippy, A. M., Coston, R., Salzberg, S. L., and Zimin, A. (2018). MUMmer4: A fast and versatile genome alignment system. *PloS Comput. Biol.* 14, e1005944. doi: 10.1371/journal.pcbi.1005944

Marchler-Bauer, A., Lu, S., Anderson, J. B., Chitsaz, F., Derbyshire, M. K., DeWeese-Scott, C., et al. (2011). CDD: a conserved domain database for the functional annotation of proteins. *Nucleic Acids Res.* 39, D225–D229. doi: 10.1093/nar/gkq1189

Massonnet, M., Cochetel, N., Minio, A., Vondras, A. M., Lin, J., Muyle, A., et al. (2020). The genetic basis of sex determination in grapes. *Nat. Commun.* 11, 2902. doi: 10.1038/s41467-020-16700-z

Ma, F., Wang, L., and Wang, Y. (2018). Ectopic expression of *VpSTS29*, a stilbene synthase gene from *Vitis pseudoreticulata*, indicates STS presence in cytosolic oil bodies. *Planta* 248, 89–103. doi: 10.1007/s00425-018-2883-0

Mi, H., Muruganujan, A., Ebert, D., Huang, X., and Thomas, P. D. (2019). PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res.* 47, D419–D426. doi: 10.1093/nar/gky1038

Nguyen, L. T., Schmidt, H. A., von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300

Niu, L. X., and He, P. C. (1996). Systematic taxonomic study on wild grapevine in China. *Acta Hortic. Sin.* 23, 209–212.

Ou, S., and Jiang, N. (2018). LTR_retriever: A highly accurate and sensitive program for identification of long terminal repeat retrotransposons. *Plant Physiol.* 176, 1410–1422. doi: 10.1104/pp.17.01310

Park, M., Vera, D., Kambrianda, D., Gajjar, P., Cadle-Davidson, L., Tsolova, V., et al. (2022). Chromosome-level genome sequence assembly and genome-wide association study of *Muscadinia rotundifolia* reveal the genetics of 12 berry-related traits. *Hortic. Res.* 9, uhab011. doi: 10.1093/hr/uhab011

Parra, G., Bradnam, K., and Korf, I. (2007). CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics* 23, 1061–1067. doi: 10.1093/bioinformatics/btm071

Patel, S., Robben, M., Fennell, A., Londo, J. P., Alahakoon, D., Villegas-Diaz, R., et al. (2020). Draft genome of the native American cold hardy grapevine *Vitis riparia* michx. 'Manitoba 37'. *Hortic. Res.* 7, 92. doi: 10.1038/s41438-020-0316-2

Pérez-Díaz, J. R., Pérez-Díaz, J., Madrid-Espinoza, J., Gonzalez-Villanueva, E., Moreno, Y., and Ruiz-Lara, S. (2016). New member of the R2R3-MYB transcription factors family in grapevine suppresses the anthocyanin accumulation in the flowers of transgenic tobacco. *Plant Mol. Biol.* 90, 63–76. doi: 10.1007/s11103-015-0394-y

Pertea, M., Kim, D., Pertea, G. M., Leek, J. T., and Salzberg, S. L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and ballgown. *Nat. Protoc.* 11, 1650–1667. doi: 10.1038/nprot.2016.095

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295. doi: 10.1038/nbt.3122

Pilati, S., Perazzolli, M., Malossini, A., Cestaro, A., Dematte, L., Fontana, P., et al. (2007). Genome-wide transcriptional analysis of grapevine berry ripening reveals a set of genes similarly modulated during three seasons and the occurrence of an oxidative burst at veraison. *BMC Genomics* 8, 428. doi: 10.1186/1471-2164-8-428

Price, A. L., Jones, N. C., and Pevzner, P. A. (2005). *De novo* identification of repeat families in large genomes. *Bioinformatics* 21, i351–i358. doi: 10.1093/bioinformatics/bti1018

Puttick, M. N. (2019). MCMCtreeR: functions to prepare MCMCtree analyses and visualize posterior ages on trees. *Bioinformatics* 35, 5321–5322. doi: 10.1093/bioinformatics/btz554

Rao, S. S., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., et al. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159, 1665–1680. doi: 10.1016/j.cell.2014.11.021

Servant, N., Varoquaux, N., Lajoie, B. R., Viara, E., Chen, C. J., Vert, J. P., et al. (2015). HiC-pro: an optimized and flexible pipeline for Hi-c data processing. *Genome Biol.* 16, 259. doi: 10.1186/s13059-015-0831-x

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, J., He, M., Cao, J., Wang, H., Ding, J., Jiao, Y., et al. (2014). The comparative analysis of the potential relationship between resveratrol and stilbene synthase gene family in the development stages of grapes (*Vitis quinquangularis* and *Vitis vinifera*). *Plant Physiol. Biochem.* 74, 24–32. doi: 10.1016/j.plaphy.2013.10.021

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212. doi: 10.1093/bioinformatics/btv351

Stanke, M., and Waack, S. (2003). Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics* 19, ii215–ii225. doi: 10.1093/bioinformatics/btg1080

Sun, R. Z., Cheng, G., Li, Q., He, Y. N., Wang, Y., Lan, Y. B., et al. (2017). Light-induced variation in phenolic compounds in cabernet sauvignon grapes (*Vitis vinifera* l.) involves extensive transcriptome reprogramming of biosynthetic enzymes, transcription factors, and phytohormonal regulators. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.00547

Sun, R. Z., Cheng, G., Li, Q., Zhu, Y. R., Zhang, X., Wang, Y., et al. (2019). Comparative physiological, metabolomic, and transcriptomic analyses reveal developmental stage-dependent effects of cluster bagging on phenolic metabolism in Cabernet sauvignon grape berries. *BMC Plant Biol.* 19, 583. doi: 10.1186/s12870-019-2186-z

Tang, S., Lomsadze, A., and Borodovsky, M. (2015). Identification of protein coding regions in RNA transcripts. *Nucleic Acids Res.* 43, e78. doi: 10.1093/nar/gkv227

Tarailo-Graovac, M., and Chen, N. (2009). Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr. Protoc. Bioinf.* 25, 4.10.1–4.10.14. doi: 10.1002/0471250953.bi0410s25

Tatusov, R. L., Natale, D. A., Garkavtsev, I. V., Tatusova, T. A., Shankavaram, U. T., Rao, B. S., et al. (2001). The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res.* 29, 22–28. doi: 10.1093/nar/29.1.22

Terrier, N., Torregrosa, L., Ageorges, A., Vialet, S., Verries, C., Cheynier, V., et al. (2009). Ectopic expression of *VvMybPA2* promotes proanthocyanidin biosynthesis in grapevine and suggests additional targets in the pathway. *Plant Physiol.* 149, 1028–1041. doi: 10.1104/pp.108.131862

Vaser, R., Sovic, I., Nagarajan, N., and Sikic, M. (2017). Fast and accurate *de novo* genome assembly from long uncorrected reads. *Genome Res.* 27, 737–746. doi: 10.1101/gr.214270.116

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., et al. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PloS One* 9, e112963. doi: 10.1371/journal.pone.0112963

Wang, W. T. (1988). Notulae de vitaceis guangxiensibus. *Guihaia* 8, 109–119.

Wang, F., Huang, Y., Wu, W., Zhu, C., Zhang, R., Chen, J., et al. (2020). Metabolomics analysis of the peels of different colored citrus fruits (*Citrus reticulata* cv. 'Shatangju') during the maturation period based on UHPLC-QQQ-MS. *Molecules* 25, 396. doi: 10.3390/molecules25020396

Wang, Y., Xin, H., Fan, P., Zhang, J., Liu, Y., Dong, Y., et al. (2021). The genome of shanputao (*Vitis amurensis*) provides a new insight into cold tolerance of grapevine. *Plant J.* 105, 1495–1506. doi: 10.1111/tpj.15127

Wei, X., Ju, Y., Ma, T., Zhang, J., Fang, Y., and Sun, X. (2021). New perspectives on the biosynthesis, transportation, astringency perception and detection methods of grape proanthocyanidins. *Crit. Rev. Food Sci. Nutr.* 61, 2372–2398. doi: 10.1080/10408398.2020.1777527

Wheeler, T. J., Clements, J., Eddy, S. R., Hubley, R., Jones, T. A., Jurka, J., et al. (2013). Dfam: a database of repetitive DNA based on profile hidden Markov models. *Nucleic Acids Res.* 41, D70–D82. doi: 10.1093/nar/gks1265

Wu, D.-D., Cheng, G., Li, H.-Y., Zhou, S.-H., Yao, N., Zhang, J., et al. (2020). The cultivation techniques and quality characteristics of a new germplasm of *Vitis adenoclada* hand.-mazz grape. *Agronomy* 10, 1851. doi: 10.3390/agronomy10121851

Xie, W., Chen, F., and Chen, Z. (2021). New materials of *Vitis* l. @ in zhejiang. *Guihaia* 41, 1391–1400. doi: 10.11931/guihaia.gxzw201909046

Xu, Z., and Wang, H. (2007). LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res.* 35, W265–W268. doi: 10.1093/nar/gkm286

Yang, Z. (1997). PAML: a program package for phylogenetic analysis by maximum likelihood. *Bioinformatics* 13, 555–556. doi: 10.1093/bioinformatics/13.5.555

Yang, B., He, S., Liu, Y., Liu, B., Ju, Y., Kang, D., et al. (2020). Transcriptomics integrated with metabolomics reveals the effect of regulated deficit irrigation on anthocyanin biosynthesis in Cabernet sauvignon grape berries. *Food Chem.* 314, 126170. doi: 10.1016/j.foodchem.2020.126170

Yan, C., Yang, N., Wang, X., and Wang, Y. (2021). *VqBGH40a* isolated from Chinese wild *Vitis quinquangularis* degrades trans-piceid and enhances trans-resveratrol. *Plant Sci.* 310, 110989. doi: 10.1016/j.plantsci.2021.110989

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

Zhang, Y., Zhang, X., and Liu, H. F. (2020). Preliminary study on the coloring mechanism of *Vitis amurensis* pericarp transcriptome. *Mol. Plant Breed. (2003).* 18, 6000–6006. doi: 10.13271/j.mpb.018.006000

Zhou, Y., Minio, A., Massonnet, M., Solares, E., Lv, Y., Beridze, T., et al. (2019). The population genetics of structural variants in grapevine domestication. *Nat. Plants* 5, 965–979. doi: 10.1038/s41477-019-0507-8

Zhu, Z., Li, G., Liu, L., Zhang, Q., Han, Z., Chen, X., et al. (2019). A R2R3-MYB transcription factor, *VvMYBC2L2*, functions as a transcriptional repressor of anthocyanin biosynthesis in grapevine (*Vitis vinifera* l.). *Molecules* 24, 92. doi: 10.3390/molecules24010092

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 26 January 2023
DOI 10.3389/fpls.2023.1136616

OPEN ACCESS

EDITED BY
Lin Xi,
University of Hohenheim, Germany

REVIEWED BY
Bowen Liang,
Hebei Agricultural University, China
Libo Xing,
Northwest A&F University, China

*CORRESPONDENCE
Xiaozhao Xu
201801006@qau.edu.cn
Zhengnan Li
lizhengnan@imau.edu.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Functional and Applied Plant Genomics, a section of the journal Frontiers in Plant Science

RECEIVED 03 January 2023
ACCEPTED 17 January 2023
PUBLISHED 26 January 2023

CITATION
Tian Q, Xu M, Wu D, Wang C, Wang X, Che Q, Li Z and Xu X (2023) Integrated transgene and transcriptome reveal the molecular basis of *MdWRKY87* positively regulate adventitious rooting in apple rootstock. *Front. Plant Sci.* 14:1136616. doi: 10.3389/fpls.2023.1136616

Integrated transgene and transcriptome reveal the molecular basis of *MdWRKY87* positively regulate adventitious rooting in apple rootstock

Qiuye Tian[1†], Mengli Xu[2†], Dongchen Wu[3], Chaoping Wang[4], Xianlin Wang[5], Qinqin Che[1], Zhengnan Li[6*] and Xiaozhao Xu[1*]

[1]College of Horticulture, Qingdao Agricultural University, Qingdao, China, [2]Engineering Laboratory of Genetic Improvement of Horticultural Crops of Shandong Province, Qingdao Agricultural University, Qingdao, China, [3]Laboratory of Quality & Safety Risk Assessment for Fruit (Qingdao), Ministry of Agriculture and Rural Affairs, Qingdao Agricultural University, Qingdao, China, [4]Shandong Academy of Grape, Shandong Academy of Agricultural Sciences, Jinan, China, [5]Weihai Yingjuval Nursery Limited Company, Weihai International Port Economic and Technological Development District, Weihai, Shandong, China, [6]College of Horticulture and Plant Protection, Inner Mongolia Agricultural University, Hohhot, China

For most fruit and forest species vegetative propagated from elite genotypes, adventitious rooting is essential. The ability to form adventitious roots significantly decreased during the juvenile to adult phase change. Apart from the miR156-SPL pathway, whether there is another regulation mechanism controlling age-dependent adventitious rooting ability remained largely unknown. In the present study, we showed that *MdWRKY87* expression level was positively correlation with adventitious rooting ability. In addition, over-expressing of *MdWRKY87* in tobacco leads to enhanced adventitious rooting ability, more adventitious root number and accelerated adventitious rooting process. Comparative transcriptome profiling indicated that *MdWRKY87* overexpression can activate the expression of adventitious rooting-induced genes, such as *WOX11* and *AIL*. In addition, *MdWRKY87* overexpression can inhibit the transcription of adventitious rooting-repressed genes, such as *AUX/IAAs* and type-B cytokinin *RRs*. Collectively, here we demonstrated that higher expression level of *MdWRKY87* contributes to age-dependent adventitious rooting-competent in juvenile apple rootstock.

KEYWORDS

auxin, adventitious root formation, transcriptome, apple, rootstock

1 Introduction

The root system is able to adapt its architecture and morphology to the soil environment and physiological requirements of the plant in a flexible manner (Casimiro et al., 2003). Plant root systems typically consist of primary roots, lateral roots, and adventitious roots (Ji et al., 2022). Among these root types, ARs display high phenotypic plasticity in response to a

variety of environmental stimuli (Li et al., 2022). Monocotyledons, like rice and maize, produce adventitious roots during normal development. In most tree species, especially poplars, cuttings are the primary means of propagation, and species differ in their ability to form adventitious root significantly. For most fruit and forest species propagated from elite genotypes, adventitious rooting is essential for proliferation. Plant tissue totipotency enables the adventitious root formation from non-root tissues, which has been extensively utilized for vegetative propagation of agricultural and forestry plants (Bellini et al., 2014). Most fruit species, such as apple rootstock, adventitious root formation is not easy to limit good varieties of asexual reproduction. Despite some preliminary studies on the mechanism of adventitious roots formation in difficult-to-root trees, and more in-depth research is needed.

In most tree species, the ability to form adventitious roots significantly decreased during the juvenile to adult phase change. Extensive research has been conducted to overcome problems related to the loss or reduction of the ability of difficult-to-root trees to form adventitious roots (Levy et al., 2014; Xu et al., 2017; Wang et al., 2019; Li et al., 2021). In our previous study, we have showed that the rooting rates of cuttings from juvenile and rejuvenated donor plants were significantly higher than those of cuttings from adult trees in *Malus xiaojinensis* (Xu et al., 2017). The high expression of miR156 is positively correlated with auxin-induced adventitious roots formation (Levy et al., 2014; Xu et al., 2017). MiR156 functions *via* its target gene *MxSPL26* in regulating adventitious root formation (Xu et al., 2017). Our previous research has indicated that *MxSPL26* inhibited *MxHB13* expression by directly binding to its promoter (Li et al., 2021). During the adult phase, MxSPL26 interacts with auxin-induced MxTIFY9 and co-represses *MxHB13* expression, leading to reduced AR formation (Li et al., 2021). Although several pathways and components of auxin-mediated molecular regulatory networks underlying adventitious root formation in apple have been identified, but the molecular mechanisms need to be investigated further.

The WRKY proteins are a superfamily of transcription factors found exclusively in plant (Rushton et al., 2010). The name of WRKY is derived from the highly conserved amino acid sequence containing WRKYGQK and the zinc finger-like motifs (Cys2-His2 or Cys2-HisCys) (Eulgem et al., 2000; Rushton et al., 2010). Based on both the number of WRKY domains and their zinc-finger motif, WRKYs can be divided into three distinct groups (Eulgem et al., 2000). The WRKYs have been found to function in seed dormancy, embryo and trichome formation, senescence, hormone synthesis, signal transduction, defense responses and abiotic stresses (Johnson et al., 2002; Lagace and Matton, 2004; Zhang et al., 2004; Li et al., 2015; Singh et al., 2017; Hu et al., 2018; Liang et al., 2020; Kang et al., 2021; Zhu et al., 2021). To date, only a small number of research have reported that WRKYs participate in adventitious roots formation. The group IIe *WRKY* gene of *Catalpa* Scop, *CbNN1* expression increased with increasing adventitious rooting ability (Wang et al., 2019). PuWRKY75, as a transcription activator, controls the low phosphorus driven adventitious root formation through up-regulating *PuLRP1* and *PuERF003* transcription in *Populus ussuriensis* (Wang et al., 2022). The functions of WRKYs in adventitious root formation remain to be investigated.

In this study, we identified a group IIe subfamily *WRKY* gene, *MdWRKY87* from apple rootstock. The *MdWRKY87* protein was found located in the nucleus and functions as a transcriptional repressor in both yeast and plant cells. Our results also indicated that *MdWRKY87* promoted adventitious rooting through regulating root-related gene involved in auxin signaling pathway.

2 Materials and methods

2.1 Plant materials

The leafy stem cuttings of *M. xiaojinensis* (Mx) was used as the materials, because Mx has a high apomictic rate to ensure the juvenile materials stability (Xu et al., 2017). Semi-lignified leafy cuttings (8-10 cm in length) were excised from basal suckers (juvenile phase, Mx-J) and shoots from the canopy of reproductively mature trees (adult phase, Mx-A). The bases of leafy cuttings were immersed 1~2 cm in depth into a 3.0 g L^{-1} indole butyric acid (IBA, Sigma-Aldrich, St. Louis, MO, USA) solution for 1 min (Xu et al., 2017). Cutting dipped in IBA-free water was used as a control. After plugging the cuttings into 50 cell trays containing fine sand, they were incubated in a solar greenhouse. The rooting ability was evaluated at 35 days after treatment. Three biological replicates, each with at least 50 leafy cuttings, were used for the experiment to manage experiment errors.

Tissue-cultured 'M9T337' plantlets were sub-cultured in Murashige and Skoog (MS) medium containing 7.5 g L^{-1} agar and 30 g L^{-1} sugar (pH 5.8) with 0.5 mg L^{-1} IBA and 0.2 mg L^{-1} 6-benzylaminopurine (6-BA) (Cheng et al., 2020). After 30 days, stem cuttings were transferred into 1/2 medium containing rooting 7.5 g L^{-1} agar and 30 g L^{-1} sugar (pH 5.8) with 0.5 mg L^{-1} IBA and 0.1 mg L^{-1} 1-naphthalene acetic acid (NAA) for rooting. The tobacco (*Nicotiana tabacum*) plants were sub-cultured in MS medium without hormone. The plantlets were grown under a 16 h light/8 h dark photoperiod with day/night temperatures of 25 ± 1°C and 20 ± 1°C.

2.2 Histological analysis

Paraffin sections of stem bases were prepared as previously described (Xu et al., 2017; Cheng et al., 2020; Jin et al., 2022), with some modifications. The bases of 'M9T337' stem cuttings were collected at 6, 9, and 12 days after transplanting on 1/2 MS medium with 0.5 mg L-1 IBA and 0.1 mg L-1 NAA. The bases of tobacco stem cuttings were excised at 2, 4 and 6 days after subculture on hormone-free MS medium. The samples were fixed in FAA solution (70% ethanol: formaldehyde: acetic acid, 95:5:5 [v/v/v]) for 2 days at room temperature, and store at 4 °C. Then samples were dehydrated with a graded ethanal series (50%, 70%, 85%, 95%, and 100%), infiltrated with xylene, and embedded in paraffin. Cross sections with a 10 μm in thickness were cut with a Leica RM2245 (Leica Microsystems, Wetzlar, Germany) rotary microtome, transferred onto glass slides, deparaffined with xylene, and re-hydrated through an ethanol series, and stained with toluidine blue. Slides were observed using an optical microscope DM2500 (Leica Microsystems, Wetzlar, Germany) and photos were obtained using an attached digital camera DFC420 (Leica Microsystems, Wetzlar, Germany).

2.3 Gene expression analysis

Total RNA was extracted from approximately 0.5 g of frozen sample using the TIANGEN Plant RNA Kit (TIANGEN biotech CO., LTD, Beijing, China, DP305). For each sample, 1 μg DNase-treated RNA was used to synthesize first-strand cDNA with oligo d(T) or random primer and HiScript® II Q RT SuperMix (Cat. R223-01, Vazyme, China). A LightCycler 480 instrument (Roche, Basel, Switzerland) and ChamQ SYBR Color qPCR Master Mix (Vazyme, Nanjing, China) were used for qRT-PCR. The relative expression levels of genes were normalized to the reference gene *EF1α* and calculated using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). All reactions were performed with at least three biological replicates. The primers are listed in Supplemental Table S1.

2.4 Histochemical GUS staining

The *MdWRKY87* promoter fragment (-2000 bp to -0 bp from the *MdWRKY87* ATG start codon) was inserted into pCambia1391 vector generating the *proMdWRKY87: GUS* construct. Tobacco leaves were transformed with the *Agrobacterium tumefaciens* strain GV3101 cells harboring a *proMdWRKY87: GUS* or *DR5: GUS* construct. Agrobacterium cells were re-suspended in buffer with (10 mM $MgCl_2$, 10 mM MES-KOH, pH 5.6; adding 200 μM acetosyringone immediately prior to use) to an OD_{600} of 0.8~1.0. After injected with 1 ml needleless syringes, the leaves were treated with 50μM IBA and collected 6 hours later. The leaves were submerged in the GUS staining solution for 24 h at 37°C. After staining, tissues were cleared by immersing in 70% ethanol. All primers used are listed in Supplemental Table S1.

2.5 Subcellular localization

The ORF fragment (stop codon removed) of *MdWRKY87* containing *Sma*I and *Xba*I sites were inserted into the Super1300-GFP vector to generate the *pSuper : MdWRKY87-GFP* construct. Subcellular localization was conducted as previously described (Cheng et al., 2021). *A. tumefaciens* cells (GV3101) expressing *pSuper : MdWRKY87-GFP* and a Cherry-labelled nuclear marker (NF-YA4-mCherry) was re-suspended using the buffer (10 mM MgCl2, 10 mM MES-KOH, pH 5.6; 200 μM acetosyringone). *pSuper : NF-YA4-mCherry* was used as a nuclear marker. The tobacco leaves were injected with the re-suspended *A. tumefaciens* cells using a 1 ml needleless syringe. Three days after infiltration, fluorescence signals of the infiltrated leaves were detected using a laser scanning confocal microscope (Leica TCS SP5 II, Wetzlar, Germany). The primers used for construction are listed in Supplemental Table S1.

2.6 Transcriptional activation analysis in yeast

The coding fragment of *MdWRKY87* were fused to the GAL4-BD in pBD-GAL4 vector. The transcriptional activation analysis was conducted as previously described (Cheng et al., 2021). The transactivation activity was verified by the growth of yeast AH109 harboring full-length of MdWRKY87 on SD/-Trp and SD/-Trp-His plates and was confirmed by a X-α-Gal staining assay. All primers used are listed in Supplemental Table S1.

2.7 Transcriptional activation analysis in *N. benthamiana*

The *MdWRKY87* ORF sequence without stop codon was cloned into the pBD-VP16 vector (Han et al., 2016). The reporter vector contained a GAL4-luciferase (LUC) containing five copies of the GAL4-binding element and a minimal CaMV35S promoter at the 5' end of the LUC gene (Han et al., 2016).The effector vectors or reporter vectors were introduced into *A. tumefaciens* strain GV3101. The *A. tumefaciens* cells was re-suspended to an OD_{600} of 1.0 using the buffer (10 mM $MgCl_2$, 10 mM MES-KOH, pH 5.6; 200 μM acetosyringone). *A. tumefaciens* cells harboring effector vector and reporter vector were mixed 1:1, then injected into the tobacco (*N. benthamiana*) leaves by using a 1 mL needleless syringe. After spraying 1 mM luciferin onto the leaves, luciferase imaging was performed using NEWTON 7.0 (VILBER LOURMAT, Paris, France). An assessment of LUC and REN activities was conducted using the Duo-Lite Luciferase Assay System (DD1205–01, Vazyme, Nanjing, China) and BioStack Ready (BioTek Instruments Inc., Winooski, Vermont, USA). LUC/REN ratio was used to calculate the results. The primers used for construction are listed in Supplemental Table S1.

2.8 Transgenic tobacco generation

The *MdWRKY87* ORF sequence without stop codon was cloned into pRI101 vector to generate *35S:MdWRKY87-OE* construct. The construct was introduced into wild type tobacco (*N. tabacum*) leaves by *A. tumefaciens*-mediated transformation as previously described (Zhu et al., 2022). The infected leaves were selected on MS medium containing 100 mg L^{-1} kanamycin and 300 mg L^{-1} cefotaxime sodium to generate *MdWRKY87*-overexpressing (*MdWRKY87*-OE) transgenic lines. Transgenic plants were propagated by subculture on hormone-free MS medium. The primers used for construction are listed in Supplemental Table S1.

2.9 RNA-Seq

Total RNA was extracted from the stem bases of MdWRKY87-OE and wild type plants using the TIANGEN Plant RNA Kit (TIANGEN biotech CO., LTD, Beijing, China, DP305). A total amount of 3 μg RNA per sample was used in RNA-seq library construction. An Illumina Hiseq (Illumina, CA, USA) system was used for RNA sequencing by Novogene (Novogene, Tianjin, China). A quality assessment was performed on raw data using FastQC. Following Trimmomatic filtering out adapters and unpaired reads, the remaining clean reads were used to calculate the expression of gene by using Kallisto, an RNA-seq quantification program (Bray et al., 2016). The N. tabacum genome was used as the reference genome (Edwards et al., 2017). The count of reads was normalized to Transcripts Per kilobase of exon model per Million mapped reads (TPM). The log2TPM values were subjected to generate the heat map by TBtools software (Chen et al., 2020). The RNA-seq data were

deposited in the NCBI Sequence Read Archive (accession number PRJNA917351).

3 Results

3.1 *MdWRKY87* expression correlates positively with adventitious rooting

According to our previous research, semi-lignified leafy cuttings from Mx-J and shoots from the canopy of Mx-A were used in this study (Xu et al., 2017; Li et al., 2021). As the previous results (Xu et al., 2017; Li et al., 2021), Mx-J cuttings exhibited a high adventitious rooting ability (Figures 1A–C). After IBA treatment, the rooting percentage of Mx-J cuttings (85.14%) was significantly higher than that of Mx-A cuttings (3.57%, Figure 1B). Neither the cuttings of Mx-A nor Mx-J exhibited the ability of adventitious rooting (Figure 1B). Moreover, the adventitious root number per cutting of Mx-J was significantly more than that of Mx-A (Figure 1C). We previously identified the expression of the WRKY transcription factor family genes in the cutting stems of Mx-A and Mx-J after IBA treatment (Che et al., 2021). The expression of *MdWRKY87* gene was significantly induced in the Mx-J cutting treated with IBA (Figure 1C). However, there was no difference in the mRNA levels of *MdWRKY87* in the Mx-A cuttings treated with IBA or untreated control (Figure 1C).

To further define the relationship between *MdWRKY87* expression levels and adventitious rooting formation, we next examined the expression pattern of *MdWRKY87* during adventitious rooting formation of tissue culture plantlets in apple. According to the paraffin sections of stem bases of apple plantlets, we found that primordia with

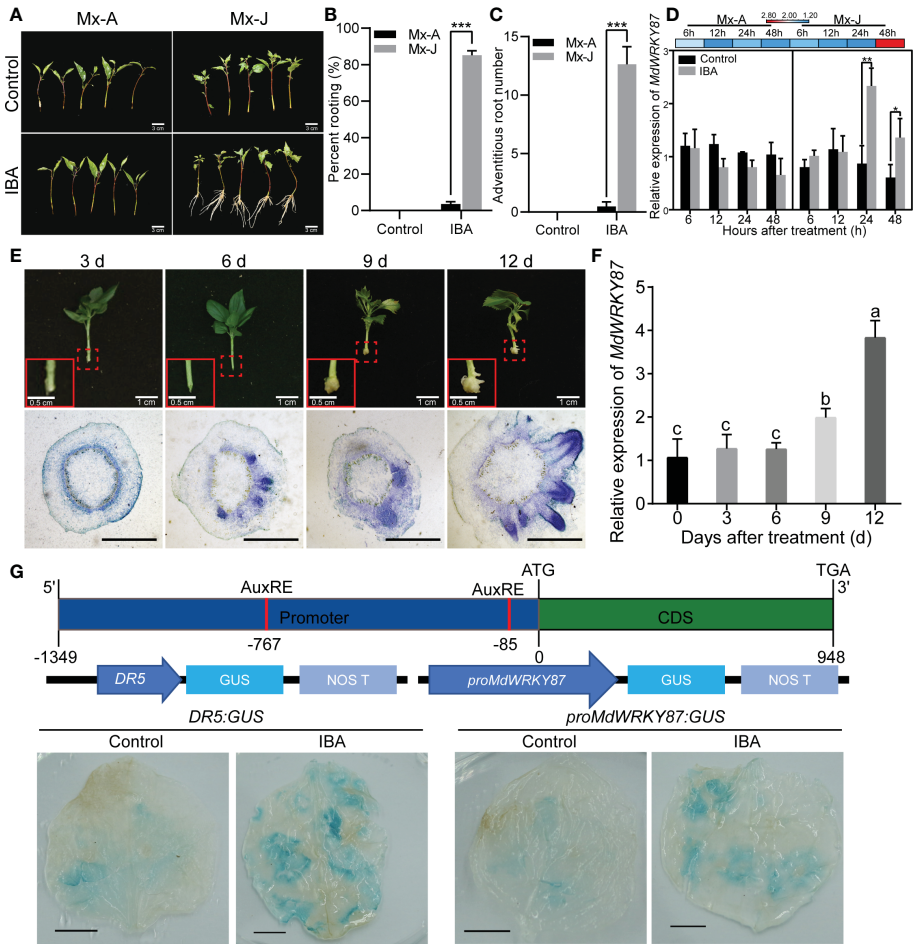


FIGURE 1
The expression of *MdWRKY87* is positively correlated with auxin-induced adventitious root formation. **(A)** Indole butyric acid (IBA) treatment induced adventitious root formation of leafy cutting from juvenile (Mx-J) not adult (Mx-A) phase in *M. xiaojinensis*. Scale Bar = 3 cm. **(B)** Adventitious rooting percentage of leafy cutting from Mx-A and Mx-J phase of *M. xiaojinensis* after IBA treatment. **(C)** Number of adventitious roots. For B and C, the mean values ± SD are shown for three biological replicates. Student's *t*-test, *** $P < 0.001$. **(D)** Relative expression of *MdWRKY87* gene in leafy cutting of Mx-A and Mx-J after IBA treatment. The mean values ± SD are shown for three biological replicates. Stem bark samples of 0.5 to 1.0 cm basal sections of 20 Mx-J or Mx-A cuttings treated with IBA-free solution or IBA were pooled together as one biological replicates. Student's *t*-test, ** $P < 0.01$. **(E)** The process of adventitious root formation of tissue culture plantlets growing on 1/2 MS medium containing IBA. Top, representative pictures of adventitious root formation of plantlets. The red dotted boxes indicate portion of stem base magnified in red solid line boxes (Scale bar = 0.5 cm). Scale bar = 1 cm. Bottom, transverse sections of stem base during the adventitious root formation. Scale bar = 1 mm. **(F)** Relative expression of *MdWRKY87* gene in stem base of tissue culture plantlets. Different letters indicate statistically significant differences ($P < 0.05$) at by Duncan's test. **(G)** GUS activity of *N. benthamiana* leaves transiently transformed with *proMdWRKY87:GUS* or *DR5:GUS* after IBA treatment. Top, schematic representation of *MdWRKY87*. Medium, schematic representation of the *proMdWRKY87:GUS* and *DR5:GUS* constructs. Bottom, GUS activity analysis. Scale bar = 1 cm. *p < 0.05.

dome-shaped adventitious structures were clearly visible at 6~9 days after transplanting (Figure 1E) as our previous results (Cheng et al., 2020). After 12 days of transplantation, adventitious roots began to appear. As expected, the gene expression levels of *MdWRKY87* increased significantly with the emergence of adventitious root from stem base (Figure 1F).

To further validate *MdWRKY87* expression in response to IBA treatment, *Agrobacterium tumefaciens* cells (GV3101) harboring the *proMdWRKY87:GUS* or *DR5:GUS* construct (auxin-responsive reporter) were transiently transformed into tobacco leaves. After IBA treatment, the levels of GUS proteins obviously increased in both of leaves transformed with *proMdWRKY87:GUS* and *DR5:GUS* construct (Figure 1G), suggesting that the promoter activity of *MdWRKY87* responds to auxin. These results indicate high *MdWRKY87* expression correlates positively with adventitious rooting and may regulate auxin-mediated adventitious root development.

3.2 *MdWRKY87* is located in the nucleus and functions as a transcriptional repressor

Multiple sequence alignments of WRKY22 homologs from apple indicated that *MdWRKY87* harbored a conserved WRKYGQ domain and a C2H2 (C-X5-C-X23-H-X1-H) zinc-finger motif at its C terminus (Figure 2A) and belonged to WRKY group IIe (Eulgem et al., 2000). To confirm whether *MdWRKY87* functions as a transcription factor, we expressed *MdWRKY87* fused to green fluorescence protein in tobacco leaves and observed that the fusion protein localized to the nucleus (Figure 2B). A transactivation assay in yeast indicated that the *MdWRKY87* protein has transcriptional activation activity (Figure 2C). We also performed a dual-luciferase transactivation assay in tobacco leaves. The results showed that the luciferase activity of co-expression of the reporter with pBD-WRKY87 -VP16 was significantly lower than that in the pBD-VP16 control (Figures 2D, E). These results suggested that *MdWRKY87* is indeed a transcription repressor of WRKY group IIe subfamily.

3.3 *MdWRKY87* positively regulates adventitious root development in tobacco

To characterize the role of *MdWRKY87* in adventitious rooting, we generated overexpression lines of *MdWRKY87* (*MdWRKY87*-OE) in tobacco plants by *Agrobacterium*-mediated transformation (Supplemental Figure S1). We tested the effects of *MdWRKY87* overexpression on adventitious root formation in *MdWRKY87*-OE

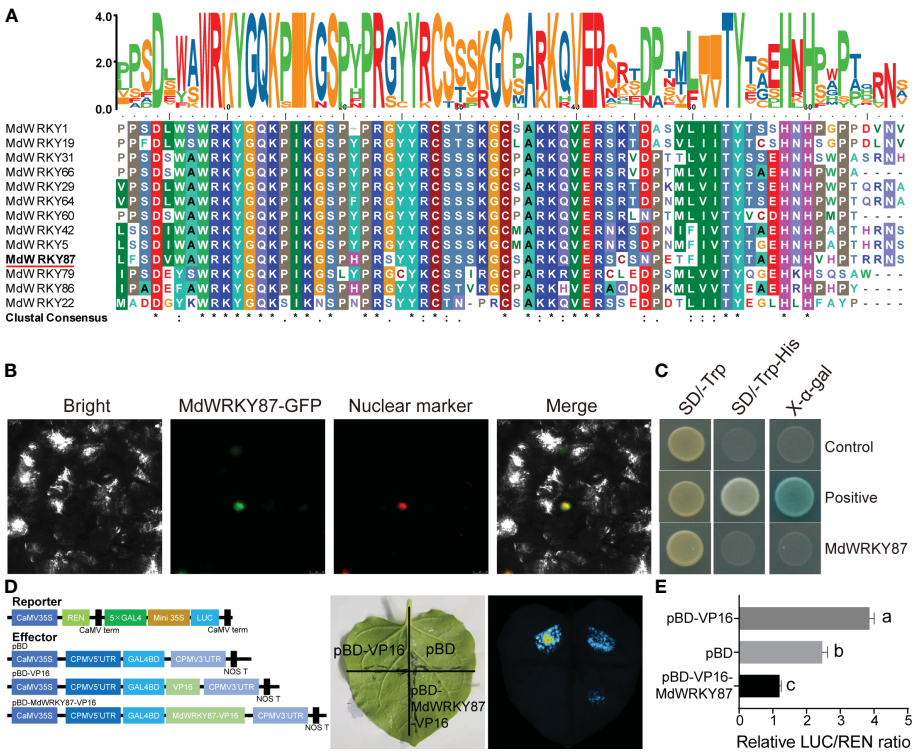


FIGURE 2
MdWRKY87 is a transcription repressor of WRKY family. **(A)** Sequence comparison of WRKY domains from WRKY IIe proteins, shaded by conserved amino acids. The accession numbers of all genes used are listed in the Methods section. **(B)** Subcellular localization of MdWRKY87 in *N. benthamiana* leaves. *Agrobacterium* carrying the pSuper : *MdWRKY87-GFP* and pSuper : *NF-YA4-mCherry* as a nuclear marker genes was co-infiltrated into *N. benthamiana* leaves. Images were captured 2 d following agroinfiltration. **(C)** Transcriptional activity analysis of MdWRKY87 protein in yeast. The transactivation activity was verified by the growth of yeast AH109 harboring *MdWRKY87* ORF on SD/-Trp and SD/-Trp-His plates and was confirmed by a X-α-Gal staining assay. **(D)** Transcriptional activity analysis of MdWRKY87 in *N. benthamiana* leaves. Reporter and effector constructs were co-infiltrated into *N. benthamiana* leaves. Left, schematic representation of the effector and reporter constructs. Right, live imaging of *N. benthamiana* leaves expressing reporter and effector constructs. **(E)** The dual-luciferase activity assay. The values were determined by calculating the ratio of LUC activity to REN activity (LUC/REN). Different letters indicate statistically significant differences ($P < 0.05$) at by Duncan's test.

lines #9, #5 and #6 (Figures 3A, B). During the adventitious rooting process, there were obvious morphological differences between *MdWRKY87*-OE transgenic plants and non-transformed wild type (WT) (Figure 3A). The adventitious rooting rate, root number and root length per stem were significantly higher in *MdWRKY87*-OE lines than that in WT plants (Figures 3C–E and Supplemental Movie S1). These results supported the notion that high *MdWRKY87* expression correlates positively with adventitious rooting.

To check whether *MdWRKY87* affects the initiation of adventitious root primordia, we conducted the cross sections of the stems of WT and *MdWRKY87-OE* transgenics lines during adventitious rooting. The initiation of adventitious root primordia was accelerated in *MdWRKY87-OE* transgenics lines from 2 days after subculture on MS (Figure 4). Moreover, the adventitious root primordium in *MdWRKY87-OE* transgenics lines were well-developed compared with the WT plantlets at 4 days after subculture, suggesting that high *MdWRKY87* expression accelerates the initiation and development of adventitious root primordia.

3.4 Interaction between *MdWRKY87* and auxin during adventitious rooting

To determine whether *WRKY87* regulates adventitious root formation through modulating auxin polar transport, we examined adventitious rooting capacity in wild-type, *MdWRKY87*-OE tobacco plants stem cuttings grown on MS medium supplemented with 1-N-naphthylphthalamic acid (NPA). Adventitious rooting was almost absolutely inhibited in both wild-type and *MdWRKY87*-OE tobacco plants treated with 20μM NPA (Figure 5A). To check whether NPA affects the initiation of adventitious root primordia, we conducted the cross sections of the stems of WT and *MdWRKY87*-OE lines during adventitious rooting. The initiation of adventitious root primordia was both inhibited in wild-type and *MdWRKY87*-OE transgenic lines under NPA treatment (Figure 5B). Dense aerial roots developed on the stems of transgenic and wild-type plants under NPA treatment (Figure 5C).

3.5 Transcriptome profiling of *MdWRKY87*-dependent gene expression during adventitious root formation

To understand how *MdWRKY87* regulates adventitious root development, we conducted a comparative transcriptome analysis of the stem of WT and *MdWRKY87*-OE tobacco plants. A total 6690 differentially expressed genes (DEGs) in tobacco were identified, including 2000 downregulated and 4690 upregulated genes (Supplemental Figure S2). These DEGs were then subjected to Gene Ontology (GO) functional classification. Within the three GO categories identified, a greatest number of DEGs was significantly

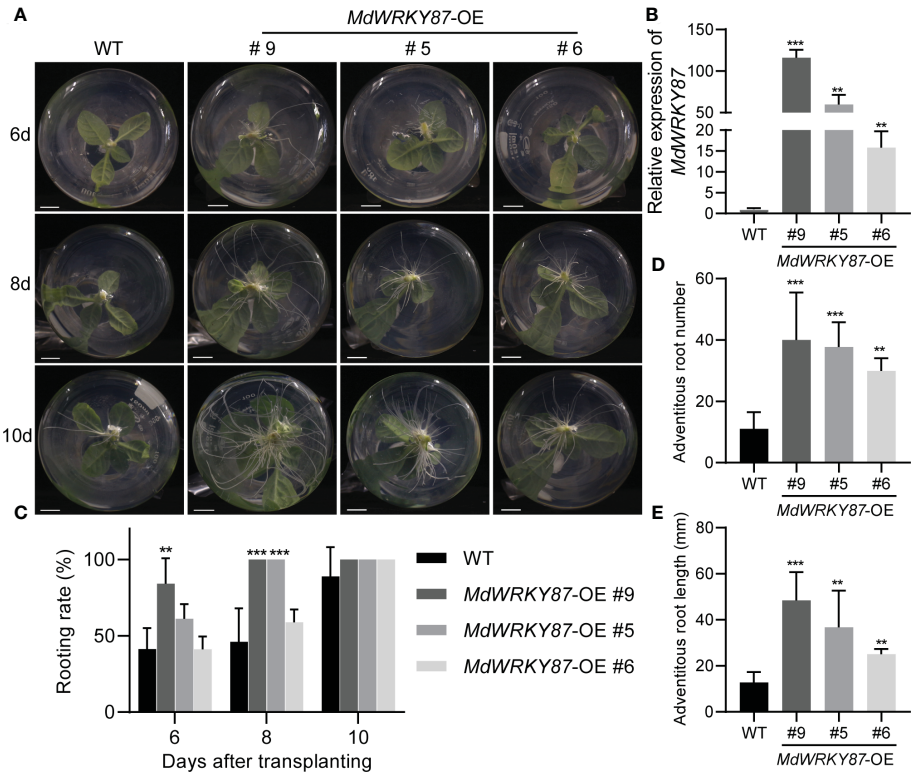


FIGURE 3
MdWRKY87 accelerates adventitious root formation in transgenic tobacco (*N. tabacum*) plants. **(A)** Phenotypes of adventitious roots in wild-type (WT) and independent transgenic lines (*MdWRKY87*-OE9, *MdWRKY87*-OE5, and *MdWRKY87*-OE6) in the absence of auxin at 6, 8 and 10 days after transplanting into hormone-free MS medium. **(B)** qRT-PCR analysis of *MdWRKY87* expression in WT and *MdWRKY87*-OE lines. **(C)** Adventitious rooting percentage of wild-type (WT) and independent transgenic lines at 6, 8 and 10 days after transplanting into hormone-free MS medium. **(D, E)** Number **(D)** and length **(E)** of adventitious roots at 10 days after transplanting into hormone-free MS medium. The mean values ± SD are shown for three biological replicates. Asterisks indicate significant differences between WT and each transgenic lines by Student's *t*-test (***P*< 0.01; ****P*< 0.001).

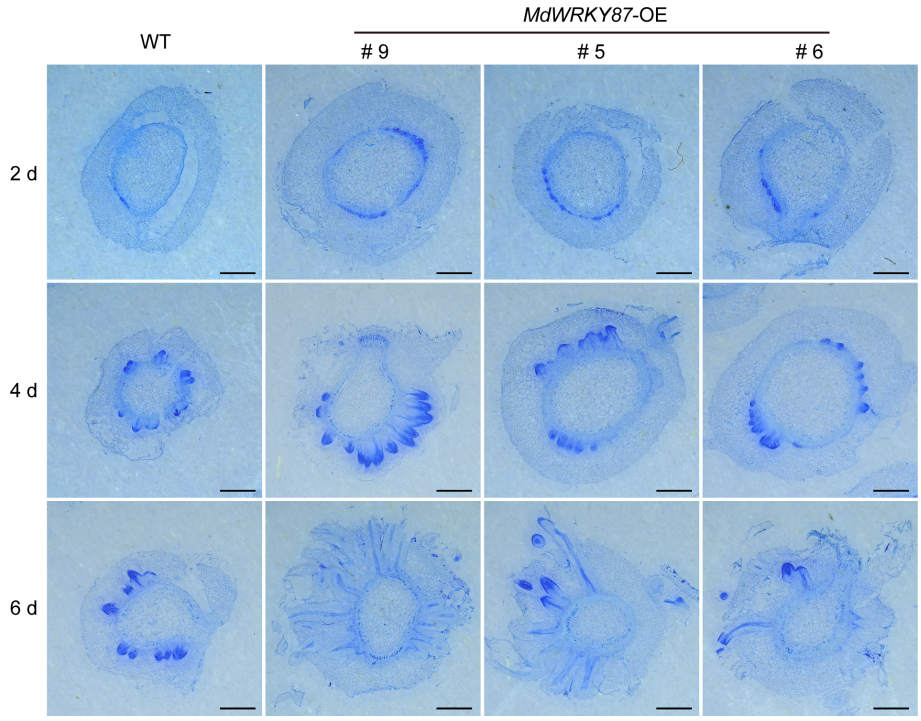


FIGURE 4
MdWRKY87 accelerates the formation of adventitious root primordium in transgenic tobacco (*N. tabacum*) plants. Cross-sections of stem bases at 2, 4 and 6 days after transplanting during adventitious root formation in WT and *MdWRKY87*-OE lines. Sections were stained by toluidine blue. Scale bar = 1 mm.

enriched in GO categories 'biological process' (Supplemental Figure S3). Within the 'biological process' GO categories, the top 4 GO terms were "regulation of cellular process", "RNA biosynthetic process", "nucleic acid-templated transcription" and "transcription, DNA-templated", respectively (Figure 6A). To further dig out the key genes regulated by *MdWRKY87*, we analyzed these four GO categories. Venn diagram analysis showed that 272 overlapping DEGs were identified in these four GO categories (Figure 6B). Most of these overlapping genes were annotated as transcription factor (TF) genes and transcriptional regulator (TR) genes (Figure 6C). Among these overlapping DEGs, there were 242 TF genes, 17 TR genes, and 13 other genes (Figure 6C). Within these TF genes, *AP-ERFBP*, *NAC*, and *HB* type *TF* genes account for a large proportion (Figure 6C, Supplemental Figure S4). Most of these TR genes are of the types Orphans and *AUX/IAA* (Figure 6C, Supplemental Figure S4).

Further analysis indicated these TF and TR genes mainly enriched in auxin and cytokinin signaling pathway. It is well known auxin and cytokinin appear to play antagonistic roles in the adventitious rooting process. Among these TF and TR genes family, previous study has demonstrated that *AUX/IAAs*, *AINTEGUMENTA* (*ANT*), *AINTEGUMENTA LIKE1* (*AIL*), *WUSCHEL-RELATED HOMEOBOX* (*WOX*), and type-B *cytokinin Response Regulator* (*RR*) family genes involved in adventitious rooting process (Ramírez-Carvajal et al., 2009; Rigal et al., 2012; Lakehal et al., 2019; Geng et al., 2023). In genes up-regulated by *MdWRKY87*, *WUSCHEL-RELATED HOMEOBOX11* (*WOX11*), and *AINTEGUMENTA LIKE1* (*AIL*) are positive regulators of adventitious root formation. For genes down-regulated by *MdWRKY87*, *AUX/IAAs* and type-B *Response Regulator* genes are negative regulator of adventitious root formation. Hence, negative regulation of *AUX/IAAs* and type-B *Response Regulator* genes and positive regulation of *WOX11* and *AIL* by MdWRKY87 contributes to the enhanced adventitious rooting ability in transgenic tobacco plants. The expression level of these genes was confirmed by quantitative real-time (qRT)-PCR, thus supporting the RNA-seq results (Figure 7).

4 Discussion

Although, it is well known that juvenile or rejuvenated phase leafy cuttings are much easier to root than the adult ones in perennial woody plants, the underlying molecular mechanism that mediates these differences is largely unknown.

4.1 Up-regulate *MdWRKY87* contributed to age-dependent adventitious rooting-competent in apple rootstock

For rooting recalcitrant woody plants, juvenility is necessary for efficient adventitious rooting. In general, rooting rates in adult plants are usually lower than in juvenile plants (Xu et al., 2017). Recent studies have provided a paradigm for the molecular basis of age-dependent adventitious rooting ability (Sun and Zhu, 2021). According to previous reports, we can understand the molecular basis from at least two independent signaling pathways: (i) *via* the accumulation of EIN3 protein in adult plants, which directly suppresses expression of *WUSCHEL RELATED HOMEOBOX* (*WOX*) genes to inhibit rooting (Ma et al., 2020); (ii) the miR156-

10.3389/fpls.2023.1136616

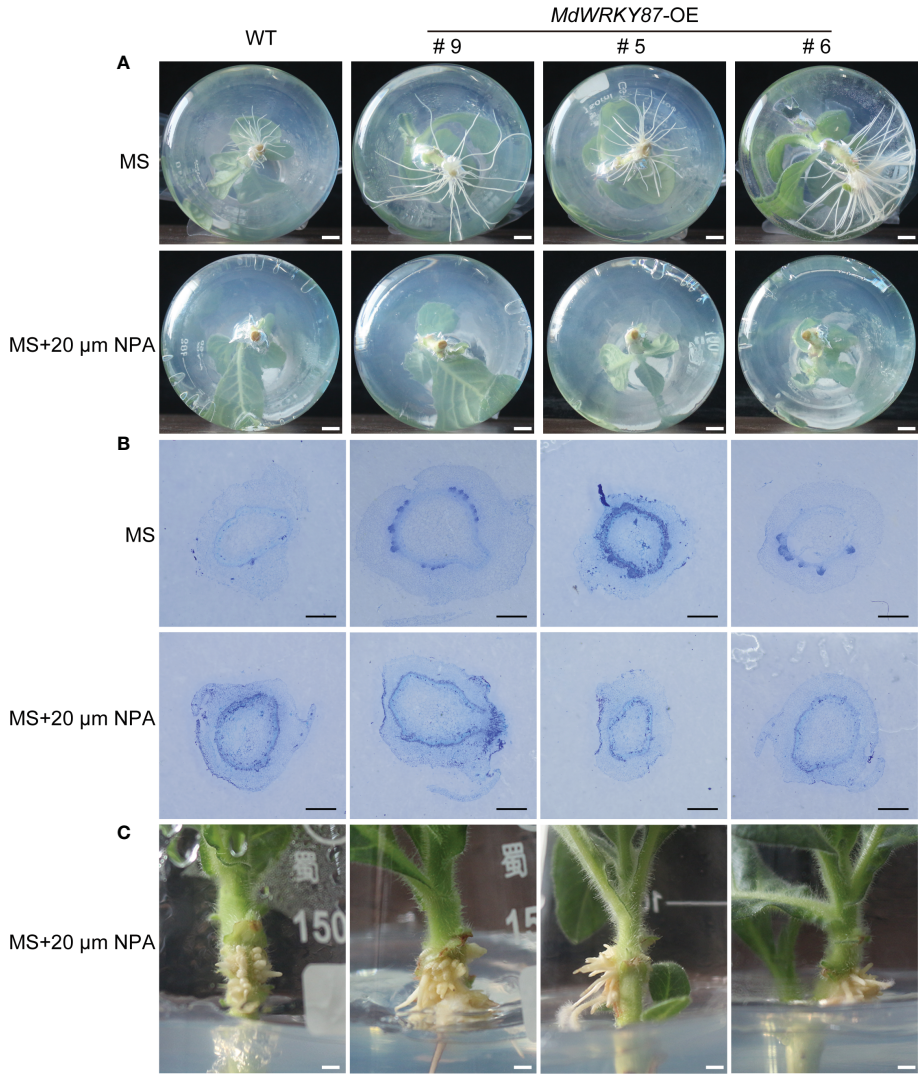


FIGURE 5
Effects of NPA on adventitious root formation in transgenic tobacco and wild-type plants. **(A)** Phenotypes of 10-day-old plants (WT, and *MdWRKY87*-OE lines) grown on MS medium with 20 μM NPA. **(B)** Cross-sections of stem bases at 5 days after transplanting during adventitious root formation in WT and *MdWRKY87*-OE lines. Sections were stained by toluidine blue. Scale bars = 500 μm. **(C)** Stem air root of wild type and *MdWRKY87*-OE lines at 15 days after transplant to MS medium with 20 μM NPA.

SPLs pathway, which modulates root regeneration by crosstalk with auxin signaling pathway (Xu et al., 2016; Xu et al., 2017; Ye et al., 2020a; Ye et al., 2020b; Li et al., 2021). Whether there is other regulation mechanism controlling age-dependent adventitious rooting ability? Here we demonstrated that higher expression level of *MdWRKY87* contributing to adventitious rooting-competent in juvenile apple rootstock independent of miR156/SPL pathway. As revealed by qRT-PCR, the expression level of *MdWRKY87* was positively correlated with adventitious rooting ability (Figures 1D, F). Over-expressing of *MdWRKY87* in tobacco leads to enhanced adventitious rooting ability, more adventitious root number and accelerated adventitious rooting process (Figures 3, 4 and Supplemental Movie S1). In addition, *SPL* family genes was not found in the transcriptome data of *MdWRKY87-OE* transgenic plants. It has been widely reported that WRKY transcription factors participate in the regulation of plant growth and development, abiotic stress responses, and disease response. However, the function of WRKY involved in adventitious rooting remains largely unknown. Here, we found WRKY87 transcription factors play essential role in adventitious root formation. In agreement to this result, PuWRKY75 was identified to control the low phosphorus driven adventitious root formation in *Populus ussuriensis* (*Pu*) (Wang et al., 2022). In addition, it was interesting that PuWRKY75 act as a transcriptional enhancer, but MdWRKY87 act as a transcriptional inhibitor (Figures 2C–E). There will be more *WRKY* genes, which be identified involving in adventitious rooting process in the future research.

4.2 *MdWRKY87* involved in auxin signaling pathway during adventitious rooting

It was well known that auxin play a dominant role in regulation of adventitious root formation (Lakehal et al., 2019). To further identify how *MdWRKY87* regulate adventitious root formation, we analyzed

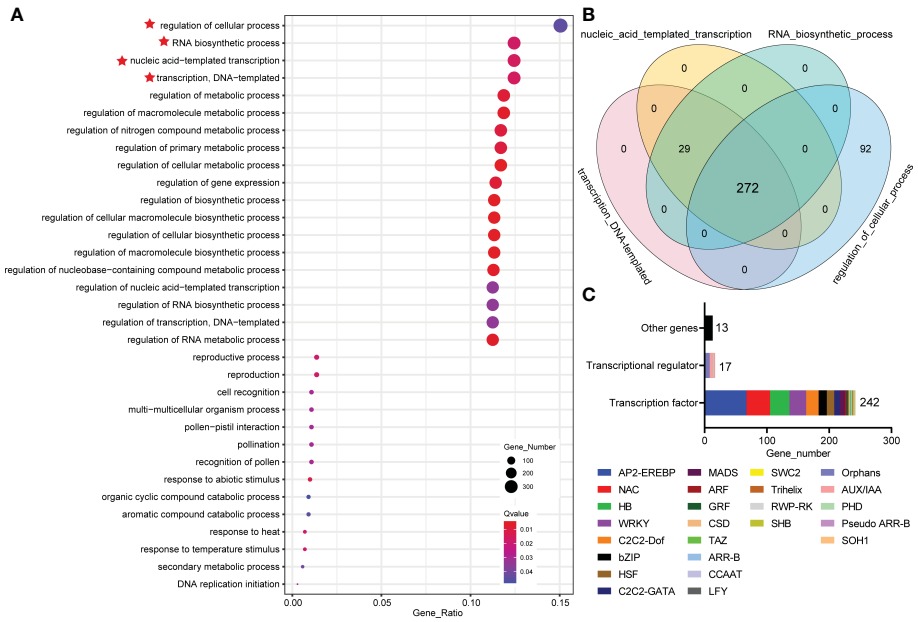


FIGURE 6
Comparative transcriptome analysis of the stem of WT and *MdWRKY87*-OE tobacco plants. **(A)** Functional categorization of DEGs based on the biological process of Gene Ontology (GO). The symbol red star represent the top 4 GO terms. **(B)** Venn diagram of DEGs in top 4 GO terms based on the biological process of GO. **(C)** The analysis of overlapping DEGs involved in "transcription, DNA-templated", "nucleic acid-templated transcription", "RNA biosynthetic process" and "regulation of cellular process".

the interaction between *MdWRKY87* and auxin. *MdWRKY87* has significantly enhanced expression levels in response to exogenous IBA treatment (Figures 1D, F). In addition, *pro MdWRKY87*: *GUS* transiently transformed tobacco leaves exhibited induction of GUS activity after spraying with IBA and the promoter region of *MdWRKY87* has auxin response element (Figure 1G). These suggest that *MdWRKY87* acts downstream of auxin to regulate adventitious root formation. However, the adventitious rooting

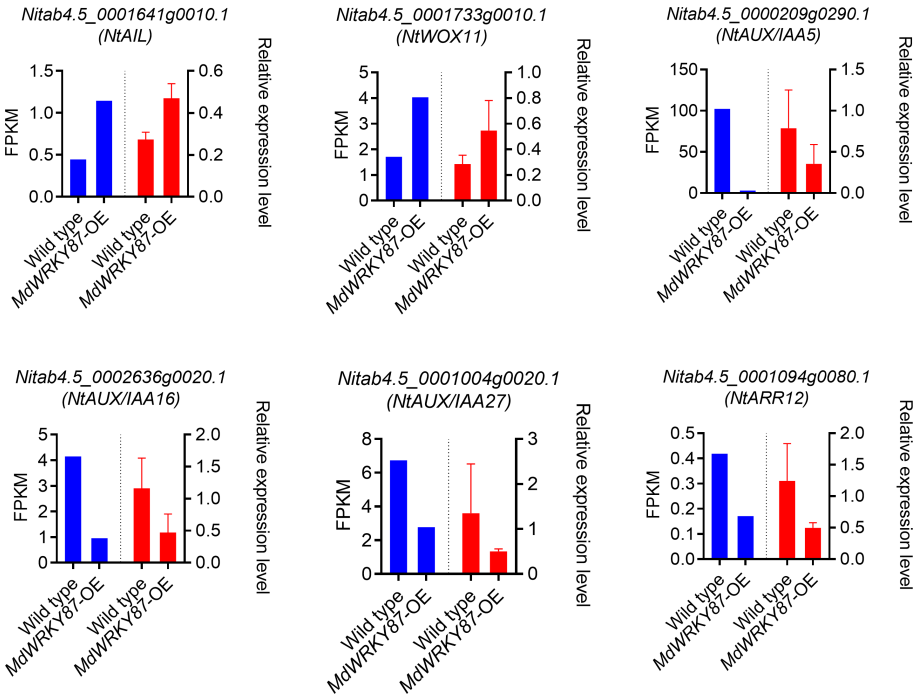


FIGURE 7
qRT-PCR validation of DEGs in RNA-seq data. Six genes were selected from RNA-seq differentially expressed genes (DEGs) to be validated by qRT-PCR. The data were reported as the means ± SE of three biological replicates. A comparison between the gene expression ratios obtained from RNA-seq data and qRT-PCR.

ability was inhibited in *MdWRKY87*-OE transgenic tobacco lines upon treatment with the polar auxin transport inhibitor NPA (Figure 5). Taken together, we conclude *MdWRKY87* act not only downstream of auxin, but also feedback regulation during adventitious rooting process. Consistent with this, previous data also demonstrate that WRKY71/EXB1 play pivotal roles in shoot branching by regulating auxin pathways (Guo et al., 2015). However, specific mechanism underlying the crosstalk between MdWRKY87 and auxin signaling needs to be further demonstrated.

4.3 *MdWRKY87*-dependent regulation of adventitious rooting related genes in transgenic tobacco

MdWRKY87 was known to control adventitious root formation (Figure 3), but the regulatory role of *MdWRKY87* was unknown. Comparative transcriptome profiling between the wild type and *MdWRKY87-OE* transgenic plants was conducted in this study (Figure 6). A set of 272 overlapping DEGs were identified through GO enrichment analysis, implying their potential importance for MdWRKY87-dependent adventitious rooting formation (Figure 6B). Among these DEGs, we found *MdWRKY87* overexpression can activate the expression of adventitious rooting-induced genes, such as *WOX11* and *AIL*, and the counterparts of *WOX11* and *AIL* promote adventitious root formation in *A.thaliana* and poplar (Rigal et al., 2012; Hu and Xu, 2016; Geng et al., 2023). In addition, *MdWRKY87* overexpression can inhibit the transcription of adventitious rooting-repressed genes, such as *AUX/IAAs* and *RRs*, and the counterparts of *AUX/IAAs* and type-B *RRs* inhibit adventitious root formation in Arabidopsis, apple, and poplar (Ramírez-Carvajal et al., 2009; Lakehal et al., 2019; Zhao et al., 2020). Similarly, it was demonstrated that PuWRKY75 interacted with PuMYB40 and directly co-regulate *PuLRP1* and *PuERF003* to promote adventitious root formation in *P. ussuriensis* (Wang et al., 2022).

In conclusion, a relatively high expression level of *MdWKRY87* contribute to improving adventitious rooting ability. Based on this potential mechanisms, artificial methods for adventitious rooting ability improving will be created *via* manipulating *MdWRKY87* gene expression, especially for rooting recalcitrant woody perennial species.

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

XX, and ZL conceived and designed the experiment. QT and MX conducted the experiment and data analysis. DW, CW, XW and QC contributed to the data analysis. QT and MX wrote the manuscript. DW, CW, XW and QC drafted the discussion and revised the manuscript. All authors contributed to the article and approved the final version.

Funding

This research was funded by National Natural Science Foundation of China (grant number 32172523, and 31801824); the Breeding Plan of Shandong Provincial Qingchuang Research Team, China (2019); The Fruit innovation team of Shandong modern agricultural industry technology system (SDAIT-06-05); Shandong Provincial Cooperative Promotion Plan of Major Agricultural Technologies (SDNYXTTG-2022-08).

Conflict of interest

Author XW was employed by Weihai Yingjuval Nursery Limited Company.

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1136616/full#supplementary-material

SUPPLEMENTAL MOVIE S1
Time-lapse video showing the adventitious rooting of *MdWRKY87-OE* transgenic tobacco plants.

References

Bellini, C., Pacurar, D. I., and Perrone, I. (2014). Adventitious roots and lateral roots: similarities and differences. *Annu. Rev. Plant Biol.* 65, 639–666. doi: 10.1146/annurev-arplant-050213-035645

Bray, N. L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* 34 (5), 525–527. doi: 10.1038/nbt.3519

Casimiro, I., Beeckman, T., Graham, N., Bhalerao, R., Zhang, H., Casero, P., et al. (2003). Dissecting arabidopsis lateral root development. *Trends Plant Sci.* 8 (4), 165–171. doi: 10.1016/S1360-1385(03)00051-7

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 (8), 1194–1202. doi: 10.1016/j.molp.2020.06.009

Cheng, C., Che, Q., Su, S., Liu, Y., Wang, Y., and Xu, X. (2020). Genome-wide identification and characterization of *Respiratory burst oxidase homolog* genes in six *Rosaceae* species and an analysis of their effects on adventitious rooting in apple. *PloS One* 15 (9), e0239705. doi: 10.1371/journal.pone.0239705

Cheng, C., Yu, Q., Wang, Y., Wang, H., Dong, Y., Ji, Y., et al. (2021). Ethylene-regulated asymmetric growth of the petal base promotes flower opening in rose (Rosa hybrida). *Plant Cell* 33 (4), 1229–1251. doi: 10.1093/plcell/koab031

Che, Q., Zhang, Z., Song, W., Li, Q., Xu, X., and Wang, Y. (2021). Comparison of *WRKY* gene family bioinformatics in 'GDDH13' and *M. baccata* and expression analysis of *MdWRKY* gene family. *Mol. Plant Breed. (China)* 19 (5), 1484–1502. doi: 10.13271/j.mpb.019.001484

Edwards, K. D., Fernandez-Pozo, N., Drake-Stowe, K., Humphry, M., Evans, A. D., Bombarely, A., et al. (2017). A reference genome for *Nicotiana tabacum* enables map-based cloning of homeologous loci implicated in nitrogen utilization efficiency. *BMC Genomics* 18 (1), 448. doi: 10.1186/s12864-017-3791-6

Eulgem, T., Rushton, P. J., Robatzek, S., and Somssich, I. E. (2000). The WRKY superfamily of plant transcription factors. *Trends Plant Sci.* 5 (5), 199–206. doi: 10.1016/S1360-1385(00)01600-9

Geng, L., Li, Q., Jiao, L., Xiang, Y., Deng, Q., Zhou, D. X., et al. (2023). *WOX11* and *CRL1* act synergistically to promote crown root development by maintaining cytokinin homeostasis in rice. *New Phytol.* 237 (1), 204–216. doi: 10.1111/nph.18522

Guo, D., Zhang, J., Wang, X., Han, X., Wei, B., Wang, J., et al. (2015). The WRKY transcription factor *WRKY71/EXB1* controls shoot branching by transcriptionally regulating *RAX* genes in arabidopsis. *Plant Cell* 27 (11), 3112–3127. doi: 10.1105/tpc.15.00829

Han, Y. C., Kuang, J. F., Chen, J. Y., Liu, X. C., Xiao, Y. Y., Fu, C. C., et al. (2016). Banana transcription factor MaERF11 recruits histone deacetylase MaHDA1 and represses the expression of *MaACO1* and expansins during fruit ripening. *Plant Physiol.* 171 (2), 1070–1084. doi: 10.1104/pp.16.00301

Hu, Z., Wang, R., Zheng, M., Liu, X., Meng, F., Wu, H., et al. (2018). TaWRKY51 promotes lateral root formation through negative regulation of ethylene biosynthesis in wheat (*Triticum aestivum* l.). *Plant J.* 96 (2), 372–388. doi: 10.1111/tpj.14038

Hu, X., and Xu, L. (2016). Transcription factors WOX11/12 directly activate *WOX5/7* to promote root primordia initiation and organogenesis. *Plant Physiol.* 172 (4), 2363–2373. doi: 10.1104/pp.16.01067

Ji, X. L., Li, H. L., Qiao, Z. W., Zhang, J. C., Sun, W. J., You, C. X., et al. (2022). The BTB protein MdBT2 recruits auxin signaling components to regulate adventitious root formation in apple. *Plant Physiol.* 189 (2), 1005–1020. doi: 10.1093/plphys/kiac084

Jin, Y., Li, J., Zhu, Q., Du, X., Liu, F., Li, Y., et al. (2022). *GhAPC8* regulates leaf blade angle by modulating multiple hormones in cotton (*Gossypium hirsutum* l.). *Int. J. Biol. Macromol* 195, 217–228. doi: 10.1016/j.ijbiomac.2021.11.205

Johnson, C. S., Kolevski, B., and Smyth, D. R. (2002). *TRANSPARENT TESTA GLABRA2*, a trichome and seed coat development gene of arabidopsis, encodes a WRKY transcription factor. *Plant Cell* 14 (6), 1359–1375. doi: 10.1105/tpc.001404

Kang, G., Yan, D., Chen, X., Yang, L., and Zeng, R. (2021). HbWRKY82, a novel IIc WRKY transcription factor from *Hevea brasiliensis* associated with abiotic stress tolerance and leaf senescence in arabidopsis. *Physiol. Plant* 171 (1), 151–160. doi: 10.1111/ppl.13238

Lagace, M., and Matton, D. P. (2004). Characterization of a WRKY transcription factor expressed in late torpedo-stage embryos of *Solanum chacoense*. *Planta* 219 (1), 185–189. doi: 10.1007/s00425-004-1253-2

Lakehal, A., Chaabouni, S., Cavel, E., Le Hir, R., Ranjan, A., Raneshan, Z., et al. (2019). A molecular framework for the control of adventitious rooting by TIR1/AFB2-Aux/IAA-dependent auxin signaling in arabidopsis. *Mol. Plant* 12 (11), 1499–1514. doi: 10.1016/j.molp.2019.09.001

Levy, A., Szwerdszarf, D., Abu-Abied, M., Mordehaev, I., Yaniv, Y., Riov, J., et al. (2014). Profiling microRNAs in *Eucalyptus grandis* reveals no mutual relationship between alterations in miR156 and miR172 expression and adventitious root induction during development. *BMC Genomics* 15, 524. doi: 10.1186/1471-2164-15-524

Liang, W., Sun, F., Zhao, Y., Shan, L., and Lou, H. (2020). Identification of susceptibility modules and genes for cardiovascular disease in diabetic patients using WGCNA analysis. *J. Diabetes Res.* 2020, 4178639. doi: 10.1155/2020/4178639

Li, X., Shen, F., Xu, X., Zheng, Q., Wang, Y., Wu, T., et al. (2021). An HD-ZIP transcription factor, MxHB13, integrates auxin-regulated and juvenility-determined control of adventitious rooting in *Malus xiaojinensis*. *Plant J.* 107 (6), 1663–1680. doi: 10.1111/tpj.15406

Li, P., Song, A., Gao, C., Jiang, J., Chen, S., Fang, W., et al. (2015). The over-expression of a chrysanthemum WRKY transcription factor enhances aphid resistance. *Plant Physiol. Biochem.* 95, 26–34. doi: 10.1016/j.plaphy.2015.07.002

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-delta delta C(T)) method. *Methods* 25 (4), 402–408. doi: 10.1006/meth.2001.1262

Li, Q. Q., Zhang, Z., Zhang, C. X., Wang, Y. L., Liu, C. B., Wu, J. C., et al. (2022). Phytochrome-interacting factors orchestrate hypocotyl adventitious root initiation in arabidopsis. *Development* 149 (10). doi: 10.1242/dev.200362

Ma, Z., Li, W., Wang, H., and Yu, D. (2020). WRKY transcription factors WRKY12 and WRKY13 interact with SPL10 to modulate age-mediated flowering. *J. Integr. Plant Biol.* 62 (11), 1659–1673. doi: 10.1111/jipb.12946

Ramírez-Carvajal, G. A., Morse, A. M., Dervinis, C., and Davis, J. M. (2009). The cytokinin type-b response regulator *PtRR13* is a negative regulator of adventitious root development in populus. *Plant Physiol.* 150 (2), 759–771. doi: 10.1104/pp.109.137505

Rigal, A., Yordanov, Y. S., Perrone, I., Karlberg, A., Tisserant, E., Bellini, C., et al. (2012). The *AINTEGUMENTA LIKE1* homeotic transcription factor *PtAIL1* controls the formation of adventitious root primordia in poplar. *Plant Physiol.* 160 (4), 1996–2006. doi: 10.1104/pp.112.204453

Rushton, P. J., Somssich, I. E., Ringler, P., and Shen, Q. J. (2010). WRKY transcription factors. *Trends Plant Sci.* 15 (5), 247–258. doi: 10.1016/j.tplants.2010.02.006

Singh, A. K., Kumar, S. R., Dwivedi, V., Rai, A., Pal, S., Shasany, A. K., et al. (2017). A WRKY transcription factor from *Withania somnifera* regulates triterpenoid withanolide accumulation and biotic stress tolerance through modulation of phytosterol and defense pathways. *New Phytol.* 215 (3), 1115–1131. doi: 10.1111/nph.14663

Sun, L., and Zhu, Z. (2021). The molecular basis of age-modulated plant *de novo* root regeneration decline in *Arabidopsis thaliana*. *Plant Cell Physiol.* 62 (1), 3–7. doi: 10.1093/pcp/pcaa134

Wang, P., Ma, L., Wang, S., Li, L., Wang, Q., Yang, R., et al. (2019). Identification and analysis of a candidate *WRKY transcription factor* gene affecting adventitious root formation using association mapping in *Catalpa scop*. *DNA Cell Biol.* 38 (4), 297–306. doi: 10.1089/dna.2018.4528

Wang, H., Pak, S., Yang, J., Wu, Y., Li, W., Feng, H., et al. (2022). Two high hierarchical regulators, *PuMYB40* and *PuWRKY75*, control the low phosphorus driven adventitious root formation in *Populus ussuriensis*. *Plant Biotechnol. J.* 20 (8), 1561–1577. doi: 10.1111/pbi.13833

Xu, M., Hu, T., Zhao, J., Park, M. Y., Earley, K. W., Wu, G., et al. (2016). Developmental functions of miR156-regulated *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL)* genes in *Arabidopsis thaliana*. *PloS Genet.* 12 (8), e1006263. doi: 10.1371/journal.pgen.1006263

Xu, X., Li, X., Hu, X., Wu, T., Wang, Y., Xu, X., et al. (2017). High miR156 expression is required for auxin-induced adventitious root formation *via MxSPL26* independent of *PINs* and *ARFs* in *Malus xiaojinensis*. *Front. Plant Sci.* 8, 1059. doi: 10.3389/fpls.2017.01059

Ye, B. B., Shang, G. D., Pan, Y., Xu, Z. G., Zhou, C. M., Mao, Y. B., et al. (2020a). AP2/ERF transcription factors integrate age and wound signals for root regeneration. *Plant Cell* 32 (1), 226–241. doi: 10.1105/tpc.19.00378

Ye, B. B., Zhang, K., and Wang, J. W. (2020b). The role of miR156 in rejuvenation in *Arabidopsis thaliana*. *J. Integr. Plant Biol.* 62 (5), 550–555. doi: 10.1111/jipb.12855

Zhang, Z. L., Xie, Z., Zou, X., Casaretto, J., Ho, T. H., and Shen, Q. J. (2004). A rice *WRKY* gene encodes a transcriptional repressor of the gibberellin signaling pathway in aleurone cells. *Plant Physiol.* 134 (4), 1500–1513. doi: 10.1104/pp.103.034967

Zhao, D., Wang, Y., Feng, C., Wei, Y., Peng, X., Guo, X., et al. (2020). Overexpression of *MsGH3.5* inhibits shoot and root development through the auxin and cytokinin pathways in apple plants. *Plant J.* 103 (1), 166–183. doi: 10.1111/tpj.14717

Zhu, H., Jiang, Y., Guo, Y., Huang, J., Zhou, M., Tang, Y., et al. (2021). A novel salt inducible *WRKY* transcription factor gene, *AhWRKY75*, confers salt tolerance in transgenic peanut. *Plant Physiol. Biochem.* 160, 175–183. doi: 10.1016/j.plaphy.2021.01.014

Zhu, W., Zhang, M., Li, J., Zhao, H., Zhang, K., and Ge, W. (2022). Key regulatory pathways, microRNAs, and target genes participate in adventitious root formation of *Acer rubrum* l. *Sci. Rep.* 12 (1), 12057. doi: 10.1038/s41598-022-16255-7

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 24 February 2023
DOI 10.3389/fpls.2023.1137630

OPEN ACCESS

EDITED BY
Chenxia Cheng,
Qingdao Agricultural University, China

REVIEWED BY
Xianyan Zhao,
Qilu University of Technology, China
Hongjian Wan,
Zhejiang Academy of Agricultural Sciences, China
Sun Xun,
Nanjing Agricultural University, China

*CORRESPONDENCE
Ming Tan
tanming@hebau.edu.cn
Jianzhu Shao
yysjz@hebau.edu.cn

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

SPECIALTY SECTION
This article was submitted to Functional and Applied Plant Genomics, a section of the journal Frontiers in Plant Science

RECEIVED 04 January 2023
ACCEPTED 07 February 2023
PUBLISHED 24 February 2023

CITATION
Li G, Li J, Zhang H, Li J, Jia L, Zhou S, Wang Y, Sun J, Tan M and Shao J (2023) ASSVd infection inhibits the vegetative growth of apple trees by affecting leaf metabolism.
Front. Plant Sci. 14:1137630.
doi: 10.3389/fpls.2023.1137630

ASSVd infection inhibits the vegetative growth of apple trees by affecting leaf metabolism

Guofang Li[1†], Jinghong Li[1†], He Zhang[1], Jiuyang Li[1], Linguang Jia[2], Shiwei Zhou[1], Yanan Wang[3], Jianshe Sun[1], Ming Tan[1*] and Jianzhu Shao[1*]

[1]College of Horticulture, Hebei Agricultural University, Baoding, China, [2]Changli Institute of Pomology, Hebei Academy of Agricultural and Forestry Science, Changli, China, [3]College of Plant Protection, Hebei Agricultural University, Baoding, China

Apple scar skin viroid (ASSVd) can infect apple trees and cause scar skin symptoms. However, the associated physiological mechanisms are unclear in young saplings. In this study, ASSVd-infected and control 'Odysso' and 'Tonami' apple saplings were examined to clarify the effects of ASSVd on apple tree growth and physiological characteristics as well as the leaf metabolome. The results indicated that leaf ASSVd contents increased significantly after grafting and remained high in the second year. Leaf size, tree height, stem diameter, branch length, and leaf photosynthetic efficiency decreased significantly in viroid-infected saplings. In response to the ASSVd infection, the chlorophyll a and b contents decreased significantly in 'Odysso', but were unchanged in 'Tonami'. Moreover, the N, P, K, Fe, Mn, and Ca contents decreased significantly in the leaves of viroid-infected 'Odysso' or 'Tonami'. Similarly, the CAT and POD contents decreased significantly in the viroid-infected saplings, but the SOD content increased in the viroid-infected 'Tonami' saplings. A total of 15 and 40 differentially abundant metabolites were respectively identified in the metabolome analyses of 'Odysso' and 'Tonami' leaves. Specifically, in the viroid-infected 'Odysso' and 'Tonami' samples, the L-2-aminobutyric acid, 6″-O-malonyldaidzin, and D-xylose contents increased, while the coumarin content decreased. These metabolites are related to the biosynthesis of isoflavonoids and phenylpropanoids as well as the metabolism of carbohydrates and amino acids. These results imply that ASSVd affects apple sapling growth by affecting physiological characteristics and metabolism of apple leaves. The study data may be useful for future investigations on the physiological mechanisms underlying apple tree responses to ASSVd.

KEYWORDS

apple, ASSVd, tree phenotype, photosynthesis, nutrient element, metabolome

Introduction

Apple scar skin viroid (ASSVd) can decrease the marketability of apple fruits, with detrimental effects on the apple industry. Specifically, ASSVd infections can decrease the size of apple fruits and lead to scarring and staining of the fruit peel (Koganezawa, 1985; Kim et al., 2010; Lee et al., 2017; Li et al., 2021). The incidence of ASSVd infections in the main apple-producing regions in Northern China, reportedly ranges from 4.8% to 48.6% (Zha et al., 2022). For the transmission of ASSVd, the most common mode is contact transmission (e.g., mechanical and agricultural operations) (Verhoeven et al., 2010). In terms of apple, the spread of scions, rootstocks, and saplings during asexual propagation is the main route for ASSVd transmission. It has been detected in various apple tissues (e.g., leaves, stems, skins, stocks, fruits, and roots) (Desvignes et al., 1999; Walia et al., 2014; Wu et al., 2015; Kim et al., 2019). However, symptoms of ASSVd infection are not appeared on the apple leaves and stems (Kim et al., 2010; Wu et al., 2015; Di Serio et al., 2018).

Replication of the ASSVd, which is a noncoding RNA molecule consisting of ~330 nucleotides, rely on host functional enzymes through an RNA-to-RNA rolling circle mechanism (Flores et al., 2009; Li et al., 2021). In order to establish systemic infection, these RNAs can traffic from an initially infected cell into neighboring cells and ultimately throughout a whole plant *via* plasmodesmata and phloem (Takeda and Ding, 2009). The method used for detecting ASSVd mainly involves the identification of botanical symptoms and the application of molecular techniques (Heo et al., 2019; Li et al., 2020). Because of the latent incubation period after ASSVd infection, and some infected plants lack obvious symptoms (Desvignes et al., 1999; Wang et al., 2000; Jiang et al., 2017), which may be difficult to identify ASSVd infection with botanical symptoms. The previously established quantitative real-time PCR (qRT-PCR) assay for detecting ASSVd is more sensitive and rapid than the alternative methods (e.g., conventional RT-PCR, spot hybridization, and botanical identification) (Wu et al., 2015; Xi et al., 2020; Heo and Chung, 2021). Plant metabolomics research mainly involves qualitative and quantitative analyses of metabolic processes under certain stress conditions to elucidate specific biological activities (Han et al., 2019). For example, the examination of the phloem of Verticillium wilt-infected mulberry plants revealed changes in various sugar, amino acid, and organic acid metabolites (Gai et al., 2014). The changes in the phloem sap composition were greater than the changes detected in leaves. The analysis of the poplar metabolome indicated several metabolites, such as p-Octylphenol, plant alcohol, catechol, and eugenol, are related to rust resistance (Zhou et al., 2013). Hence, exploring the apple metabolome may lead to the identification of metabolites responsive to an ASSVd infection.

In this study, 'Odysso' and 'Tonami' apple saplings were selected as materials for an investigation of the ASSVd content in growing saplings and its influence on young sapling physiological characteristics. Furthermore, leaf metabolomics experiments were conducted to systematically evaluate the effects of ASSVd on young sapling growth and physiological characteristics. The generated data were used to elucidate the mechanism by which ASSVd damages apple saplings. The findings of this study may provide the theoretical basis for future studies conducted to clarify the mechanism underlying ASSVd infections, while also highlighting the necessity of virus-free cultivation.

Materials and methods

Plant materials

This study was carried out at the fruit demonstration station of Hebei Agricultural University, Shunping county, Hebei province (38.97N, 114.92E), China. The newly grafted (NG) axillary buds and 1-year-old (OY) saplings examined in this study were from 'Odysso' (red-fleshed apple fruits with high anthocyanin contents bred by Markus Kobelt in Switzerland) and 'Tonami' (*Malus domestica* 'Tonami') apple cultivars. The virus-free *M. robusta* rootstocks used for grafting were rapidly propagated in a tissue culture system. Viroid-infected 'Odysso' was obtained by grafting onto 'Tonami' apple trees infected with ASSVd. Axillary buds on the annual branches of all scions were collected at the Hebei Agricultural University nursery in early March. The 'Odysso' and 'Tonami' saplings infected with ASSVd (i.e., treatments) were compared with the virus-free control 'Odysso' and 'Tonami' saplings. The grafted saplings were transferred to 45 cm plastic pots. Ten saplings were used per treatment. All saplings were grown under consistent conditions.

Quantitative real-time PCR analysis of ASSVd and production of a standard curve

Total RNA was extracted from sapling leaves using the polysaccharide polyphenol plant total RNA extraction kit (Tiangen, Beijing). The integrity, purity, and concentration of the isolated RNA were determined on the basis of the A260/A280 and A260/A230 ratios using the NanoDrop 2000 spectrophotometer (Thermo Scientific, USA). The RNA served as the template for synthesizing cDNA using the TransScript One-Step gDNA Removal and cDNA Synthesis SuperMix kit (TransGen Biotech, Beijing). The qRT-PCR assays were completed using the LightCycler® 96 PCR system (Roche, Switzerland) as previously described (Wu et al., 2015). We performed qRT-PCR with a SYBR Premix Ex Taq II kit (Takara, Beijing) in a 25-μl reaction system. The cycling protocol consisted of 95 °C for 30 s, followed by 39 cycles of 95 °C for 15 s, 60 °C for 20 s and 72 °C for 20 s, followed by 39 cycles to construct a melting curve. The ASSVd sequence-specific primers were developed in our previous studies (Table S1, Chen et al., 2012; Wu et al., 2015).

To obtain cRNA, a 106 bp ASSVd fragment was recovered and inserted into the pWASY-T3 vector (TransGen Biotech, Beijing). The recombinant plasmid was transferred into DH5α cells (TransGen Biotech, Beijing), which were screened on LB solid medium containing IPTG and X-gal. The white colonies were subsequently transferred to LB medium supplemented with Ampicillin to obtain single colonies, which were analyzed by

PCR. Plasmids were extracted from the colonies with the correct sequence. The extracted plasmid was digested, after which the T7 *in vitro* Transcription Kit was used to generate cRNA.

The standard cRNA solution (1.24×10^{12} copies μl^{-1}) was serially diluted 10-fold (1.24×10^{5}–1.24×10^{11} copies μl^{-1}) for the qRT-PCR conducted to obtain the corresponding linear equation. The data for the biological replicates and replicate tests were used to construct a standard curve, with the cRNA concentration (logarithm value) on one axis and the cycle threshold (Ct) value on the other axis. To assess the reliability of the standard curve, the standard deviation (SD) and variable coefficient (VC) were calculated following the statistical analysis of the Ct values.

Measurement of tree growth

The tree height was measured from the grafting site using steel tape, whereas the stem diameter was measured at 5 cm above the grafting site using vernier calliper. A portable leaf area meter CI-203 (CID BioScience, USA) and ruler were used to measure leaf parameters. These measurements were performed from April to November for 2 years.

After the leaves fell, the branch compositions were analyzed and the lengths of the short (≤5 cm), medium (5-15 cm), long (15-30 cm), and super-long shoots (>30 cm) were recorded.

Measurement of photosynthetic parameters and chlorophyll contents in leaves

The net photosynthetic rate (Pn), stomatal conductivity (Gs), transpiration rate (Tr), and intercellular CO_2 concentration (Ci) of the OY saplings were measured using the LI-6400 Portable Photosynthesis System (Li-Cor, Lincoln NE, USA) on August 10. Five healthy and mature leaves were selected for each treatment. The diurnal changes were measured once every 2 h from 6:00 to 18:00. According to a modified version of an established method (Lichtenhaler and Wellburn, 1983), apple leaves were placed in 10 ml tubes. After adding 5 ml 80% acetone, the tubes were incubated at 4 °C in a black box. The samples were mixed several times every 4 h. As the color intensity of the solutions decreased, the chlorophyll content per unit area was determined by measuring the absorbance at 663 and 645 nm.

Determination of nutrient elements and antioxidant activities in leaves

For each treatment, 30 healthy and mature leaves were selected, washed, and dried. Following a 30-min incubation at 105 °C, the leaves were dried at 75 °C, crushed, and screened. The ground leaves were collected in a self-sealing bag and stored in a dryer. Leaf samples were boiled in H_2SO_4–$HClO_4$ and HNO_3–$HClO_4$ solutions and then the nutrient element contents were determined using AutoAnalyzer 3 with XY-2 Sampler (SEAL Analytical, UK) and an inductively coupled plasma emission spectrometer (Prodigy Spec, Teledyne, USA).

On August 10, 10 randomly selected leaves (per treatment) were obtained from OY saplings to measure the catalase (CAT), peroxidase (POD), and superoxide dismutase (SOD) contents as previously described (Wang et al., 2018). CAT was determined by monitoring the decomposition of H_2O_2. SOD was assayed by monitoring the inhibition of the photochemical reduction of nitro blue tetrazolium. POD was determined by monitoring the oxidation reaction of guaiacol.

Extraction, sampling, and analysis of leaf metabolites

Six leaves were collected from 'Odysso' and 'Tonami' OY saplings, frozen in liquid nitrogen, and stored at −80 °C prior to the leaf metabolomics analysis performed by Novogene Bioinformatics Technology Co. Ltd. The analysis was completed using six biological replicates. Briefly, leaf (100 mg) extracts were prepared using a mixture (methanol: acetonitrile: nitrile water=2:2:1, v/v), after which the 100 µl mixture (acetonitrile: water =1:1, v/v) was added to redissolve the extract. The supernatant was retained for the analysis. Samples were separated using the Agilent 1290 ultra-high performance liquid chromatography system (Agilent, USA) and then analyzed using the TripleTOF 5600 mass spectrometer (AB SCIEX, USA). Accurate quality numbers and secondary spectrograms were used to identify and annotate metabolites according to the Novogene source database.

To evaluate the similarity among biological replicates and the differences between samples, a correlation heatmap was constructed and a principal component analysis (PCA) and a partial least squares discriminant analysis (PLS-DA) were performed on the basis of the metabolite profiles (Lu et al., 2017). A variable importance in projection (VIP) value >1 and $P < 0.05$ were set as the criteria for identifying significant differentially abundant metabolites between samples (Lu et al., 2017; Han et al., 2021). Figures for the heatmaps, volcano maps, Venn diagram, and KEGG pathways were obtained using the Majorbio Cloud Platform (www.majorbio.com).

Statistical analysis

Significant differences were detected using one-way ANOVA ($P < 0.05$), which was performed using the IBM SPSS Statistics 20. Values in figures represent the mean ± SD. The figures presenting physiological and biochemical data were produced using Microsoft Office Home and Student 2019, whereas the figures presenting the metabolomics data were generated using the Majorbio Cloud Platform.

Results

Detection and quantification of ASSVd in 'Odysso' and 'Tonami' leaves

The presence of ASSVd in the leaves of 'Odysso' and 'Tonami' apple saplings was determined on the basis of gene cloning and a

qRT-PCR analysis. The results indicated ASSVd was undetectable in the control 'Odysso' and 'Tonami' saplings, but it was present in the leaves of infected 'Odysso' and 'Tonami' saplings (Figure S1A). The cloned sequences were analyzed using NCBI BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi), which revealed substantial similarities (97.28%–99.09%) to 39 aligned sequences (the most similar sequence was X71599.1; Figure 1A, Figure S1A). For the quantitative analysis of the viroid levels, the correlation coefficient (R^2) was 0.998, the amplification efficiency was 90%, and the maximum coefficient of variation among the biological replicates and the replicated tests was 1.35%–1.81% (Figure 1A, Table S2), which reflected the reliability of the standard curve.

In viroid-infected 'Odysso' and 'Tonami' NG saplings, the initial ASSVd content was low, but it increased significantly starting on July 15 (approximately the start of the autumn growing season) for both 'Tonami' and 'Odysso' (Figure 1B). In the viroid-infected 'Odysso' and 'Tonami' OY saplings, the ASSVd content decreased from April 15 to July 15 (after sprouting) and increased from July 15 to October 15 (autumn growing season) (Figure 1B). Among the OY saplings, the viroid level was higher in 'Tonami' than in 'Odysso', except on April 15 (Figure 1B).

Effects of ASSVd on apple sapling leaf size

To quantify the effects of ASSVd on leaf size, leaf area, length, width and perimeter were analyzed in the leaves of NG and OY saplings. Compared with the control, the leaf area and leaf length of the viroid-infected 'Odysso' saplings decreased significantly by 11.30% and 12.27% (NG) and 31.95% and 10.63% (OY), respectively (Figure 2). The leaf width did not differ between NG and OY. There was also no difference in the NG leaf circumference between the viroid-infected and control 'Odysso' saplings. In contrast, the OY leaf circumference of the viroid-infected 'Odysso' saplings decreased significantly by 19.99% (Figure 2). Compared with control NG saplings, the leaf area and leaf circumference of the viroid-infected 'Tonami' saplings decreased significantly by 18.57% and 10.63%, respectively, but there were no significant differences in the leaf length and leaf width (Figure 2A). In terms of the OY samples, the leaf area, length, width, and circumference of the viroid-infected 'Tonami' saplings decreased significantly by 20.15%, 9.95%, 12.42%, and 12.26%, respectively (compared with the control leaves) (Figure 2B). These results indicated that ASSVd can significantly affect the growth of 'Odysso' and 'Tonami' leaves.

Effects of ASSVd on the stem height and diameter, and shoot length of grafted scions

To clarify the effects of ASSVd on shoot growth, stem height and diameter, and shoot length were analyzed in NG and OY saplings. The 'Odysso' and 'Tonami' NG saplings grew to a similar height. Both cultivars grew rapidly from May to September, but the growth rate decreased after September (Figure 3A). In November,

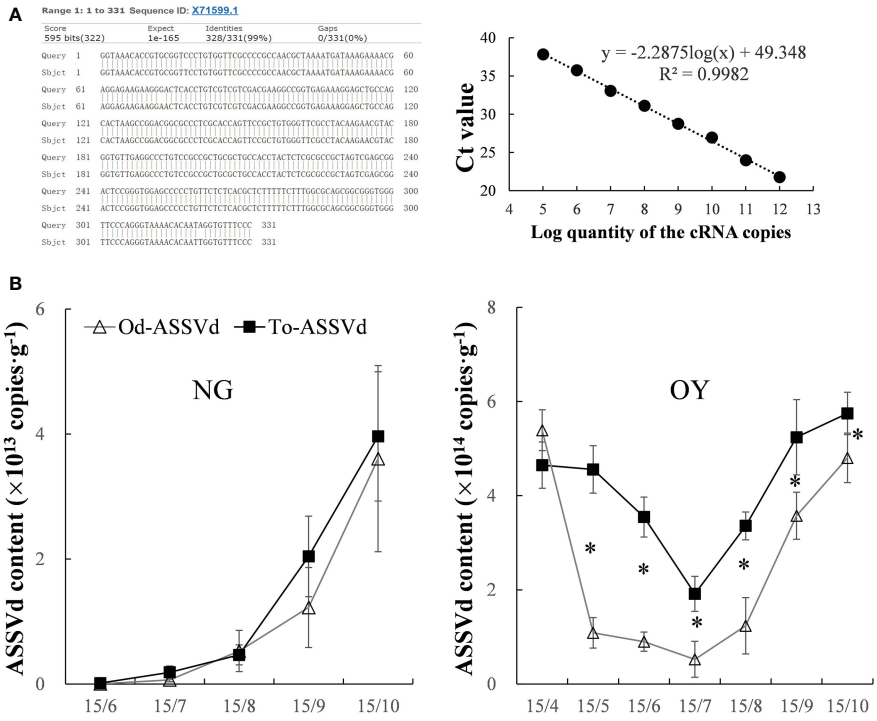


FIGURE 1
Detection and quantification of ASSVd in leaves. **(A)** Alignment of the most similar sequences revealed by the NCBI BLAST search and standard curve for the qRT-PCR analysis. **(B)** Dynamic changes in the ASSVd content of the NG and OY leaves. Data are presented as the mean ± SD (n=3). The asterisk indicates a significant difference as determined by Student's *t*-test ($P < 0.05$).

10.3389/fpls.2023.1137630

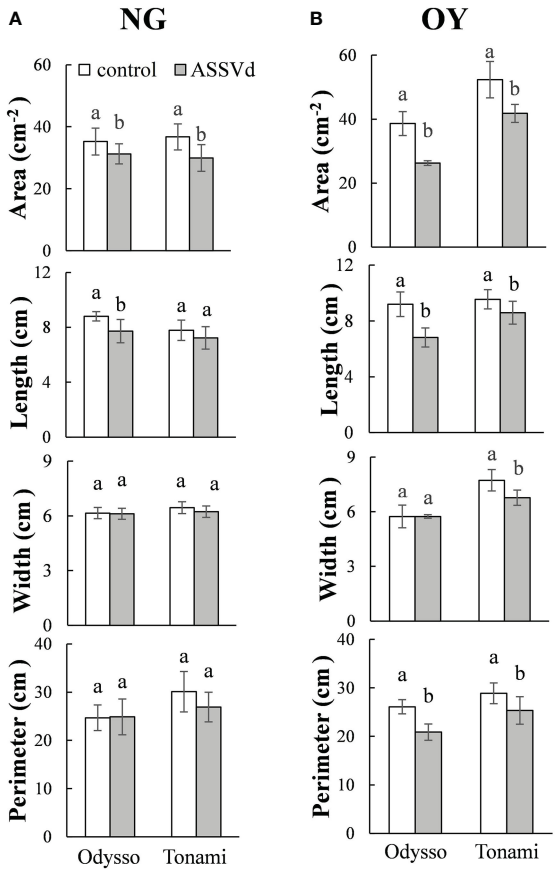


FIGURE 2
Effects of ASSVd on the size of 'Odysso' and 'Tonami' leaves. Leaf area, length, width, and perimeter in NG **(A)** and OY **(B)** saplings. Data are presented as the mean ± SD (n=5). Lowercase letters indicate significant differences as determined by Student's *t*-test ($P < 0.05$).

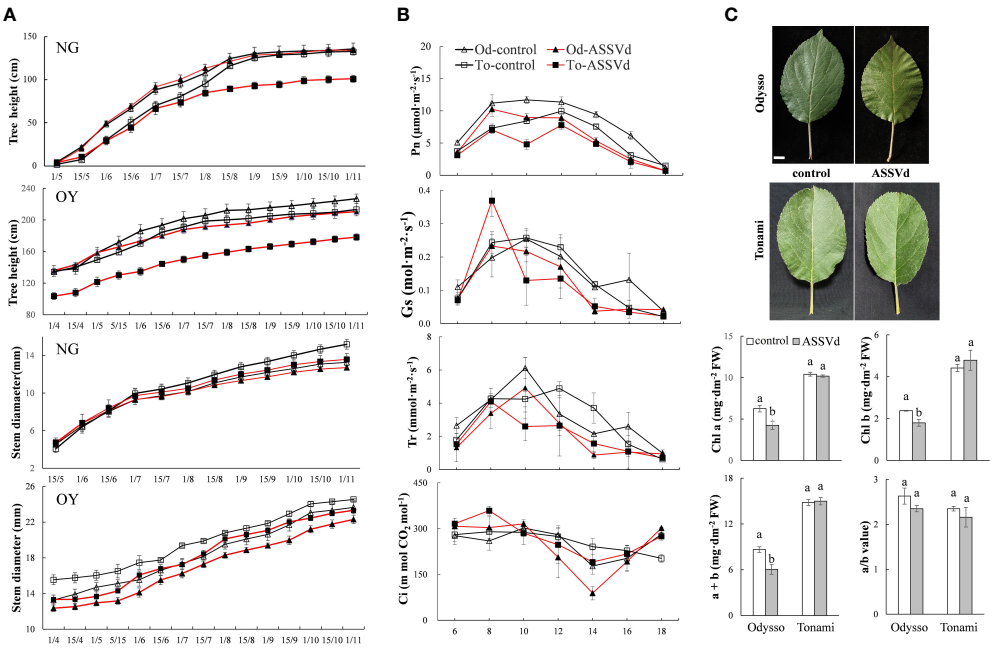


FIGURE 3
Effects of ASSVd on tree height, stem diameter, photosynthetic activity, and chlorophyll content. **(A)** Tree height and stem diameter curves for the NG and OY saplings of 'Odysso' and 'Tonami'. **(B)** Diurnal variation in the net photosynthetic rate (Pn), stomatal conductance (Gs), transpiration rate (Tr), and intercellular CO_2 concentration (Ci) in OY leaves. **(C)** Leaf phenotypes and chlorophyll contents in OY leaves. Data are presented as the mean ± SD (n=5). Scale bar = 1.0 cm. Lowercase letters indicate significant differences as determined by Student's t-test ($P < 0.05$).

there were no significant differences between the control and viroid-infected 'Odysso' NG saplings (Figure 3A). Before August 1, the differences in the sapling height were insignificant between the control and viroid-infected 'Tonami' samples (Figure 3A). However, the growth rate and height were significantly greater for the control saplings than for the viroid-infected 'Tonami' saplings after August 15 (Figure 3A). The increase in 'Odysso' and 'Tonami' plant height decreased significantly on September 1 and there was essentially no increase by November (Figure 3A). Thus, the viroid-infected 'Tonami' saplings were significantly shorter than the other three saplings. The comparison between OY and NG indicated the changes in sapling height were generally the same for 'Odysso' and 'Tonami' (Figure 3A). For NG and OY, the stem diameters were similar between 'Odysso' and 'Tonami' and continued to increase throughout the growing season (Figure 3A). The OY sapling stem diameter differed only between the control and viroid-infected 'Odysso' samples (Figure 3A). However, the NG sapling stem diameter was greater for the control 'Tonami' samples than for the viroid-infected 'Tonami' samples after August (Figure 3A). For the NG and OY saplings, the total branch length (i.e., short, medium, long, and super-long branches and scion height) of the viroid-infected 'Odysso' and 'Tonami' scions decreased significantly (compared with the control saplings) (Table S3).

Effects of ASSVd on photosynthetic efficiency and the chlorophyll content

The Pn value differed between the viroid-infected and control saplings (Figure 3B). In the controls, Pn was highest in 'Odysso' and 'Tonami' leaves at 8:00–12:00 and 12:00, respectively. For the viroid-infected 'Tonami' samples, Pn was lower at 10:00 than at 8:00 or 12:00. In the viroid-infected 'Odysso' samples, Pn peaked at 8:00. Generally, Pn was higher in the controls than in the viroid-infected saplings, although Pn was high for both the viroid-infected and control 'Odysso' saplings. The Gs value indicates the degree of stomatal opening, which influences the rate of CO_2 and water vapor exchange. The daily variation curve indicated Gs initially increased and then decreased (Figure 3B). The Gs was highest at 8:00 for the viroid-infected 'Tonami' saplings, whereas it was highest at 10:00 for the other three materials. In addition, Gs was lower for the viroid-infected 'Odysso' and 'Tonami' samples than for the corresponding controls at 10:00–16:00. Leaf transpiration is critical for the upward transport of water through plants. The changes in Tr were similar to the Gs trends (Figure 3B). The Tr value was highest for the viroid-infected 'Tonami' saplings at 8:00, but it peaked for the 'Odysso' saplings (control and viroid-infected) and the control 'Tonami' saplings at 10:00 and 12:00, respectively. Additionally, for both 'Odysso' and 'Tonami', Tr was lower for the viroid-infected samples than for the controls at 12:00–16:00. The Ci values fluctuated over the analyzed time-period (Figure 3B). The Ci value for the viroid-infected 'Tonami' samples was highest at 8:00, whereas it peaked at 10:00 and then decreased significantly to its lowest point at 14:00 for the viroid-infected 'Odysso' samples. Chlorophyll a and b are important photosynthetic pigments. Compared with the controls, the chlorophyll a and b contents were significantly higher in the viroid-infected 'Odysso' samples, but the chlorophyll a/b ratio was unchanged (Figure 3C). In 'Tonami', the ASSVd infection did not significantly affect the chlorophyll content and chlorophyll a/b ratio.

The above results showed that the effects of ASSVd on leaf photosynthetic activities and chlorophyll contents differed between 'Odysso' and 'Tonami'. More specifically, the ASSVd infection affected the chlorophyll content in 'Odysso', whereas it decreased the photosynthetic efficiency in 'Tonami'.

Effect of ASSVd on nutrient elements and antioxidant activities in leaves

Nutrient elements are essential for plant growth and development as well as the composition of plant tissues and organs. Compared with the control saplings, the N, Mg, Cu, and Mn contents in the viroid-infected 'Odysso' saplings decreased significantly by 46.63%, 10.98%, 9.43%, and 5.41%, respectively, while the B content increased significantly by 6.96% (Figure 4A). In the 'Tonami' leaves infected with ASSVd, the N, P, K, Mg, Mn, and Ca contents decreased by 46.58%, 8.47%, 25.89%, 29.65%, 8.29%, and 14.75%, respectively, whereas the Fe content increased by 30.18% (Figure 4A). The rank-order of the decrease in the nutrient element contents was N > Mg > K > Cu > Ca > Mn > P > Zn > Fe > B for viroid-infected 'Odysso' and N > Mn > K > Ca > P > Mg > Zn > B > Cu > Fe for viroid-infected 'Tonami' (Figure 4A).

Compared with the controls, the leaf CAT and POD contents were lower for the viroid-infected saplings (Figure 4B), suggestive of a decrease in oxidative stress resistance. However, the SOD content was lower in the control 'Tonami' saplings than in the viroid-infected 'Tonami' samples (Figure 4B).

Effects of ASSVd on the leaf metabolome

To characterize the effects, the content and species of metabolites in the leaves from 'Odysso' and 'Tonami' OY saplings were analyzed. Of the 12,751-pos and 12,465-neg peaks detected during the metabolomics analysis of the viroid-infected and control saplings, 136-pos and 72-neg metabolites were identified (Table S4). According to the correlation heatmaps based on positive and negative ions, the correlation coefficients between biological replicates ranged from 0.91 to 0.96 and 0.89 (two replicates of the control 'Odysso') to 0.96 for the same samples (Figure 5A, B). On the basis of the PCA and PLS-DA data, the confidence level of six biological replicates for each sample exceeded 95%. Additionally, different samples could be distinguished (Figure 5A, B). These results indicated the high-quality metabolomics data were suitable for the subsequent analysis.

A Venn diagram revealed 15 and 40 differentially abundant metabolites between the viroid-infected 'Odysso' and 'Tonami' samples and the corresponding controls, respectively (Figure 5C, Table S5). More differentially abundant metabolites were detected by the comparisons between the two viroid-infected and two control saplings. Furthermore, eight and seven up-regulated and

10.3389/fpls.2023.1137630

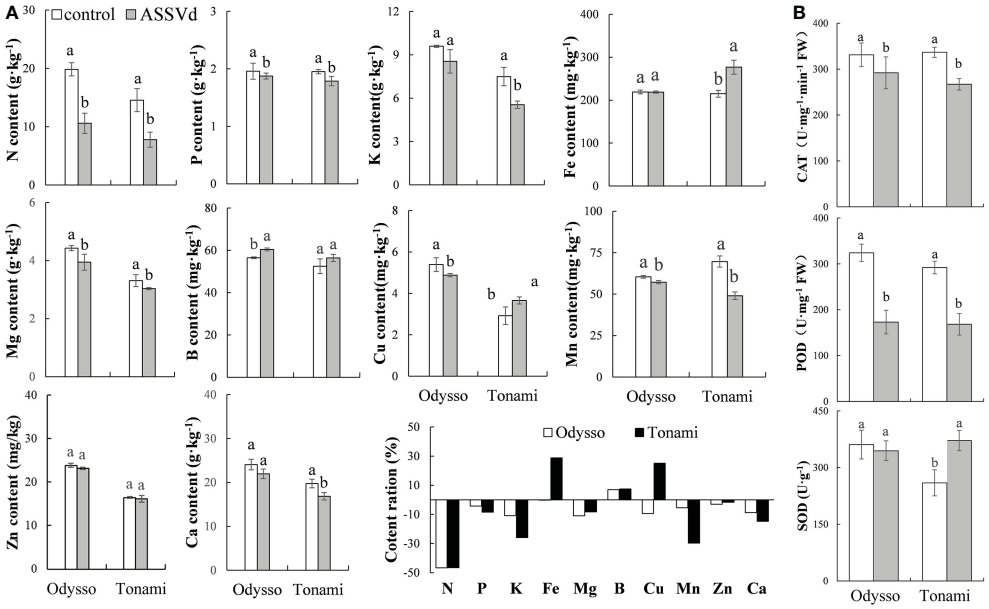


FIGURE 4
Effects of ASSVd on the nutrient element and antioxidant enzyme contents in leaves. **(A)** N, P, K, Fe, Mg, B, Cu, Mn, Zn, and Ca contents and ratios in the OY leaves of 'Odysso' and 'Tonami'. **(B)** CAT, POD, and SOD contents in OY leaves. Data are presented as the mean ± SD (n=3). Lowercase letters indicate significant differences as determined by Student's *t*-test ($P < 0.05$).

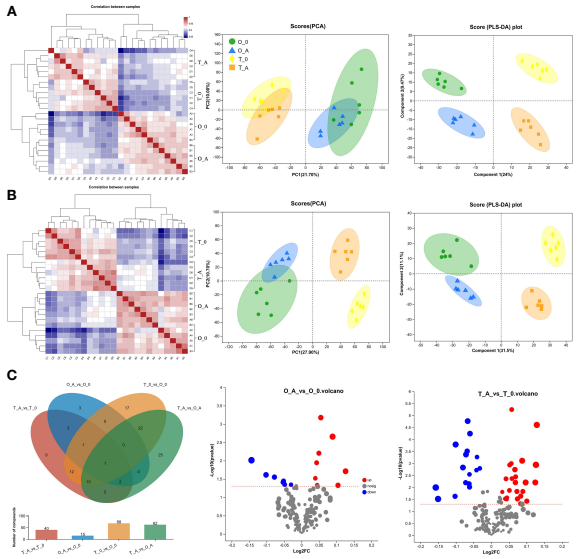


FIGURE 5
Heatmap, PCA and PLS-DA scores, and differentially abundant metabolites in leaves. Analysis of the positive **(A)** and negative **(B)** metabolites. In the heatmap, each lattice represents the correlation between two samples. The different colors represent the correlation coefficient between samples, whereas the length of the cluster branch represents the relative distance between samples. Each dot in the PCA and PLS-DA score maps represents the leaf metabolomic profile of a single sample. **(C)** In the Venn diagram, different colors represent different comparisons. Overlapping and non-overlapping numbers respectively represent the number of shared and unique metabolites. The bar graphs present the number of metabolites in the corresponding comparisons. In the volcano plot, the abscissa presents the fold-change in the metabolite abundance between two samples, whereas the ordinate presents the significance of the difference. Red, blue, and gray dots represent the up-regulated, down-regulated, and unchanged metabolites, respectively.

down-regulated metabolites, respectively, were identified in 'Odysso', whereas 24 and 16 up-regulated and down-regulated metabolites, respectively, were identified in 'Tonami' (Figure 5C). The metabolite cluster analysis divided the metabolites into 10 subclusters. The metabolites in subclusters 1, 2, 4, and 7 clearly differed among samples (Figure 6). More specifically, the abundance of the metabolites in subclusters 10 and 6 increased in 'Odysso' and 'Tonami' in response to the ASSVd infection, whereas the opposite trend was observed for the subcluster 8 metabolites (Figure 6). These findings confirmed the metabolomes of the two apple varieties vary, but the ASSVd infection-induced changes in the contents of some metabolites were consistent between 'Odysso' and 'Tonami'.

Compared with the controls, the L-2-aminobutyric acid, 6″-O-malonyldaidzin, and D-xylose contents increased in the viroid-infected 'Odysso' and 'Tonami' samples, which was in contrast to the decrease in the coumarin content (Figure 7A). The VIP values of these four metabolites were significantly greater than 1, indicative of significant differences between the viroid-infected and control saplings (Figure 7A). The KEGG enrichment analysis revealed the differentially abundant metabolites between the viroid-infected and control saplings were significantly associated with lipid metabolism, carbohydrate metabolism, the biosynthesis of other secondary metabolites, amino acid metabolism, and membrane transport (Figure 7B).

Discussion

ASSVd affects apple sapling growth, photosynthetic activity, and nutrient element contents

Symptoms of an ASSVd infection are not appeared on the apple leaves and stems (Kim et al., 2010; Wu et al., 2015; Di Serio et al., 2018). However, leaf curling and leaf discoloration (e.g., faded and mottled) are included among the symptoms on litchi plants infected with ASSVd; however, these symptoms may not develop in some varieties (Jiang et al., 2017). An infection by ASSVd may weaken persimmon trees (Ito et al., 2013). In the current study, prior to detecting and quantifying the ASSVd content, it was difficult to determine the cause of the phenotypic changes in the examined saplings. In the viroid-infected 'Odysso' and 'Tonami' saplings, the decrease in Pn, Gs, and Tr may have been primarily due to the decrease in the chlorophyll content in the 'Odysso' leaves and the decline in photosynthetic efficiency in the 'Tonami' leaves (no change in the chlorophyll content) (Wang et al., 2018). These observations indicate that ASSVd infections can decrease the chlorophyll content or the photosynthetic efficiency in apple leaves, which may explain the decreased tree height, stem diameter, and total branch length detected for the viroid-infected 'Odysso' and 'Tonami' samples. In addition, Mg, Fe, and Cu are involved in the synthesis of chlorophyll (Gong et al., 2020). Compared with the 'Odysso' leaves, the Fe and Cu contents were significantly higher in the 'Tonami' leaves, which may help to explain the observed lack of change in the chlorophyll content.

The nutrient status of plants reportedly influences disease resistance (Williams and Salt, 2009). In the present study, the N, P, K, Mg, Mn, Zn, and Ca contents were significantly lower in the ASSVd-infected leaves than in the control leaves. Insufficient amounts of some of these nutrient elements may adversely affect the uptake of the other elements, thereby exacerbating the detrimental effects of ASSVd on plants (Gong et al., 2020). Several studies confirmed that these elements are involved in the biosynthesis of certain compounds, enzymatic reactions, and metabolic pathways (Liang et al., 2017; Gong et al., 2020). Accordingly, the ASSVd-induced changes in the contents of these elements may have contributed to the inhibited growth of the viroid-infected 'Odysso' and 'Tonami' saplings.

ASSVd affects the leaf metabolomes of young apple saplings

Metabolomics-based analyses may identify important metabolites in specific physiological periods or conditions (Goodacre et al., 2004). Metabolic changes may be closely related to plant physiological performance. Recently reported changes in metabolites have helped to elucidate plant responses to various stresses (Moussaieff et al., 2013; Booth et al., 2015; Pidatala et al., 2016; Han et al., 2022). An infection by ASSVd may affect multiple tissues by altering nutrient accumulation and distribution throughout the plant. In apple saplings, the leaves are the main organs in which nutrients are synthesized, but they are also the primary tissues affected by ASSVd. In this study, 15 and 40 differentially abundant metabolites were identified in the 'Odysso' and 'Tonami' leaves, respectively, which implies that ASSVd has relatively limited effects on leaf metabolites. According to the KEGG analysis, most of the differentially abundant metabolites are associated with lipid, carbohydrate, and amino acid metabolism and the biosynthesis of other secondary metabolites. Thus, an ASSVd infection of 'Odysso' and 'Tonami' leaves significantly modulates specific metabolic pathways.

In the pathway mediating starch and sucrose metabolism, D-xylose is an intermediate during the synthesis of sucrose, which serves as a source of energy for various plant metabolic pathways and activities (Pan et al., 2016). We observed that the D-xylose content increased significantly in the 'Odysso' and 'Tonami' leaves, which reflected enhanced respiration and a decrease in photosynthetic productivity. A network analysis identified three metabolites (L-2-aminobutyric acid, 6″-O-malonyldaidzin, and coumarin) associated with signal transduction or stress resistance pathways. These changes in metabolite contents in 'Odysso' and 'Tonami' might be part of a specific response to an ASSVd infection. The differences in the phenotypes, photosynthetic activities, nutrient element contents, and metabolite contents between 'Odysso' and 'Tonami' suggest different apple varieties vary regarding their response to ASSVd. Compared with other apple varieties, the ASSVd-infected 'Odysso' saplings in this study had a lower ASSVd content, longer branches, fewer differentially abundant metabolites, and a greater anthocyanin content (i.e., 30% higher) (Pomiferous.com). However, determining the main

10.3389/fpls.2023.1137630

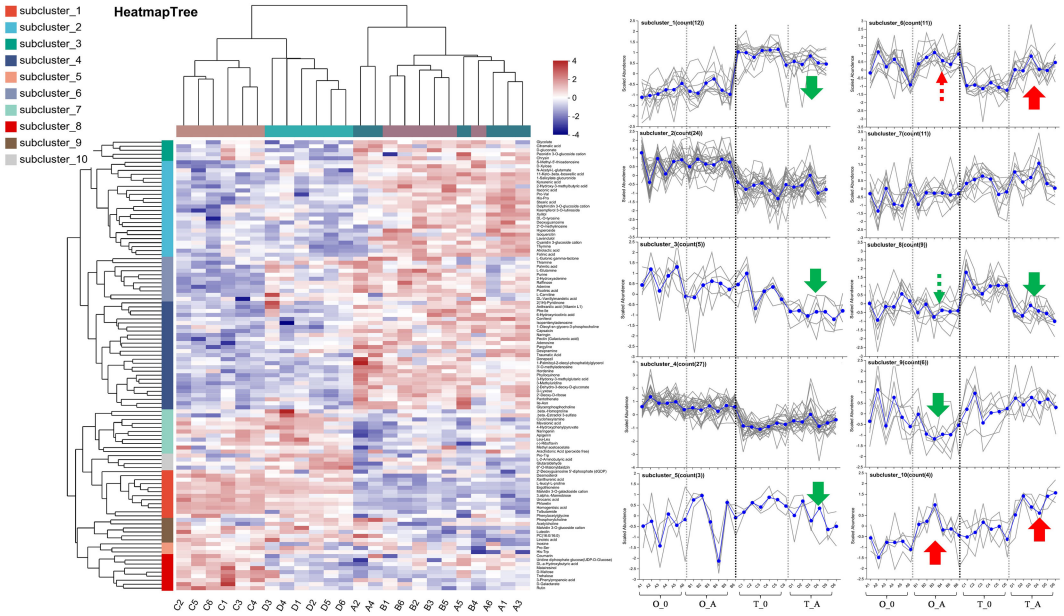


FIGURE 6
Cluster analysis of the leaf metabolites affected by the ASSVd infection. Each column represents a sample (code), whereas each row represents a metabolite (name). The color indicates the relative metabolite abundance. The dendrograms of the metabolite subclusters and sample clusters are presented on the left and at the top, respectively. The metabolite abundance in each subcluster was analyzed using vertical and horizontal coordinates representing the samples and the scaled abundance of metabolites, respectively. Green and red arrows indicate up-regulated and down-regulated metabolites, respectively. The solid and dashed lines indicate significant changes and certain trends, respectively.

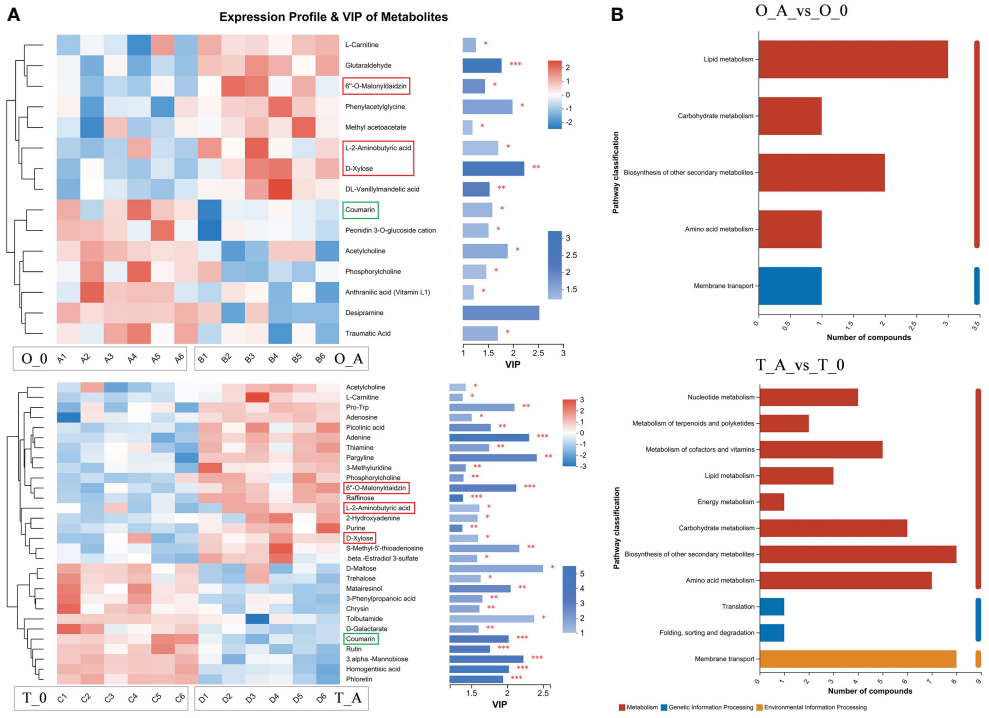


FIGURE 7
Heatmap and KEGG analysis of differentially abundant metabolites. **(A)** Left: metabolite cluster dendrogram, with each column representing one sample and each row representing one metabolite; the color indicates the relative metabolite abundance. Right: bar length indicates the contribution of the metabolite to the difference between samples. The blue bar indicates the significance of the differences in the metabolites between samples (*P < 0.05, **P < 0.01, and ***P < 0.001). **(B)** The ordinate presents the KEGG metabolic pathways, whereas the abscissa presents the number of metabolites. Red, blue, and yellow bars indicate the pathways belong to the Metabolism, Genetic Information Processing, and Environmental Information Processing categories.

common effects of ASSVd on apple trees may provide the theoretical basis for characterizing the mechanism mediating ASSVd infections and disease symptom development.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

Author contributions

GFL, JHL, HZ, YNW, MT, and JZS participated in the data analysis. JHL, LGJ, SWZ, JYL, JSS and JZS performed material sampling, field measurements and the laboratory data measurement. GFL, JHL, YNW, JSS, MT and JZS participated in the paper writing and manuscript amending. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Program of Apple Dwarfing Cultivation in the National Apple Industry Technology System of Agriculture Ministry of China (CARS-27), National Natural Science Foundation of China (31901981, 32002068), Scientific Research Project of Hebei Agricultural University (YJ2020053), Key Research and Development Program of the Ministry of Science and Technology (2019YFD1001400), and Apple Innovation Team of Modern Agricultural Industrial Technology System of Hebei Province (HBCT2021100409).

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1137630/full#supplementary-material

References

Booth, S. C., Weljie, A. M., and Turner, R. J. (2015). Metabolomics reveals differences of metal toxicity in cultures of pseudomonas pseudoalcaligenes KF707 grown on different carbon sources. *Front. Microbiol.* 6. doi: 10.3389/fmicb.2015.00827

Chen, H. X., Shao, J. Z., Sun, J. S., Cao, X. F., and Qiao, X. H. (2012). Quick detection of apple latent viruses by two-step multiplex RT-PCR assay. *J. Fruit Sci.* 29, 695–701. doi: 10.13925/j.cnki.gsxb.2012.04.008

Desvignes, J. C., Grasseau, N., Boye, R., Cornaggia, D., Aparicio, F., Di Serio, F., et al. (1999). Biological properties of apple scar skin viroid: isolates, host range, different sensitivity of apple cultivars, elimination, and natural transmission. *Plant Dis.* 83, 768–772. doi: 10.1094/PDIS.1999.83.8.768

Di Serio, F., Ambros, S., Sano, T., Flores, R., and Navarro, B. (2018). Viroid diseases in pome and stone fruit trees and koch's postulates: A critical assessment. *Viruses-basel* 10, 612. doi: 10.3390/v10110612

Flores, R., Gas, M. E., Molina-Serrano, D., Nohales, M. A., Carbonell, A., Gago, S., et al. (2009). Viroid replication: Rolling-circles, enzymes and ribozymes. *Viruses-basel* 1, 317–334. doi: 10.3390/v1020317

Gai, Y. P., Han, X. J., Li, Y. Q., Yuan, C. Z., Mo, Y. Y., Guo, F. Y., et al. (2014). Metabolomic analysis reveals the potential metabolites and pathogenesis involved in mulberry yellow dwarf disease. *Plant Cell Environ.* 37, 1474–1490. doi: 10.1111/pce.12255

Gong, Z. Z., Xiong, L. M., Shi, H. Z., Yang, S. H., Herrera-Estrella, L. R., Xu, G. H., et al. (2020). Plant abiotic stress response and nutrient use efficiency. *Sci. China-Life Sci.* 63, 635–674. doi: 10.1007/s11427-020-1683-x

Goodacre, R., Vaidyanathan, S., Dunn, W. B., Harrigan, G. G., and Kell, D. B. (2004). Metabolomics by numbers: acquiring and understanding global metabolite data. *Trends Biotechnol.* 22, 245–252. doi: 10.1016/j.tibtech.2004.03.007

Han, T., Liang, Y. P., Wu, Z. N., Zhang, L., Liu, Z. W., Li, Q. F., et al. (2019). Effects of tetracycline on growth, oxidative stress response, and metabolite pattern of ryegrass. *J. Hazard. Mater.* 380, 120885. doi: 10.1016/j.jhazmat.2019.120885

Han, T., Mi, Z. R., Chen, Z., Zhao, J. J., Zhang, H. G., Lv, Y., et al. (2022). Multi-omics analysis reveals the influence of tetracycline on the growth of ryegrass root. *J. Hazard. Mater.* 435, 129019. doi: 10.1016/j.jhazmat.2022.129019

Han, T., Sun, M. Y., Zhao, J. J., Dai, C. Y., Li, Y., Zhang, P., et al. (2021). The roles of cadmium on growth of seedlings by analysing the composition of metabolites in pumpkin tissues. *Ecotoxicol. Environ. Saf.* 226, 112817. doi: 10.1016/j.ecoenv.2021.112817

Heo, S., and Chung, Y. S. (2021). Rapid real–time detection method of ACLSV and ASSVd for apple quarantine field. *Plant Biotechnol. Rep.* 15, 187–195. doi: 10.1007/s11816-021-00663-8

Heo, S., Kim, H. R., and Lee, H. J. (2019). Development of a quantitative realtime nucleic acid sequence based amplification (NASBA) assay for early detection of apple scar skin viroid. *Plant Pathol. J.* 35, 164–171. doi: 10.5423/PPJ.OA.10.2018.0206

Ito, T., Suzaki, K., Nakano, M., and Sato, A. (2013). Characterization of a new apscaviroid from American persimmon. *Arch. Virol.* 158, 2629–2631. doi: 10.1007/s00705-013-1772-x

Jiang, J. H., Zhang, Z. X., Hu, B., Hu, G. B., Wang, H. Q., Faure, C., et al. (2017). Identification of a viroid-like RNA in a lychee transcriptome shotgun assembly. *Virus Res.* 240, 1–7. doi: 10.1016/j.virusres.2017.07.012

Kim, D. H., Kim, H. R., Heo, S., Kim, S. H., Kim, M. A., Shin, I. S., et al. (2010). Occurrence of apple scar skin viroid and relative quantity analysis using real-time RT-PCR. *Res. Plant Dis.* 16, 247–253. doi: 10.5423/RPD.2010.16.3.247

Kim, N. Y., Lee, H. J., Kim, N. K., Oh, J., Lee, S. H., Kim, H., et al. (2019). Occurrence pattern of viral infection on pear in Korea and genetic characterization of apple scar skin viroid isolates. *Hortic. Sci. Technol.* 37, 767–778. doi: 10.7235/HORT.20190076

Koganezawa, H. (1985). Transmission to apple seedlings of low molecular weight RNA from apple scar skin diseased trees. *Ann. Phytopathol. Soc. Jpn.* 51, 176–182. doi: 10.3186/jjphytopath.51.176

Lee, S. H., Kwon, Y. S., Shin, H. M., Nam, S. Y., Hong, E. Y., Kim, B. K., et al. (2017). Survey on virus infection for commercial nursery trees of major apple cultivars in korea. *Res. Plant Dis.* 23, 355–362. doi: 10.5423/RPD.2017.23.4.355

Li, Z. T., Cao, Y. H., Li, N., Meng, X. L., Hu, T. L., Wang, S. T., et al. (2021). Molecular variation and phylogenetic relationship of apple scar skin viroid in seven cultivars of apple. *Sci. Agric. Sin.* 54, 4326–4336. doi: 10.3864/j.issn.0578-1752.2021.20.007

Li, S., Li, G. X., Khoo, Y. W., Yang, H., Zhang, G. R., Dai, Q. D., et al. (2020). Occurrence and molecular characterization of apple scar skin viroid in malus pumila 'Saiwaihong' (small apple tree) from north China. *J. Plant Pathol.* 102, 899–902. doi: 10.1007/s42161-020-00527-0

Liang, B. W., Li, C. Y., Ma, C. Q., Wei, Z. W., Wang, Q., Huang, D., et al. (2017). Dopamine alleviates nutrient deficiency-induced stress in malus hupehensis. *Plant Physiol. Biochem.* 119, 346–359. doi: 10.1016/j.plaphy.2017.09.012

Lichtenhaler, K., and Wellburn, A. R. (1983). Determinations of total carotenoids and chlorophylls a and b of leaf extracts in different solvents. *Biochem. Soc. Trans.* 11, 591–592. doi: 10.1042/bst0110591

Lu, J. H., Chen, B. X., Chen, T. T., Guo, S. Y., Xue, X. L., Chen, Q., et al. (2017). Comprehensive metabolomics identified lipid peroxidation as a prominent feature in human plasma of patients with coronary heart diseases. *Redox Biol.* 12, 899–907. doi: 10.1016/j.redox.2017.04.032

Moussaieff, A., Rogachev, I., Brodsky, L., Malitsky, S., Toal, T. W., Belcher, H., et al. (2013). High-resolution metabolic mapping of cell types in plant roots. *Proc. Natl. Acad. Sci. U.S.A.* 110, 1232–1241. doi: 10.1073/pnas.1302019110

Pan, Y., Meng, X. T., Che, F. B., Xue, S. L., Zhang, T., Zhao, S. R., et al. (2016). Metabolic profiles of sugar metabolism and respiratory metabolism of korla pear (Pyrus sinkiangensis yu) throughout fruit development and ripening. *Sci. Agric. Sin.* 49, 3391–3412. doi: 10.3864/j.issn.0578-1752.2016.17.013

Pidatala, V. R., Li, K., Sarkar, D., Ramakrishna, W., and Datta, R. (2016). Identification of biochemical pathways associated with lead tolerance and detoxification in chrysopogon zizanioides l. Nash (Vetiver) by metabolic profiling. *Environ. Sci. Technol.* 50, 2530–2537. doi: 10.1021/acs.est.5b04725

Takeda, R., and Ding, B. (2009). Viroid intercellular trafficking: RNA motifs, cellular factors and broad impacts. *Viruses* 1, 210–221. doi: 10.3390/v1020210

Verhoeven, J. T. J., Huner, L., Marn, M. V., Plesko, I. M., and Roenhorst, J. W. (2010). Mechanical transmission of potato spindle tuber viroid between plants of brugmansia suaveoles, solanum jasminoides and potatoes and tomatoes. *Eur. J. Plant Pathol.* 128, 417–421. doi: 10.1007/s10658-010-9675-0

Walia, Y., Dhir, S., Ram, R., Zaidi, A. A., and Hallan, V. (2014). Identification of the herbaceous host range of apple scar skin viroid and analysis of its progeny variants. *Plant Pathol.* 63, 684–690. doi: 10.1111/ppa.12118

Wang, Z., Li, G., Sun, H., Ma, L., Guo, Y., Zhao, Z., et al. (2018). Effects of drought stress on photosynthesis and photosynthetic electron transport chain in young apple tree leaves. *Biol. Open* 7, 35279. doi: 10.1242/bio.035279

Wang, J. X., Liu, Z., Xie, X. H., Wu, B., and Tong, Z. G. (2000). The reaction of virus-free apple tree growth and fruit production. *Acta Hortic. Sin.* 27, 157–160. doi: 10.3321/j.issn:0513-353X.2000.03.001

Williams, L., and Salt, D. E. (2009). The plant ionome coming into focus. *Curr. Opin. Plant Biol.* 12, 247–249. doi: 10.1016/j.pbi.2009.05.009

Wu, R., Li, J. Y., Shao, J. Z., Sun, J. S., and Zhang, H. (2015). Establishment of detection method for *Apple scar skin viroid* (ASSVd) by real-time PCR. *J. Fruit Sci.* 32, 150–155. doi: 10.13925/j.cnki.gsxb.20140288

Xi, N. N., Li, Z. T., Zhang, J. Y., Meng, X. L., Wang, Y. N., and Cao, K. Q. (2020). Detection of apple scar skin viroid by real-time fluorescence quantitative reverse transcription PCR and its movement in an apple tree. *J. Plant Prot.* 47, 1304–1312. doi: 10.13802/j.cnki.zwbhxb.2020.2020033

Zha, F. R., Wang, Z. H., Xu, Q. Z., Chen, J., Yin, Y. M., Shen, M. Q., et al. (2022). Molecular characteristics analysis of apple scar skin viroid isolates associated with dapple and scar symptoms of apple fruits. *Acta Phytopathol. Sin.*, 1-13. doi: 10.13926/j.cnki.apps.000655

Zhou, Y. B., Zhu, J. F., and Qin, S. J. (2013). Metabolomics of poplar resistance based on GC-MS technique. *J. Northeast Forestry Univ.* 41, 62–65.

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 23 March 2023
DOI 10.3389/fpls.2023.1153490

OPEN ACCESS

EDITED BY
Lin Xi,
University of Hohenheim, Germany

REVIEWED BY
Xu Xiaozhao,
China Agricultural University, China
Guixia Xu,
Institute of Botany (CAS), China

*CORRESPONDENCE
Silan Dai
silandai@sina.com

SPECIALTY SECTION
This article was submitted to Functional and Applied Plant Genomics, a section of the journal Frontiers in Plant Science

RECEIVED 29 January 2023
ACCEPTED 14 March 2023
PUBLISHED 23 March 2023

CITATION
Li J, Zhang Q, Kong D, Pu Y, Wen X and Dai S (2023) Genome-wide identification of the MIKCc-type MADS-box gene family in *Chrysanthemum lavandulifolium* reveals their roles in the capitulum development. *Front. Plant Sci.* 14:1153490. doi: 10.3389/fpls.2023.1153490

Genome-wide identification of the MIKCc-type MADS-box gene family in *Chrysanthemum lavandulifolium* reveals their roles in the capitulum development

Junzhuo Li, Qiuling Zhang, Deyuan Kong, Ya Pu, Xiaohui Wen and Silan Dai*

Beijing Key Laboratory of Ornamental Plants Germplasm Innovation and Molecular Breeding, National Engineering Research Center for Floriculture, Beijing Laboratory of Urban and Rural Ecological Environment, Key Laboratory of Genetics and Breeding in Forest Trees and Ornamental Plants of Education Ministry, School of Landscape Architecture, Beijing Forestry University, Beijing, China

Chrysanthemum ×morifolium is well known throughout the world for its diverse and exquisite flower types. However, due to the complicated genetic background of *C. ×morifolium*, it is difficult to understand the molecular mechanism of its flower development. And it limits the molecular breeding of improving chrysanthemum flower types. *C. ×morifolium* has the typical radial capitulum, and many researches showed that the members of the MIKCc-type MADS box gene family play a key role in the formation and development of the capitulum. However, it has been difficult to isolate the important MIKCc and investigate their roles in this process due to the lack of genomic information in chrysanthemum. Here, we identified MIKCc-type MADS box genes at whole genome-wide level in *C. lavandulifolium*, a diploid species closely related to *C. ×morifolium*, and investigated their roles in capitulum development by gene expression pattern analysis and protein interaction analysis. A total of 40 ClMIKCc were identified and were phylogenetically grouped into 12 clades. Members of all clades showed different enriched expression patterns during capitulum formation. We speculate that the E-class genes in *C. lavandulifolium* underwent subfunctionalization because they have a significantly expanded, more diverse expression patterns, and specifically tissue expression than *AtSEPs*. Meanwhile, we detected the C-class expressed in disc floret corolla, which could be the clue to explore the morphological differences between disc and ray floret corolla. In addition, the potential roles of some MIKCcs in complex inflorescence formation were explored by comparing the number and phylogenetic relationship of MIKCc subfamily members in Asteraceae with different capitulum types. Members of the FLC branch in Asteraceae were found to be possibly related to the differentiation and development of the ray floret.

KEYWORDS

chrysanthemum, *Chrysanthemum lavandulifolium*, capitulum development, MIKCc-type gene family, complex inflorescences

1 Introduction

In angiosperms, flowers are the reproductive organs. The yield and quality of crop and ornamental plants are directly affected by the development of their floral organs. In *Arabidopsis thaliana*, the floral organs have been divided into four rounds according to their location and developmental sequence, which are the sepals, petals, stamens, and pistils from the outside to the inside. Each round of floral organ primordia was under the regulation of a different set of genes to determine its identity (Coen and Meyerowitz, 1991; Rounsley et al., 1995; Causier et al., 2010). Currently, the best explanation for floral organ specification in plants is the ABC(D)E model. This model assumes that the different floral primordia are synergistically regulated by several genes: sepals (A+E), petals (A+B +E), stamens (B+C+E), pistils (C+E) and ovules (D+E). The ABC (D)E model contains the A class gene *APETALLA1* (*AP1*), *APETALLA2* (*AP2*), the B class genes *PISTILLATA* (*PI*), *APETALLA3* (*AP3*), the C class gene *AGAMOUS* (*AG*), the D class genes *SEEDSTICK* (*STK*), *SHATTERPROOF1/2* (*SHP1/2*) and the E class genes *SEPALLATA1/2/3/4* (*SEP1/2/3/4*). They all belong to the MIKCc clade of the MADS box gene family except *AP2* (Becker and Theiben, 2003; Theiβen et al., 2016). Other members of the MIKCc clade have also been associated with flowering processes such as flower buds resting (Theiben et al., 2000; Wu et al., 2017; Nishiyama et al., 2021), apical meristematic differentiation (Li et al., 2010), development of inflorescence morphology (Hussin et al., 2021) and flowering time (Lee and Lee, 2010; Wang et al., 2021; Ahmar et al., 2022).

However, the members of the ABC(D)E model differ in their expression patterns in plants with more primitive floral organs or more complex inflorescence structures. In addition, their functions and regulatory mechanisms are still not well characterized (Melzer et al., 2010; Zhang et al., 2017; Zhang L. et al., 2020). The capitulum of the Asteraceae is a complex inflorescence with a highly compressed inflorescence structure. Numerous florets are borne in a spiral pattern on the capitulum in a Fibonacci series with the involucre wrapped around the outside. They can be classified as homogeneous capitulum (discoid or ligulate type) and heterogeneous capitulum (radiate type) depending on the type of florets borne on the receptacle. The capitulum is the most peculiar and important organ of Asteraceae plants, gives them a unique ornamental value. Members of the MIKCc subfamily have been shown to be essential for capitulum development in numerous studies (Elomaa et al., 2018).

Although the capitulum of Asteraceae is known as pseudanthium, the process of developing floral organs and the function of ABC(D)E-class genes in the capitulum differs from that of the single flower (Elomaa et al., 2018; Zhang and Elomaa, 2021). The involucre replaces the first floral organ (sepals) to protect the inner floral organs. This results in the disappearance of the sepals or their transition to crown hairs in the capitulum to aid seed dispersal. The A-class gene *AP1* regulates the initiation and development of crown hairs in *Taraxacum mongolicum*. However, it is also widely expressed throughout the capitulum (Bremer, 1994; Harris, 1995; Wen et al., 2019a; Vijverberg et al., 2021). The structure and morphology of the second floral organ (petal) differ between the two types of florets in the radiate capitulum: Peripheral ray florets mimic the petals of individual flowers to attract pollinators. These petals are ligulate with complete upper and lower epidermis and parenchymatous tissue. The disc florets are responsible for receiving pollen and reproducing, with their petals enclosed in a corolla tube without parenchymatous tissue. B-class genes that determine petal identity are differentially expressed in the corollas of two floret types in *C. lavandulifolium* and *C. ×morifolium*, with *AP3*/*PI* both highly expressed in the ray floret corollas (Wen et al., 2019a). *GDEF2* and *GDEF3*, the orthologous genes of *AP3* in *Gerbera hybrida*, have a broader domain of expression during the differentiation of the floral organs (Broholm et al., 2010). The third floral organ (stamen) also differs in the two types of florets in the radial capitulum. Disc florets are bisexual and ray florets have only one pistil. The C class gene *AG*, which determines stamen identity, is conserved in two copies in *G. hybrida*, *C. lavandulifolium*, and *C. ×morifolium* with radial capitulum. However, the expression patterns in these three species are different (Yu et al., 1999; Wen et al., 2019a). All florets on the capitulum have a normally developed fourth floral organ (pistil) and ovule. In *Tagetes erecta*, *TeAGL11-1* shows a partial function of the D-class genes that regulate seed and petal development (Zhang C. et al., 2020). The D-class gene showed more novel function except control the seed development in Asteraceae species. A total of seven *SEP*-like genes were identified in *G. hybrida*, indicating a significant expansion of the SEP evolutionary branch. SEPs are highly functionally redundant in *A. thaliana*, but in *G. hybrida* they showed specific transcriptional patterns during capitulum formation and floral organ identity determination, suggesting subfunctionalization and neofunctionalization (Zhang et al., 2017). Thus, the relatively conserved expression patterns and functions of ABC(D)E class genes in capitulum and simple plants have resulted in the pseudanthium formation. Meanwhile, their general expanded, differentiated and subfunctionalized in Asteraceae have led to significant differences in the floral organs of capitulum and simple plants. In the meantime, ABC(D)E class genes often regulate floral organ differentiation and development as protein tetramers. Their protein interactions in Asteraceae need to be further investigated.

MIKCc members showed significant neofunctionalization during capitulum and floret development in Asteraceae compared to model plants. *ClCAL* and *ClFUL* are specifically highly expressed in *C. lavandulifolium* at the stage of floret primordia initiation. They may be involved in the differentiation of ray and disc florets (Wen et al., 2022). The B class gene *CmPI*/*CmAP3* is differentially expressed in the ray and disc florets of *C. ×morifolium*. It leads to the localization of carotenoid accumulation in the petals (Lu et al., 2022). The C-class gene *ScAG* and the D-class gene *ScAGL11* can inhibit anthocyanin synthesis in ray florets of *Senecio cruentus*, which in turn affects the formation of flower spots (Qi et al., 2022). *AGL6* plays the role of the A class gene in basal angiosperms (Kim et al., 2005). However, *GhGRCD3*, the *AGL6* ortholog in *G. hybrida*, functions as both an A- and E-class gene, affecting crown hair and corolla identity and development, as well as inflorescence meristem maintenance and floral meristem differentiation (Ruokolainen et al., 2010a; Zhang et al., 2017). The key lowering integrator *SUPPRESSOR OF OVEREXPRESSION OF CO1* (*SOC1*) affects ray floret development in *G. hybrida*. Overexpression of *GhSOC1* causes ray floret shortening and discoloration (Ruokolainen et al., 2011). The current research

indicates that MIKCc-type MADS box genes have a relatively conserved function during flower development in angiosperms. However, due to the complex structure of the capitulum, ABC(D) E-like genes appear to be commonly expanded, neofunctionalized, and subfunctionalized in Asteraceae. Other MIKCc genes have unique functions in capitulum formation and development. Therefore, the potential new functions of these classical flower development genes and their roles in the process of capitulum formation can be discovered through the mining, identification, and functional studies of MIKCc in Asteraceae.

Recently, with the release of the Asteraceae genome information (Badouin et al., 2017; Reyes-Chin-Wo et al., 2017; Song et al., 2018; Liu et al., 2020; Nakano et al., 2021; van Lieshout et al., 2022; Wen et al., 2022), the identification of gene families based on a genome-wide level has become feasible. Here, *C. lavandulifolium* (2n=2*x*=18) was selected as the object of study. It's a perennial herb and a closely related diploid of the cultivated chrysanthemum (*C.* ×*morifolium*), often used as a typical material in chrysanthemum studies. In our previous work, we have conducted a detailed study of the capitulum development process of *C. lavandulifolium*, and the study of the MIKCc gene family of *C. lavandulifolium* will help us to further elucidate its formation mechanism (Wen et al., 2019a; Wen et al., 2019b; Wen et al., 2022). In this study, the MIKCc-type MADS box gene family members in the *C. lavandulifolium* genome been identified. We analyzed their protein physicochemical properties, conserved motifs, and phylogenetic relationships. Furthermore, based on transcriptome data, we screened for highly expressed ClMIKCcs during *C. lavandulifolium* capitulum and floral organ development. And qRT-PCR, yeast two-hybrid (Y2H) and luciferase complementation assay (LCA) were used to analyze the expression patterns and protein interactions of these ClMIKCcs in nine floral organs of *C. lavandulifolium*. These results will help to gain insight into the specific functions of the MIKCc-type MADS-box gene in complex inflorescences, and the pattern of the ClMIKCc protein complex. They will also provide potential genetic resources for floral shape improvement in *C.* ×*morifolium*. Based on this, we further compared the contraction/expansion of MIKCc subfamily members in Asteraceae, discussed the differentiation characteristics of these members in different capitulum type Asteraceae plants, and provided some new perspectives for resolving MIKCc evolution correlated with capitulum types in Asteraceae.

2 Materials and methods

2.1 Collection of publicly datasets and identification of MIKCc-type MADS-box genes

A total of 16 species, including *A. thaliana* as a reference species, and 15 Asteraceae with three different capitulum types were examined in our study to determine whether the number and classification of MIKCc genes correlated with capitulum type. Among these Asteraceae, radiate type capitulum with both ray florets and disc florets such as *Artemisia annua*, *C. lavandulifolium*, *C. makinoi*, *C. nankingense*, *C. seticuspe*, and *Helianthus annuus*, *Stevia rebaudiana*, and *Smallanthus sonchifolius*. Capitulum of the discoid type with disc florets only, such as *Arctium lappa*, *Cynara cardunculus*, *Carthamus tinctorius*, and *Mikania micrantha*. Capitulum of the ligulate type with only ray florets such as *Cichorium endivia*, *Taraxacum kok-saghyz*, and *Lactuca sativa*.

The reference sequences of *A. thaliana* MIKCc-type transcription factors were available from the Plant Transcription Factor Database (http://planttfdb.gao-lab; Tian et al., 2020). The genome data of *A. annua* (PRJNA280557), *Arctium lappa* (PRJNA764011), *C. cardunculus* (PRJNA453787), *C. lavandulifolium* (PRJNA681093), *C. endivia* (PRJNA797903), *C. tinctorius* (PRJNA313459), *M. micrantha* (PRJNA528368), *S. rebaudiana* (PRJNA436363), and *S. sonchifolius* (PRJNA798108) were downloaded from the NCBI (https://www.ncbi.nlm.nih.gov/). The genome data of *H. annuu* (PRJNA345532), *L. sativa* (PRJNA173551), were downloaded from the EnsemblPlants (http://plants.ensembl.org/index.html). The genome data of *T._kok-saghyz* (PRJCA005187) was downloaded from Genome Warehouse (https://ngdc.cncb.ac.cn/gwh/). The genome data of *C. nankingense* was downloaded from Chrysanthemum Genome Database http://www.amwayabrc.com/zh-cn/index.html). *C. makinoi* genome data was downloaded at https://www.chrysanthemumgenome.wur.nl/. And *C. seticuspe* (Gojo-0 v1) genome data was downloaded from the Plant Garden database (https://plantgarden.jp/en/list/t1111766/genome/t1111766.G002).

AtMIKCc was used as a reference sequence to search against the genomic protein sequences using TBtools v1.098765 with an E-value cut-off of $1e^{-5}$ (Chen et al., 2020) to confirm the candidate MIKCc-type MADS-box genes in the above species. Each predicted sequence was then checked against NCBI and the redundant sequences were removed.

The transcriptome data of *C. lavandulifolium* (SRR14723013-SRR14723033) were downloaded from NCBI (https://www.ncbi.nlm.nih.gov/). The six stages represent the six key stages of *C. lavandulifolium* capitulum development: During the vegetative stage, there are no primordia on the conical-shaped shoot apical meristem (SAM) (stage 1, S1). During the short-day condition, SAMs began to change from conical shape to hemispherical shape (stage 2, S2). At the early stage of floret primordia, only the first-round disc floret primordia differentiated at the margin, which is a key stage for the differentiation of disc and ray florets (stage 5, S5). At the middle stage of floret primordia differentiation, ray floret primordia initiated between the involucre and the outermost disc floret primordia (stage 6, S6). At the middle stage of corolla primordia differentiation, ray floret corollas began to initiate, while disc florets differentiated to produce stamen primordia and pistil primordia (stage 9, S9). At the final stage of corolla primordia differentiation, the corolla formation of the disc florets and the ray florets was completed, and they maintained a closed state (Stage10, S10) (Wen et al., 2019b).

2.2 Chromosomal location, phylogenetic analysis, conserved structural domains and conserved motifs analysis

The chromosomal location of ClMIKCc was obtained from genome annotation files (Wen et al., 2022), and TBtools v1.098765

(Chen et al., 2020) was used to map their physical location on *C. lavandulifolium* chromosomes and display the co-linearity.

The maximum likelihood (ML) method and 5000 bootstrap replicates were used to construct the phylogenetic tree of MIKCc. TBtools v1.098765 (Chen et al., 2020) was used. In this part, sequence alignments were performed with MUSCLE (Edgar, 2004), and trimAI (–automated1 option, Capella-Gutiérrez et al., 2009) helped us trim poorly aligned sequences and reserve reliable comparison results. The best-fit model (JTT+R7) chosen according to BIC and used to construct the phylogenetic tree in IQtree v2 (Minh et al., 2020). The results were further graphically edited using iTOL (https://itol.embl.de/).

The Conserved Domains Search database (https://www.ncbi.nlm.nih.gov/cdd/?term=) was used to predict the conserved domains of MIKCc. The predicted conserved domains were visualized using TBtools v1.098765 (Chen et al., 2020).

The MEME (https://meme-suite.org/meme/tools/meme) was used to identify the conserved motifs in AtMIKCc and ClMIKCc with the following optimal parameters: the width of the motifs was from 6 to 50 amino acids and the maximum number of motifs was 10. TBtools v1.098765 (Chen et al., 2020) was used to visualize the identified motifs.

2.3 Physicochemical properties analysis and subcellular localization prediction

The molecular weight and theoretical pI of AtMIKCc and ClMIKCc were analyzed using the Expasy ProtParam online software (https://web.expasy.org/protparam/), and the subcellular localization was predicted using the WOLF PSORT online tool (https://wolfpsort.hgc.jp/).

2.4 RNA extraction and gene expression measurement

The genome-sequenced *C. lavandulifolium* G1 line was cultured at Beijing Forestry University for gene expression analysis (Wen et al., 2022). These materials were planted in 10 × 9 cm plastic pots with peat: vermiculite = 1:1. The temperature was 22 ± 1 °C and the light intensity was 4000 lx. The nutritive growth was carried out under long daylight (16 h light/8 h dark). After they grew 14 leaves, they were transferred to short daylight (12 h light/12 h dark) for reproductive growth (Fu et al., 2014). All floral organs (involucre, receptacles, corolla, pistils, ovules, stamens) were sampled at flowering, and four developmental stages of reproductive buds were collected based on Wen's literature (Wen et al., 2019b). The morphology of the sampled material is shown in Figure 1. All the material was immediately frozen in liquid nitrogen and stored at -80°C for the extraction of total RNA.

Total RNA was extracted from the collected plant materials using Plant RNA Rapid Extraction Kit (HUAYUEYANG Biotechnology, Beijing, China) and treated with RNase-free DNaseI to digest DNA. Gene expression analysis was performed by Real-time quantitative reverse transcription-PCR (qRT-PCR), which was performed using a Mini Opticon Real-time PCR System (Bio-Rad Laboratories Inc., Hercules, CA, USA) based on the SYBR Premix Ex Taq (Takara Bio Inc., Shiga, Japan). Three biological replicates were used to confirm the reliability of the results. *ClSAND* was used as an internal control gene for qRT-PCR (Qi et al., 2016). The primers for qRT-PCR were shown in Supplementary Table 2. The ΔΔCt method was used to analyze qRT-PCR data.

2.5 Protein-protein interaction assay

Pairwise yeast two-hybrid analyses were performed to verify the interactions between ClMIKCc. Gene-specific primers (Supplementary Table 3) were used to amplify protein-coding regions for all ClMIKCc tested. The PCR products were then fused into pGADT7 and pGBKT7 vectors. Sequencing was performed to verify the fusion. The constructs were then transformed into the Saccharomyces cerevisiae strain Y2H Gold and sequenced. Only the ClPI and ClSEPd constructs were found to have the ability to self-activate the expression of the HIS3 reporter genes. 10 mmol/L 3-amino-1, 2, 4-triazole (3-AT) was used to eliminate self-activation. For further interaction studies, recombinant vectors were used. Three technological replicates and three different concentrations were used in the protein-protein interaction experiments.

2.6 Luciferase complementation assay

The luciferase complementation assay (LCA) was used to verify the positive results obtained in the Y2H assay (Chen et al., 2008). Target genes were ligated separately to pCAMBIA1300-cLUC and pCAMBIA1300-nLUC using the homologous recombination method. All primer sequence used to construct the vector is shown in Supplementary Table 4. Stop codon removal was required for all genes ligated to pCAMBIA1300-nLUC. The resulting recombinant plasmid was transformed into *Agrobacterium tumefaciens* (GV3101) using the freeze-thaw method.

Empty pCAMBIA1300-cLUC and pCAMBIA1300-nLUC carriers were used as controls. Subsequently, the Agrobacterium suspension (OD600 = 1.0) was injected into the abaxial leaf surface of three-week-old *Nicotiana benthamiana* leaves. After 72 h (24 h dark and 48 h light)of incubation in the greenhouse, D-Luciferin potassium salt was applied to the abaxial side of these leaves and then placed in the dark for 5 min. The fluorescence signal was observed using LB983 NightOwl II (Berthold Technologies, Bad Wildbad, Germany).

3 Results

3.1 Identification and general information of ClMIKCc

The 42 possible MIKCc proteins were retrieved from the *C. lavandulifolium* genome using two different BLAST methods. The

10.3389/fpls.2023.1153490

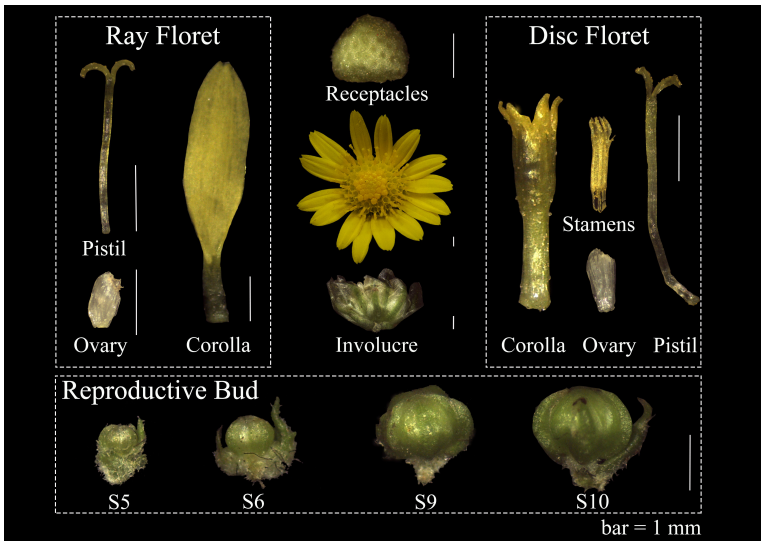


FIGURE 1
Floral organs on capitulum and four development stages reproductive bud in *C. lavandulifolium* for qRT-PCR. bar=1mm.

retrieved MIKCc proteins were further screened based on the results of multiple sequence comparisons and the Conserved Domain of the NCBI website. The final 40 ClMIKCc proteins were identified. ClMIKCc encodes a polypeptide of 183 to 296 amino acids ranging in molecular weight from 20.73 to 34 kD. The theoretical pI range was 5.07 to 9.58. Subcellular localization predictions showed that 32 of the 40 ClMIKCc proteins were localized in the nucleus, which is typical of transcription factors (Supplementary Table 1).

Based on the genome annotation information of *C. lavandulifolium*, 38 of the 40 ClMIKCc were localized to all nine chromosomes. Two ClMIKCc were not associated with chromosomes (Figure 2). These ClMIKCc were not evenly distributed among the chromosomes. Chromosomes 1, 2, and 9 contained the fewest members, with two members each, and chromosome 5 contained the most, with eight members. The co-linearity analysis shows that only two MIKCc undergone a tandem repeat (Figure 2).

3.2 Phylogenetic analysis of ClMIKCc and AtMIKCc

To elucidate the phylogenetic relationships of ClMIKCc, an ML phylogenetic tree was constructed using 40 ClMIKCc and 39 AtMIKCc (Figure 3A). ClMIKCc were distributed in 12 clades (GMM13, AP3/PI, AGL15, SVP, ANR1, FLC, SOC1, AGL12, AG/SHP/STK, AP1/FUL, AGL6, SEP), ClMIKCc underwent contraction in FLC and SOC1 clade and expansion in SVP, FUL and SEP clade.

Using the MEME website, the conserved motifs of ClMIKCc and AtMIKCc were predicted. 10 conserved motifs were identified and labeled as motifs 1-10 (Supplementary Figure 1). For ClMIKCc and AtMIKCc, the motifs were relatively consistent and were located in the same clade (Figure 3B). There are 4-8 conserved motifs in each ClMIKCc, and all ClMIKCc and AtMIKCc proteins have Motif 1 and Motif 2. Combined with the predicted results of conserved structural domains, Motif 1 and Motif 2 are the core

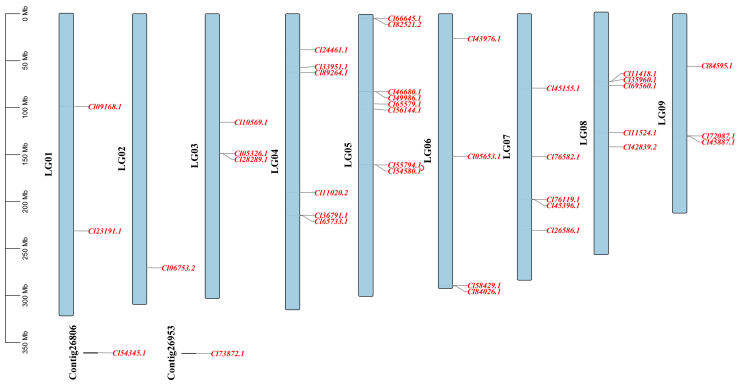


FIGURE 2
Localization and co-linearity of *ClMIKCc* on nine *C. lavandulifolium* chromosomes. The red line shown the co-linearity relationship between genes.

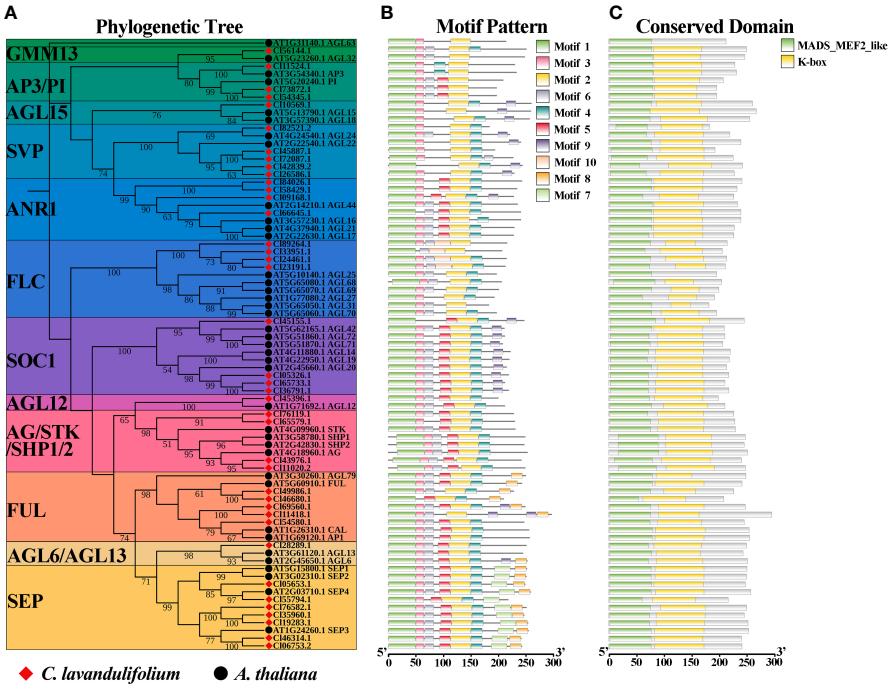


FIGURE 3
Analysis of phylogenetic, conserved motifs, and conserved structural domains of ClMIKCc and AtMIKCc. **(A)** ML phylogenetic tree of ClMIKCc and AtMIKCc. The number above branches represented bootstrap value and the values below 50 would not be shown; **(B)** Conserved motifs in ClMIKCc and AtMIKCc, 10 predicted motifs are indicated by different colored squares; **(C)** Conserved structural domains in ClMIKCc and AtMIKCc, 2 predicted structural domains are indicated by squares of different colors.

motifs of the MADS domain and K-box, respectively. Motif 7 is a specific conserved motif in At- and ClSEPs. Motif 10 is present only in ClFLCs (Figure 3C, Supplementary Figure 1).

3.3 Expression pattern of ClMIKCc during *C. lavandulifolium* capitulum development

Based on the public transcriptome dataset of *C. lavandulifolium* capitulum developmental stages (Wen et al., 2022), the expression pattern of *ClMIKCc* during *C. lavandulifolium* capitulum development was analyzed. A total of 26 *ClMIKCc* are expressed in leaves and six developmental stages (Figure 4A). These *ClMIKCc* could be mainly classified into four classes according to their expression patterns. Among them, class I, class II, and class IV *ClMIKCc* were specifically expressed in the capitulum (Supplementary Figure 2). Class I included the B-, C-class genes and *ClSEPd/e/f*, which were abundantly expressed in the late stage of capitulum development (floral organ differentiation period). These genes are mainly involved in the process of floral organ differentiation in *C. lavandulifolium* florets. Class II includes *ClAP1a*, *ClSEPa/b/c/f*, *ClSOC1a*, *ClSVPa* and *ClFLCa*, which showed high expression levels after the floral transition. They may be involved in both inflorescence primordium development and floral organ differentiation. Class III includes *ClFLCb*, *ClFULa*, and *ClSVPb/c*, which are mainly expressed in the leaf, the apical meristem before and after the floral transition, and the early stage of capitulum development. They may be involved in the floral transition process in *C. lavandulifolium*. Class IV contains *ClANR1a*, *ClSOC1b/c/d* and *ClFULb*, which are mainly expressed in the middle stage of capitulum development (S2-S6). They may regulate the floral transition and the formation of the inflorescence structure.

To further elucidate the role of *ClMIKCc* in the development of the *C. lavandulifolium* capitulum, we examined the expression of class I and class II *ClMIKCc* in the capitulum of *C. lavandulifolium* at four critical periods after the initiation of the florets primordium by qRT-PCR. As shown in Figure 4B, almost all of these *ClMIKCc* were first expressed after the initiation of the floret primordium (S5). The expression increased with increasing floret number and floral organ differentiation. Interestingly, with the flowering process of *C. lavandulifolium*, the expression of *ClSVPa* and *ClFLCa*, two genes usually considered as flowering repressors, were also up-regulated. During the development of the *C. lavandulifolium* inflorescence, they may play a different role than their paralogous genes.

3.4 ClMIKCc expression and protein interaction in *C. lavandulifolium* floral organs

In order to investigate these genes' roles in the development of different floral organs, 21 *ClMIKCc* was determined by qRT-PCR in nine floral organs. Class I and II of *ClMIKCc* were accompanied by the floral organ development, highly expression during the S9 and

10.3389/fpls.2023.1153490

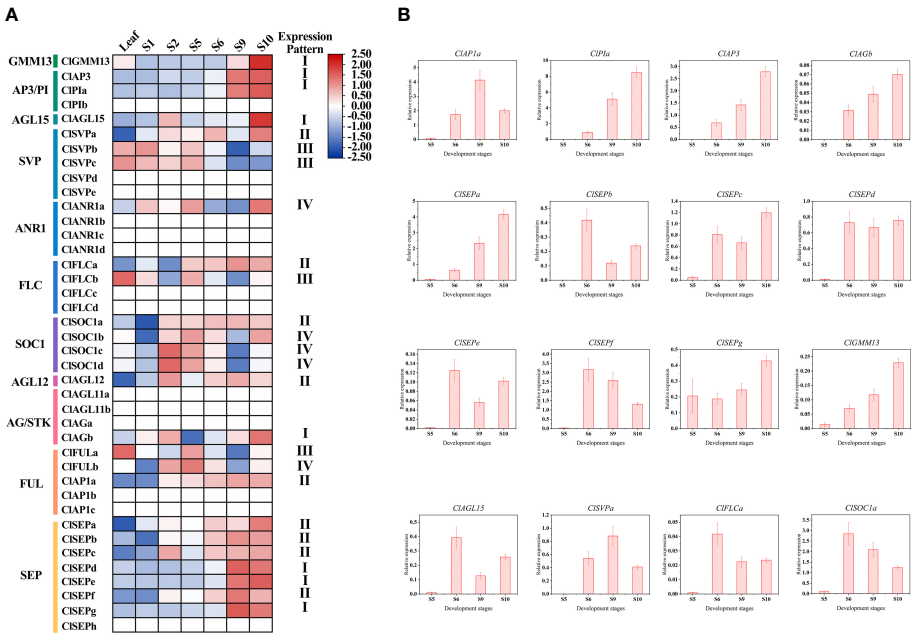


FIGURE 4
Expression pattern of *ClMIKCc* during the capitulum development stage. **(A)** Expression heat map constructed based on transcriptome data during the capitulum development stage of *C. lavandulifolium*. **(B)** qRT-PCR analysis of the expression level of class I and class II ClMIKCc in four development stages. Error bars represent the standard errors of three biological replicates.

S10. Those ClMIKCc may be involved in the process of floral organ identity determination and development. The results (Figure 5) showed that the A-class gene *ClAP1a* was widely expressed in the nine floral organs. There are three B-class genes in *C. lavandulifolium*, *ClAP3*, *ClPIa*, and *ClPIb*, among which *ClAP3* and *ClPIa* had a typical B-class gene expression pattern that expressed mainly in the stamens and corollas, also a weak expression was detected in the ovules. However, *ClPIb* was detected weakly expressed in all floral organs. A total of two C-class genes were identified in the *C. lavandulifolium* genome, *ClAGa* was not only expressed in the pistil, ovule, and stamen but also partially overlapping expressed with *ClAP1a* in the corolla of disc floret. *ClAGb* showed a low expression level and a broad expression domain which is like *ClPIb*. This expression patterns difference of paralogous genes was also present in two *ClAGL11*-like genes, *ClAGL11a* specifically expressed in the ovules, while *ClAGL11b* expressed in all floral organs but with a lower level. The E-class genes in the SEP branch showed diverse expression patterns. *ClSEPa* was mainly expressed in ovules and stamens. *ClSEPc* expressed in all floral organs except corollas and pistils of disc floret. *ClSEPd* and *ClSEPe* mainly expressed in ovules and ray floret corollas. *ClSEPb*, *ClSEPf*, and *ClSEPg* expressed in all floral organs.The result also showed that *ClAGL12*, *ClAGL15*, and *ClFLCa* weakly expressed in all floral organs. *ClGMM13* and *ClSOC1a* expressed in pistils, carpels, stamens, involucre, and receptacles. C*lSVPa* specifically expressed in involucre and receptacles (Figure 5).

The expression pattern analysis of 26 *ClMIKCc* in nine floral organs showed that they had similar patterns in the corresponding floral organs between the two floret types (Figure 5). The expression patterns of *ClMIKCc* in the floral organs of *C. lavandulifolium* were generally consistent with the ABC(D)E model in model plants (Figure 5). The B-and E-class genes were predominant in the corolla of both floret types. The stamen development was regulated by B-, C-, and E-class genes. And the pistils and ovules development were regulated by B-, C-, and E-class genes as well as *AGL11*-like. However, differences in the expression level of same genes or the tissue-specific expression patterns of homologous genes in two floret types may be the cause of their morphology distinction.

Since members of the MIKCc gene subfamily usually form homologous or heterologous protein complexes to function, we performed the yeast two-hybrid assay and LCA to analyze the protein interaction between the four most highly expressed ClMIKCc in each floral organ. The positive results obtained for all Y2H systems are shown in the Supplementary Figure 3. Since the Y2H system may give false positive results, we verified the positive results of the heterologous interactions by the LCA in tobacco (Supplementary Figure 4). These results showed that these four ClMIKCc were able to form protein complexes through different modes of interaction (Supplementary Figure 3-4). Although the A-class gene *ClAP1a* was widely expressed in the capitulum, it could only interact with ClAGa, ClAGb, and ClAGL11b from the AG clades, as well as the homolog of Bsister (Bs), ClGMM13 (Supplementary Figure 3). Among the B class genes, ClAP3 can only interact with ClPIa. However, ClPIa can extensively form protein dimers with C and E class genes. Similar to B class genes, among two copies of C class gene ClAG, ClAGa was able to form protein dimers with more ClMIKCc. ClAGb was only able to interact with A and C class genes. ClAGL11-like is highly

10.3389/fpls.2023.1153490

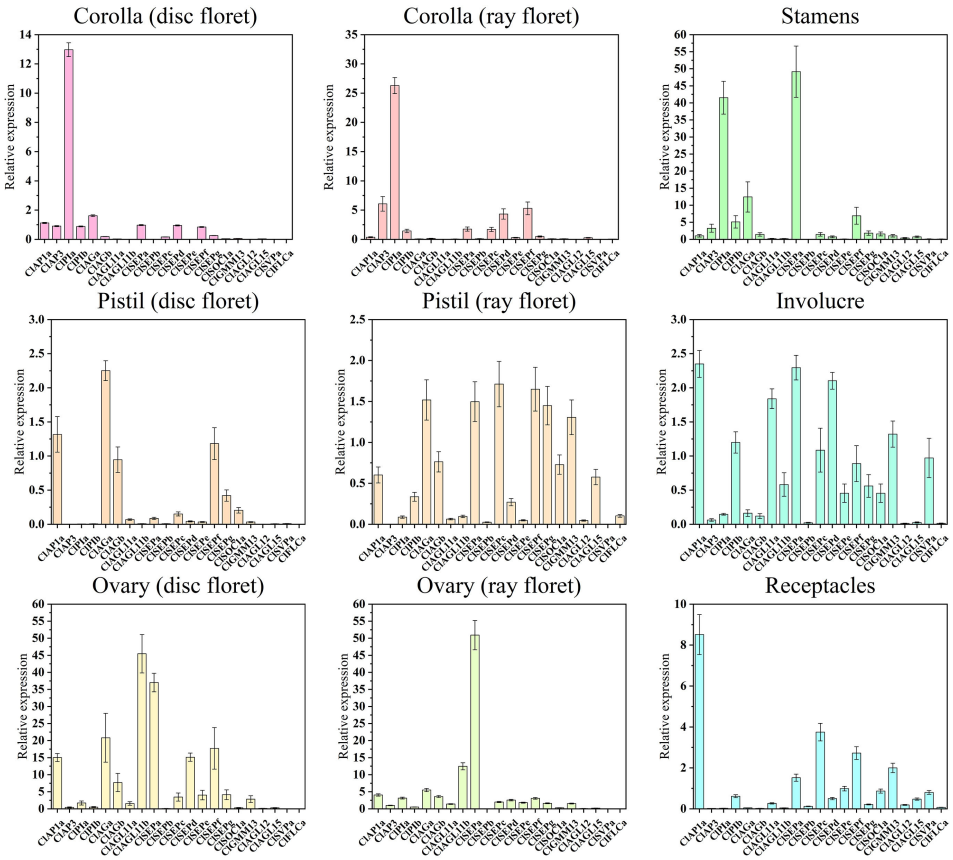


FIGURE 5
Quantitative RT-PCR analysis of the expression level of 21 ClMIKCc in nine floral organs. Error bars represent the standard errors of three biological replicates.

expressed in ovules and can interact with E-class genes. Therefore, it may function as a D-class gene in *C. lavandulifolium*. *ClSEPs* are highly expressed in all floral organs, and they also interact to a large extent with members of other clades. Thus, they appear to be key members of the multi-protein complex. ClGMM13 from the Bs evolutionary branch can interact with A- and E-class genes. Its function may also depend on the formation of the protein complex (Figure 6).

3.5 Evolutionary analysis of MIKCc in Asteraceae with different capitulum types and single flower plants

For Asteraceae MIKCc phylogenetic analysis, ML trees were constructed using MIKCc proteins from 16 species. These MIKCcs were grouped into 12 clades, and the number of proteins varied greatly among these branches. The SVP, SOC1, SEP, FLC, and AP1/FUL clades had more members (Figure 7A). Furthermore, the number of MIKCc per species in each branch was counted (Figure 7B). The results showed that the number of MIKCc was lowest in *C. cardunculus*, which has a discoid type capitulum, with 23 members. While the highest number of proteins was found in *L. sativa*, which had a ligulate type capitulum, with 85 members. SEP clade significant expansion in radial capitulum species, the number generally more than discoid and ligulate type capitulum species. This may be related to the more complex inflorescence structure of the radial capitulum. FLC clade is abundantly retained in Asteraceae with ray floret, whether they were need vernalization.

4 Discussion

4.1 Overview of MIKCc-type MADS-box gene family in *C. lavandulifolium*

A total of 40 MIKCc-type MADS box genes were identified from the whole genome sequence of *C. lavandulifolium*. Chromosome localization results showed that ClMIKCcs were unevenly distributed on nine chromosomes of *C. lavandulifolium* (Figure 2). Based on a phylogenetic tree constructed from ClMIKCc and AtMIKCc, the ClMIKCcs were divided into 11 clades (Figure 3A). The number of A-class genes *AP1*, B-class genes *PI* *AP3*, C-class genes *AG*, and flower-forming integrator *SOC1* genes in *C. lavandulifolium* genome was consistent with previous results based on the *C. lavandulifolium* transcriptome and homologous clones (Wang et al., 2014; Wen et al., 2019a). In addition, we identified many E-class genes in the *C. lavandulifolium* genome.

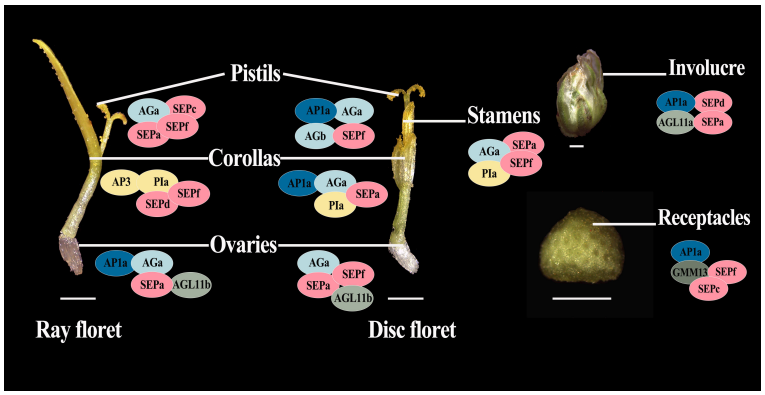


FIGURE 6
Suppositional interaction model of the four most highly expressed ClMIKCc in the nine floral organs of *C. lavandulifolium*. Different colors are used to distinguish the members of the different clades. Intersecting ellipses indicate that the protein pair has interactions. All positive results are supported by the Y2H and LCA results. bar=1 mm.

These SEP clade members showed diverse expression patterns and extensive protein interactions during capitulum and floral organ development (Figures 4–6), which may play a critical role in regulating complex inflorescence formation.

Although *C. lavandulifolium* does not require vernalization to complete floral transition (Fu et al., 2014), the flowering repressor *FLOWERING LOCUS C* (*FLC*), the key integrator of the vernalization pathway, retains four members in the *C. lavandulifolium* genome. In contrast, other nonvernalizing plants such as *Oryza sativa*, *Cucumis sativus*, *Zea mays*, *Sorghum bicolor*, and *Gossypium hirsutum* have experienced a complete loss of members of the FLC members (Arora et al., 2007; Zhao et al., 2011; Hu and Liu, 2012; Ren et al., 2017). Notably, all ClFLCs have a unique motif 10 (Figure 3B, Supplementary Figure 1). This motif is not found in other ClMIKCcs and all AtMIKCcs. These results suggest that the ClFLCs may have a specific function in the flowering process of *C. lavandulifolium*.

4.2 Role of ClMIKCc in the *C. lavandulifolium* capitulum development

In *A. thaliana*, members of the MIKCc subfamily not only regulate floral organ identity and development, but also play important roles in several biological processes, such as flowering time, nitrogen utilization, and seed development (Gan et al., 2005; Mizzotti et al., 2014; Theiβen et al., 2016). ClMIKCc has a typical MADS box domain and K box, with similar conserved motifs to the *A. thaliana* reference sequence (Figure 3), and its function may also be conserved.

However, unlike *A. thaliana*, the process of capitulum development in *C. lavandulifolium* involves five stages, which are floral induction, floral transition, involucre differentiation, floral primordia differentiation, and corolla primordia differentiation (Wen et al., 2019b). We found that *ClAP1*, *ClAP3*, *ClPIa/b*, *ClAGb*, *ClSEPs*, *ClSOC1a*, *SHORT VEGETATIVE PHASE* (*ClSVPa*), and *ClFLCa* maintained a high level of expression after the floral primordia (Figure 4B). In *A. thaliana*, *AP1* is an important floral meristem maintenance gene to maintain the limited growth of floral meristem tissues (Pelaz et al., 2001; Litt, 2007; Han et al., 2014), in addition to functioning as an A-class gene to determine the identity of sepals and petals. In *G. hybrid*, the A- and E-class genes work together to determine the fate of the inflorescence meristem tissue (Uimari et al., 2004; Ruokolainen et al., 2010b). Usually, *SOC1* considered an activator in the flowering pathway, and *SVP* and *FLC* were repressors (Blümel et al., 2015). In *A. thaliana*, *AGL24*, *SVP* and *AP1* function together to promote and maintain the floral meristem (Gregis et al., 2008). *SOC1*. In contrast, *FLC* were not found to be associated with inflorescence meristem or floral meristem formation. In view of the fact that there are specific motifs in *ClFLCs* (Figure 2B) and *ClFLCa* expression of which is upregulated with *C. lavandulifolium* flowering (Figure 4B), we hypothesize that part of *FLCs* are involved in the formation of the inflorescence meristem and floral meristem.

In this study, we found some important genes regulating the developmental process of *C. lavandulifolium* capitulum, including classical A-, B-, C- and E-class genes. Novel expression patterns were also found for some key genes in the photoperiodic pathway. However, their function needs to be further investigated.

4.3 Expression pattern and interaction model of ClMIKCc in *C. lavandulifolium* floral organs

The involucre is formed by a cluster of bracts. It is an important basis for the classification of Asteraceae (Carlquist, 1976), based on its shape and the number of whorls. The involucre is located on the outermost whorl of the capitulum. It envelops and protects the capitulum (Yang and Sun, 2009). It is similar to the first whorl of the floral organs (sepals) in individual flowers in morphology, position, and function. We observed a high expression level of A- and E-class genes expressed in the involucre of *C. lavandulifolium* (Figure 5). This expression pattern is also similar to that of sepals (Theißen et al., 2016).

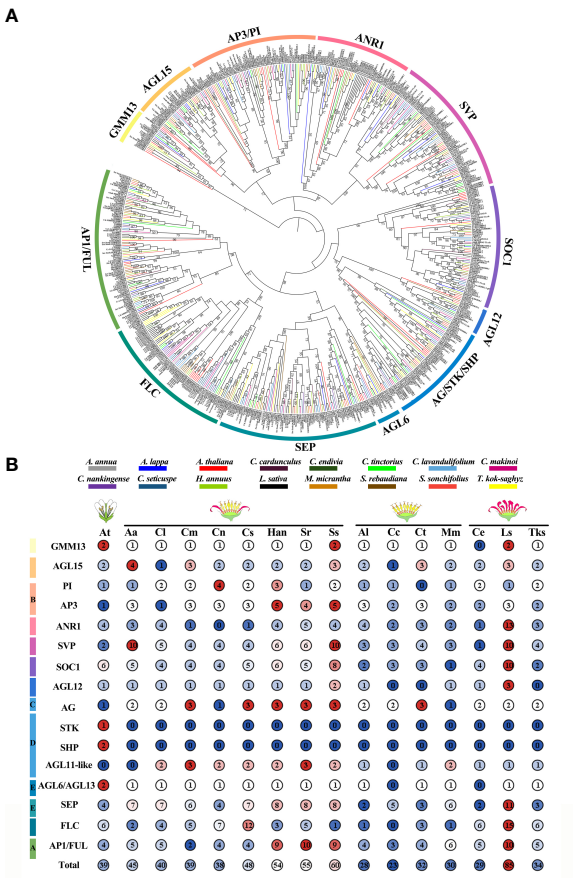


FIGURE 7
Phylogenetic analysis and statistic of MIKCc in Arabidopsis and fifteen Asteraceae. **(A)** The ML Phylogenetic tree of MIKCc in 15 Asteraceae. The bootstrap replicates were 5000. The number above branches represented bootstrap value and the values below 50 would not be shown. We used different colors to distinguish species and clades. **(B)** The member counts of each clade of MIKCc in 16 species. Their respective inflorescence types are indicated on the top. At, *A. thaliana*; Aa, *A. annua*; Cl, *C. lavandulifolium*; Cm, *C. makinoi*; Cn, *C. nankingense*; Cs, *C. seticuspe*; Han, *H. annuus*; Sr, *S. rebaudiana*; Ss, *S. sonchifolius*; Al, *A. lappa*; Cc, *C. cardunculus*; Ct, *C. tinctorius*; Mm, *M. micrantha*; Ce, *C. endivia*; Ls, *L. sativa*; Tks, *T._kok-saghyz*.

Interestingly, we also found that the D-class gene *ClAGL11a* is highly expressed in the involucre (Figure 5). It can also interact with ClAP1a and ClSEPa to form a protein complex (Figures 5, 6). In contrast, ectopic expression of STK/AGL11 in *A. thaliana* leads to the transformation of sepals into carpel-like organs with ovules, indicating the antagonism between D-class genes and sepals (Favaro et al., 2003; Pinyopich et al., 2003). This result suggests that the involucre of *C. lavandulifolium* is similar to the first whorl of floral organs (sepals) of *A. thaliana*, but not in terms of the pattern of gene expression. The A-class gene *ClAP1a* is widely expressed in all floral organs. This result is similar to previous findings in *C. lavandulifolium*, *C.* ×*morifolium*, and *G. hybrida* (Ruokolainen et al., 2010b; Wen et al., 2019a).

There is a clear difference in corolla morphology between ray and disc florets (Figure 1). B class genes have a higher expression level in the ray floret corolla. Interestingly, we observed an overlap in the expression of A- and C-class genes in the disc floret corolla of *C. lavandulifolium* (Figure 5). C-class genes are usually considered to have a negative effect on corolla formation (Wollmann et al., 2010). In the classical ABC model, the A- and C-class genes are antagonistic to each other, with the A-class genes repressing the expression of the C-class genes in order to maintain petal/stamen differentiation (Coen and Meyerowitz, 1991). However, in the disc floret corolla, *ClAP1a* and *ClAGa* not only have similar expression patterns. They can also form a heterodimer. We suggest that *ClAGa* may be a candidate to explain such differences, considering the obvious differences in length and structure between ray and disc corollas. Meanwhile, E-class genes are also differentially expressed between the two floral types (Figure 5). Studies on *G. hybrid* have also shown that E-class genes can directly affect floral organ identity and morphology (Zhang et al., 2017).

Except for the C class genes that were found to be expressed in the corolla of disc florets, the B, C, D, and E class genes in *C. lavandulifolium* have similar tissue-specific expression to the model plant (Figure 5). We also found that the retention pattern of *C. lavandulifolium* is conserved for paralogs, with some members of all four gene classes (*ClPIb*, *ClAGb*, *ClAGL11b*, *ClSEPc*, *ClSEPg*) showing a consistent expression pattern with lower expression levels and broader expression domains (Figure 5). In Asteraceae (Figure 7), expansions and duplications of ABC(D)E class genes are widespread. Expansions of A-, B-, C-, and E-class genes were also found in *G. hybrid*. Some members have low expression levels and broad expression domains. Inhibition of these members results in little phenotype alteration (Broholm et al., 2010; Ruokolainen et al., 2010a; Ruokolainen et al., 2010b; Zhang et al., 2017). These results suggest that, in response to the expansion and duplication of ABC(D)E class genes, this strategy of subfunctionalization is conserved in Asteraceae.

4.4 Correlation between the number of MIKC and the capitulum types in Asteraceae

The number of MIKCc gene subfamily members showed significant differences among Asteraceae plants with distinct capitulum types (Figure 7B). The largest account of MIKCc was 85 in *L. sativa* which has the ligulate type capitulum. The smallest number of MIKCc was in the discoid type capitulum species *C. cardunculus*, at 23. ANR1, SVP, SOC1 FLC, and SEP clades all underwent a significant expansion in *L. sativa* due to a specific whole genome triplication (WGT) event, with many MADS-box genes distributed within the triplication region (Reyes-Chin-Wo et al., 2017). In *H. annuus*, which also underwent WGT (Badouin et al., 2017), the number of MIKCcs was also significantly higher than that in other Asteraceae. These results suggest that genome duplication events are the main reason for the expansion of MIKCc gene family in Asteraceae.

In perennial Brassicaceae plants (e.g. *Arabis alpine*), *FLC* expression is repressed by low temperatures of winter and reactivated in spring, conferring a seasonal flowering pattern to the plant (Leijten et al., 2018). In this study, the amount of FLC branch members were found to be significantly different among seven Asteraceae tested. Many *FLCs* retained in Asteraceae species with ray florets, although they did not need to undergo a low-temperature

period to complete the flowering transition. Whereas in two discoid type capitulum species, *C. cardunculus* and *M. micrantha*, the *FLC* was significantly contracted or completely lost (Figure 7B). Previous studies have shown that *C. lavandulifolium* flowering is strictly dependent on short-day induction rather than low temperature (Fu et al., 2014), but four *FLC*-like (Figure 3) are retained in the genome and these genes contain a specific motif 10 (Figure 3B, Supplementary Figure 1). Transcriptomic data on the capitulum development stage showed that *ClFLCa* highly expressed at the stage of the florets and the corolla primordium differentiation (Figure 4A), which is the same expression pattern with A- and E-class genes (Figure 4). It might play a novel role in the capitulum. *CiMFL* in *Cichorium intybus*, the ortholog of *AtFLC*, was found to fail to recover the early flowering phenotype of the *flc3* mutant of *A. thaliana*, but instead resulted in abnormal leaf organ morphology. However, the morphological changes in the floral organs were not observed due to a continuous reduction in the *CiMFL* expression level during flowering in transgenic lines (Locascio et al., 2009). This result also suggests that members of the *FLC* in Asteraceae have a novel function in controlling organ morphology. Accordingly, we hypothesize that the *FLC* may have novel functions in the development of capitulum in Asteraceae, especially in the development of ray florets.

5 Conclusions

In this study, a total of 40 ClMIKCc proteins were identified and classified into 12 branches. Transcriptomic data and qRT-PCR indicated that most *ClMIKCc* are specifically expressed in the capitulum, and *ClFLCa* may have novel functions in regulating the differentiation and development of ray floret. *C. lavandulifolium* has a conservative coping strategy for the redundant MIKCc gene, where members of the same branch have similar expression patterns, but only one member is highly expressed. Y2H and LCA results revealed the interactions of ClMIKCc in different floral organs of *C. lavandulifolium*. That will help to determine the tetrameric model in *C. lavandulifolium* floral organs. The results of phylogenetic analysis of MIKCc in Asteraceae indicated that there may be a correlation between the number of SEPs and the complexity of capitulum structure, as the SEP clade was significantly expanded in Asteraceae with radial capitulum. However, FLC was almost lost in Asteraceae with discoid capitulum, which, combined with the expression pattern of *ClFLCa* in *C. lavandulifolium*, further indicates the potential role of FLC in ray floret development in Asteraceae.

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

SD conceived and designed the study. JL performed most of the experiments and data analysis. QZ performed part of gene expression analysis. JL wrote the manuscript, and DK, YP, and XW edited it. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by grants from the National Natural Science Foundation of China (No. 31530064).

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1153490/full#supplementary-material

References

Ahmar, S., Zhai, Y., Huang, H., Yu, K., Khan, M. F. U., Shahid, M., et al. (2022). Development of mutants with varying flowering times by targeted editing of multiple *SVP* gene copies in *Brassica napus* l. *Crop J.* 10, 67–74. doi: 10.1016/j.cj.2021.03.023

Arora, R., Agarwal, P., Ray, S., Singh, A. K., Singh, V. P., Tyagi, A. K., et al. (2007). MADS-box gene family in rice: genome-wide identification, organization and expression profiling during reproductive development and stress. *BMC Genom.* 8 (1), 1–21. doi: 10.1186/1471-2164-8-242

Badouin, H., Gouzy, J., Grassa, C. J., Murat, F., Staton, S. E., Cottret, L., et al. (2017). The sunflower genome provides insights into oil metabolism, flowering and asterid evolution. *Nature* 546, 148–152. doi: 10.1038/nature22380

Becker, A., and Theiβen, G. (2003). The major clades of MADS-box genes and their role in the development and evolution of flowering plants. *Mol. Phylogenet. Evol.* 29, 464–489. doi: 10.1016/S1055-7903(03)00207-0

Blümel, M., Dally, N., and Jung, C. (2015). Flowering time regulation in crops–what did we learn from arabidopsis? *Curr. Opin. Biotech.* 32, 121–129. doi: 10.1016/j.copbio.2014.11.023

Bremer, K. (1994). *Asteraceae: Cladistics and classification* (Portland: Timber Press).

Broholm, S. K., Pöllänen, E., Ruokolainen, S., Tähtiharju, S., Kotilainen, M., Albert, V. A., et al. (2010). Functional characterization of b class MADS-box transcription factors in *Gerbera hybrida*. *J. Exp. Bot.* 61, 75–85. doi: 10.1093/jxb/erp279

Capella-Gutiérrez, S., Silla-Martínez, J. M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25 (15), 1972–1973. doi: 10.1093/bioinformatics/btp348

Carlquist, S. (1976). Tribal interrelationships and phylogeny of the asteraceae. *Aliso* 8, 465–492. doi: 10.5642/aliso.19760804.10

Causier, B., Schwarz-Sommer, Z., and Davies, B. (2010). Floral organ identity: 20 years of ABCs. *Semin. Cell Dev. Biol.* 21, 73–79. doi: 10.1016/j.semcdb.2009.10.005

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, H., Zou, Y., Shang, Y., Lin, H., Wang, Y., Cai, R., et al. (2008). Firefly luciferase complementation imaging assay for protein-protein interactions in plants. *Plant Physiol.* 146, 368–376. doi: 10.1104/pp.107.111740

Coen, E. S., and Meyerowitz, E. M. (1991). The war of the whorls: genetic interactions controlling flower development. *Nature* 353, 31–37. doi: 10.1093/nar/gkh340

Elomaa, P., Zhao, Y., and Zhang, T. (2018). Flower heads in asteraceae–recruitment of conserved developmental regulators to control the flower-like inflorescence architecture. *Hortic. Res.* 5. doi: 10.1038/s41438-018-0056-8

Favaro, R., Pinyopich, A., Battaglia, R., Kooiker, M., Borghi, L., Ditta, G., et al. (2003). MADS-box protein complexes control carpel and ovule development in arabidopsis. *Plant Cell* 15, 2603–2611. doi: 10.1105/tpc.015123

Fu, J., Wang, L., Wang, Y., Yang, L., Yang, Y., and Dai, S. (2014). Photoperiodic control of *FT*-like gene *ClFT* initiates flowering in *Chrysanthemum lavandulifolium*. *Plant Physiol. Bioch.* 74, 230–238. doi: 10.1016/j.plaphy.2013.11.004

Gan, Y., Filleur, S., Rahman, A., Gotensparre, S., and Forde, B. G. (2005). Nutritional regulation of *ANR1* and other root-expressed MADS-box genes in *Arabidopsis thaliana*. *Planta* 222 (4), 730–742. doi: 10.1007/s00425-005-0020-3

Gregis, V., Sessa, A., Colombo, L., and Kater, M. M. (2008). *AGAMOUS-LIKE24* and *SHORT VEGETATIVE PHASE* determine floral meristem identity in arabidopsis. *Plant J.* 56, 891–902. doi: 10.1111/j.1365-313X.2008.03648.x

Han, Y., Zhang, C., Yang, H., and Jiao, Y. (2014). Cytokinin pathway mediates *APETALA1* function in the establishment of determinate floral meristems in arabidopsis. *Proc. Natl. Acad. Sci. U.S.A.* 111, 6840–6845. doi: 10.1073/pnas.131853211

Harris, E. M. (1995). Inflorescence and floral ontogeny in asteraceae: a synthesis of historical and current concepts. *Bot. Rev.* 61, 93–278. doi: 10.1007/BF02887192

Hu, L., and Liu, S. (2012). Genome-wide analysis of the MADS-box gene family in cucumber. *Genome* 55, 245–256. doi: 10.1139/g2012-009

Hussin, S. H., Wang, H., Tang, S., Zhi, H., Tang, C., Zhang, W., et al. (2021). *SiMADS34*, an e-class MADS-box transcription factor, regulates inflorescence architecture and grain yield in *Setaria italica*. *Plant Mol. Biol.* 105, 419–434. doi: 10.1007/s11103-020-01097-6

Kim, S., Koh, J., Yoo, M. J., Kong, H., Hu, Y., Ma, H., et al. (2005). Expression of floral MADS-box genes in basal angiosperms: Implications for the evolution of floral regulators. *Plant J.* 43, 724–744. doi: 10.1111/j.1365-313X.2005.02487.x

Lee, J., and Lee, I. (2010). Regulation and function of SOC1, a flowering pathway integrator. *J. Exp. Bot.* 61, 2247–2254. doi: 10.1093/jxb/erq098

Leijten, W., Koes, R., Roobeek, I., and Frugis, G. (2018). Translating flowering time from *Arabidopsis thaliana* to brassicaceae and asteraceae crop species. *Plants* 7, 111. doi: 10.3390/plants7040111

Li, Z. M., Zhang, J. Z., Mei, L., Deng, X. X., Hu, C. G., and Yao, J. L. (2010). *PtSVP*, an *SVP* homolog from trifoliate orange (*Poncirus trifoliata* l. raf.), shows seasonal periodicity of meristem determination and affects flower development in transgenic *Arabidopsis* and tobacco plants. *Plant Mol. Biol.* 74, 129–142. doi: 10.1007/s11103-010-9660-1

Litt, A. (2007). An evaluation of a-function: evidence from the *APETALA1* and *APETALA2* gene lineages. *Int. J. Plant Sci.* 168, 73–91. doi: 10.1086/509662

Liu, B., Yan, J., Li, W., Yin, L., Li, P., Yu, H., et al. (2020). *Mikania micrantha* genome provides insights into the molecular mechanism of rapid growth. *Nat. Commun.* 11, 1–13. doi: 10.1038/s41467-019-13926-4

Locascio, A., Lucchin, M., and Varotto, S. (2009). Characterization of a *MADS FLOWERING LOCUS c-LIKE* (*MFL*) sequence in cichorium intybus: a comparative study of *CiMFL* and *AtFLC* reveals homologies and divergences in gene function. *New Phytol.* 182, 630–643. doi: 10.1111/j.1469-8137.2009.02791.x

Lu, C., Qu, J., Deng, C., Liu, F., Zhang, F., Huang, H., et al. (2022). The transcription factor complex CmAP3-CmPI-CmUIF1 modulates carotenoid metabolism by directly regulating the carotenogenic gene *CmCCD4a-2* in chrysanthemum. *Hortic. Res.* 9. doi: 10.1093/hr/uhac020

Melzer, R., Wang, Y. Q., and Theißen, G. (2010). The naked and the dead: the ABCs of gymnosperm reproduction and the origin of the angiosperm flower. *Sem. Cell Dev. Biol.* 21, 118–128. doi: 10.1016/j.semcdb.2009.11.015

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler, A., et al. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Bio. Evol.* 37 (5), 1530–1534. doi: 10.1093/molbev/msaa131

Mizzotti, C., Ezquer, I., Paolo, D., Rueda-Romero, P., Guerra, R. F., Battaglia, R., et al. (2014). *SEEDSTICK* is a master regulator of development and metabolism in the arabidopsis seed coat. *PloS Genet.* 10 (12), e1004856. doi: 10.1371/journal.pgen.1004856

Nakano, M., Hirakawa, H., Fukai, E., Toyoda, A., Kajitani, R., Minakuchi, Y., et al. (2021). A chromosome-level genome sequence of chrysanthemum seticuspe, a model species for hexaploid cultivated chrysanthemum. *Commun. Biol.* 4 (1), 1167. doi: 10.1038/s42003-021-02704-y

Nishiyama, S., Matsushita, M. C., Yamane, H., Honda, C., Okada, K., Tamada, Y., et al. (2021). Functional and expressional analyses of apple *FLC*-like in relation to dormancy progress and flower bud development. *Tree Physiol.* 41, 562–570. doi: 10.1093/treephys/tpz111

Pelaz, S., Gustafson-Brown, C., Kohalmi, S. E., Crosby, W. L., and Yanofsky, M. F. (2001). APETALA1 and SEPALLATA3 interact to promote flower development. *Plant J.* 26, 385–394. doi: 10.1046/j.1365-313X.2001.2641042.x

Pinyopich, A., Ditta, G. S., Savidge, B., Liljegren, S. J., Baumann, E., Wisman, E., et al. (2003). Assessing the redundancy of MADS-box genes during carpel and ovule development. *Nature* 424, 85–88. doi: 10.1038/nature01741

Qi, F., Liu, Y., Luo, Y., Cui, Y., Lu, C., Li, H., et al. (2022). Functional analysis of the *ScAG* and *ScAGL11* MADS-box transcription factors for anthocyanin biosynthesis and bicolour pattern formation in *Senecio cruentus* ray florets. *Hortic. Res.* 9. doi: 10.1093/hr/uhac071

Qi, S., Yang, L., Wen, X., Hong, Y., Song, X., Zhang, M., et al. (2016). Reference gene selection for RT-qPCR analysis of flower development in *Chrysanthemum morifolium* and *Chrysanthemum lavandulifolium*. *Front. Plant Sci.* 7. doi: 10.3389/fpls.2016.00287

Ren, Z., Yu, D., Yang, Z., Li, C., Qanmber, G., Li, Y., et al. (2017). Genome-wide identification of the MIKC-type MADS-box gene family in *Gossypium hirsutum* l. unravels their roles in flowering. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.00384

Reyes-Chin-Wo, S., Wang, Z., Yang, X., Kozik, A., Arikit, S., Song, C., et al. (2017). Genome assembly with in vitro proximity ligation data and whole-genome triplication in lettuce. *Nat. Commun.* 8 (1), 1–11. doi: 10.1038/ncomms14953

Rounsley, S. D., Ditta, G. S., and Yanofsky, M. F. (1995). Diverse roles for MADS box genes in arabidopsis development. *Plant Cell* 7, 1259–1269. doi: 10.1105/tpc.7.8.1259

Ruokolainen, S., Ng, Y. P., Albert, V. A., Elomaa, P., and Teeri, T. H. (2010a). Large Scale interaction analysis predicts that the gerbera hybrida floral e function is provided both by general and specialized proteins. *BMC Plant Bio.* 10, 1–13. doi: 10.1186/1471-2229-10-129

Ruokolainen, S., Ng, Y. P., Albert, V. A., Elomaa, P., and Teeri, T. H. (2011). Over-expression of the *Gerbera hybrida At-SOC1-like1* gene *Gh-SOC1* leads to floral organ identity deterioration. *Ann. Bot.* 107, 1491–1499. doi: 10.1093/aob/mcr112

Ruokolainen, S., Ng, Y. P., Broholm, S. K., Albert, V. A., Elomaa, P., and Teeri, T. H. (2010b). Characterization of *SQUAMOSA*-like genes in gerbera hybrida, including one involved in reproductive transition. *BMC Plant Bio.* 10, 1–11. doi: 10.1186/1471-2229-10-128

Song, C., Liu, Y., Song, A., Dong, G., Zhao, H., Sun, W., et al. (2018). The *Chrysanthemum nankingense* genome provides insights into the evolution and diversification of chrysanthemum flowers and medicinal traits. *Mol. Plant* 11, 1482–1491. doi: 10.1016/j.molp.2018.10.003

Theißen, G., Becker, A., Di Rosa, A., Kanno, A., Kim, J. T., Münster, T., et al. (2000). A short history of MADS-box genes in plants. *Plant Mol. Biol.* 42, 115–149. doi: 10.1023/A:1006332105728

Theißen, G., Melzer, R., and Rümpler, F. (2016). MADS-domain transcription factors and the floral quartet model of flower development: linking plant development and evolution. *Development* 143 (18), 3259–3271. doi: 10.1242/dev.134080

Tian, F., Yang, D. C., Meng, Y. Q., Jin, J., and Gao, G. (2020). PlantRegMap: charting functional regulatory maps in plants. *Nucleic Acids Res.* 48, D1104–D1113. doi: 10.1093/nar/gkz1020

Uimari, A., Kotilainen, M., Elomaa, P., Yu, D., Albert, V. A., and Teeri, T. H. (2004). Integration of reproductive meristem fates by a *SEPALLATA*-like MADS-box gene. *Proc. Natl. Acad. Sci. U.S.A.* 101, 15817–15822. doi: 10.1073/pnas.040684410

van Lieshout, N., van Kaauwen, M., Kodde, L., Arens, P., Smulders, M. J., Visser, R. G., et al. (2022). *De novo* whole-genome assembly of *Chrysanthemum makinoi*, a key wild chrysanthemum. *G3* 12 (1), jkab358. doi: 10.1093/g3journal/jkab358

Vijverberg, K., Welten, M., Kraaij, M., van Heuven, B. J., Smets, E., and Gravendeel, B. (2021). Sepal identity of the pappus and floral organ development in the common dandelion (*Taraxacum officinale*; asteraceae). *Plants* 10, 1682. doi: 10.3390/plants10081682

Wang, Y., Huang, H., Ma, Y., Fu, J., Wang, L., and Dai, S. (2014). Construction and *de novo* characterization of a transcriptome of *Chrysanthemum lavandulifolium*: analysis of gene expression patterns in floral bud emergence. *Plant Cell Tissue Organ Cult.* 116, 297–309. doi: 10.1007/s11240-013-0404-1

Wang, J., Jiu, S., Xu, Y., Sabir, I. A., Wang, L., Ma, C., et al. (2021). *SVP*-like gene *PavSVP* potentially suppressing flowering with *PavSEP*, *PavAP1*, and *PavJONITLESS* in sweet cherries (*Prunus avium* l.). *Plant Physiol. Bioch.* 159, 277–284. doi: 10.1016/j.plaphy.2020.12.013

Wen, X., Li, J., Wang, L., Lu, C., Gao, Q., Xu, P., et al. (2022). The *Chrysanthemum lavandulifolium* genome and the molecular mechanism underlying diverse capitulum types. *Horti. Res.* 9. doi: 10.1093/hr/uhab022

Wen, X., Qi, S., Huang, H., Wu, X., Zhang, B., Fan, G., et al. (2019a). The expression and interactions of ABCE-class and CYC2-like genes in the capitulum development of *Chrysanthemum lavandulifolium* and *C.× morifolium*. *Plant Growth Regul.* 88, 205–214. doi: 10.1007/s10725-019-00491-5

Wen, X., Qi, S., Yang, L., Hong, Y., and Dai, S. (2019b). Expression pattern of candidate genes in early capitulum morphogenesis of *Chrysanthemum lavandulifolium*. *Sci. Hortic.* 252, 332–341. doi: 10.1016/j.scienta.2019.03.064

Wollmann, H., Mica, E., Todesco, M., Long, J. A., and Weigel, D. (2010). On reconciling the interactions between *APETALA2*, miR172 and *AGAMOUS* with the ABC model of flower development. *Development* 137, 3633–3642. doi: 10.1242/dev.036673

Wu, R., Tomes, S., Karunairetnam, S., Tustin, S. D., Hellens, R. P., Allan, A. C., et al. (2017). *SVP*-like MADS box genes control dormancy and budbreak in apple. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.00477

Yang, Y., and Sun, H. (2009). The bracts of *Saussurea velutina* (Asteraceae) protect inflorescences from fluctuating weather at high elevations of the hengduan mountains, southwestern China. *Arct. Antarct. Alp. Res.* 41, 515–521. doi: 10.1657/1938-4246-41.4.515

Yu, D., Kotilainen, M., Pöllänen, E., Mehto, M., Elomaa, P., Helariutta, Y., et al. (1999). Organ identity genes and modified patterns of flower development in *Gerbera hybrida* (Asteraceae). *Plant J.* 17, 51–62. doi: 10.1046/j.1365-313X.1999.00351.x

Zhang, L., Chen, F., Zhang, X., Li, Z., Zhao, Y., Lohaus, R., et al. (2020). The water lily genome and the early evolution of flowering plants. *Nature* 577, 79–84. doi: 10.1038/s41586-019-1852-5

Zhang, T., and Elomaa, P. (2021). Don't be fooled: false flowers in asteraceae. *Curr. Opin. Plant Biol.* 59, 101972. doi: 10.1016/j.pbi.2020.09.006

Zhang, C., Wei, L., Wang, W., Qi, W., Cao, Z., Li, H., et al. (2020). Identification, characterization and functional analysis of AGAMOUS subfamily genes associated with floral organs and seed development in marigold (*Tagetes erecta*). *BMC Plant Bio.* 20, 1–17. doi: 10.1186/s12870-020-02644-5

Zhang, T., Zhao, Y., Juntheikki, I., Mouhu, K., Broholm, S. K., Rijpkema, A. S., et al. (2017). Dissecting functions of *SEPALLATA*-like MADS box genes in patterning of the pseudanthial inflorescence of *Gerbera hybrida*. *New Phytol.* 216, 939–954. doi: 10.1111/nph.14707

Zhao, Y., Li, X., Chen, W., Peng, X., Cheng, X., Zhu, S., et al. (2011). Whole-genome survey and characterization of MADS-box gene family in maize and sorghum. *Plant Cell Tissue Organ Cult.* 105, 159–173. doi: 10.1007/s11240-010-9848-8

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 06 April 2023
DOI 10.3389/fpls.2023.1183398

OPEN ACCESS

EDITED BY
Chenxia Cheng,
Qingdao Agricultural University, China

REVIEWED BY
Ning Yan,
Tobacco Research Institute (CAAS), China
Qianqian Liu,
Qingdao Agricultural University, China
Biao Gong,
Shandong Agricultural University, China

*CORRESPONDENCE
Fengde Wang
wfengde@163.com
Jianwei Gao
jwg_738@163.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Functional and Applied Plant Genomics, a section of the journal Frontiers in Plant Science

RECEIVED 10 March 2023
ACCEPTED 23 March 2023
PUBLISHED 06 April 2023

CITATION
Wang L, Zhang S, Zhang Y, Li J, Zhang Y, Zhou D, Li C, He L, Li H, Wang F and Gao J (2023) Integrative analysis of physiology, biochemistry and transcriptome reveals the mechanism of leaf size formation in Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). *Front. Plant Sci.* 14:1183398. doi: 10.3389/fpls.2023.1183398

Integrative analysis of physiology, biochemistry and transcriptome reveals the mechanism of leaf size formation in Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*)

Lixia Wang[1†], Shu Zhang[1†], Ye Zhang[2†], Jingjuan Li[1], Yihui Zhang[1], Dandan Zhou[1,3], Cheng Li[1], Lilong He[1], Huayin Li[1], Fengde Wang[1*] and Jianwei Gao[1*]

[1]Institute of Vegetables, Shandong Academy of Agricultural Sciences, Jinan, China, [2]College of Life Science, Huangshan University, Huangshan, China, [3]College of Life Sciences, Shandong Normal University, Jinan, China

Introduction: The leaf, the main product organ, is an essential factor in determining the Chinese cabbage growth, yield and quality.

Methods: To explore the regulatory mechanism of leaf size development of Chinese cabbage, we investigated the leaf size difference between two high-generation inbred lines of Chinese cabbage, Y2 (large leaf) and Y7 (small leaf). Furtherly, the transcriptome and cis-acting elements analyses were conducted.

Results and Discussion: According to our results, Y2 exhibited a higher growth rate than Y7 during the whole growth stage. In addition, the significant higher leaf number was observed in Y2 than in Y7. There was no significant difference in the number of epidermal cells and guard cells per square millimeter between Y2 and Y7 leaves. It indicated that cell numbers caused the difference in leaf size. The measurement of phytohormone content confirmed that GA1 and GA3 mainly play essential roles in the early stage of leaf growth, and IPA and ABA were in the whole leaf growth period in regulating the cell proliferation difference between Y2 and Y7. Transcriptome analysis revealed that cyclins BraA09g010980.3C (CYCB) and BraA10g027420.3C (CYCD) were mainly responsible for the leaf size difference between Y2 and Y7 Chinese cabbage. Further, we revealed that the transcription factors BraA09gMYB47 and BraA06gMYB88 played critical roles in the difference of leaf size between Y2 and Y7 through the regulation of cell proliferation.

Conclusion: This observation not only offers essential insights into understanding the regulation mechanism of leaf development, also provides a promising breeding strategy to improve Chinese cabbage yield.

KEYWORDS

leaf size, cell cycle, cyclins, MYB transcription factor, Chinese cabbage

1 Introduction

Leaves are essential to plant organs because of their roles in photosynthesis, respiration, photo-perception and transpiration (Tsukaya, 2002). Leaves provide plant growth and development energy through photosynthesis and respiration and store organic matter and mineral nutrients. Leaves arise from the shoot apical meristem (Scanlon, 2000), and the founder cells expand into a young leaf primordium in this stage (Poethig, 1997). In a previous study, the leaf outgrowth is determined by the cell division, which produces a certain number of cells with dense cytoplasm and the cell expansion, which make a specific cell size by cytoplasmic growth (Kalve et al., 2014). Cell division and cell expansion are complementary (Tsukaya, 2002; Beemster et al., 2003). Cell growth must be balanced by cell division and expansion for stable tissue growth and ideal leaf morphology (Sablowski and Dornelas, 2013). Thus, exploring the regulatory mechanism of leaf size can enrich the theoretical basis of leaf development and provide a basis for leaf size regulation and breed improvement.

The process of leaf development is a complex process regulated by genetic, environmental and plant hormonal factors, among which genetic factors are the intrinsic factors and play essential roles in leaf development (Tsukaya, 2003). The coordination of the two processes, cell division and expansion, is the basis for the final leaf size. The duration changes of either of these two processes can affect the leaf's final size. In previous reports, *cyclin* (*CYC*) genes are involved in cell division and expansion, called the endoreduplication cycle (Xie et al., 2010). The combinatorial interactions between different CYCs and CDKs promote cell cycle phases (Komaki and Sugimoto, 2012). CYCDs in Arabidopsis primarily bind to CDKA to drive the G1 to S transition, while the CYCA2 combine with CDKBs to promote the G2 to M transition (Boudolf et al., 2009; Van Leene et al., 2010). It has been confirmed that the expression of *CYCD*s necessarily correlates with the presence of mitogens (De Veylder, 2019). Therefore, *MYBs* could activate *CYCAs* and *CYCBs* to promote the cell cycle in tobacco (Takatsuka et al., 2021).

Multiple genes control the plant cell division, and this regulatory network mainly includes the ANT (AINTEGUMENTA) pathway and TCP-GRF (TEOSINTE BRANCHED/CYCLOIDEA/PCF-GROWTH REGULATION FACTOR) pathway (Powell and Lenhard, 2012; Czesnick and Lenhard, 2015). Previous studies showed that in the ANT pathway, the *ARGOS* (*AUXIN-REGULATED GENE INVOLVED IN ORGAN SIZE*) overexpression promoted the cell division, resulting in larger organs, and downstream the *AXR1 (auxin-resistant 1)* to mediate cell division (Hu et al., 2003; Wang et al., 2010). As a downstream gene of *ARGOS*, *ANT* allows a more extended growth period of plant leaves by positively regulating the expression of *CYCD3*, increasing the size of plant leaves and floral organs (Mizukami and Fischer, 2000). Furthermore, ectopic expression of a *BrANT* increased the organ size and stomatal density of Arabidopsis (Ding et al., 2018). In the TCP-GRF pathway, *TCP* regulated by *miR319* negatively regulates the leaf organ size, and the overexpression of *miR319* caused a down-regulated expression of *TCP* (Class II), leading to excessive proliferation extended to the margins (Nag et al., 2009; Bresso et al., 2018). Moreover, the positive regulation of *TCP* (Class I) on cell proliferation has been identified (Li et al., 2005). In Arabidopsis, nine *GRFs* can regulate cell proliferation to promote leaf growth, *AtGRF1* and *AtGRF2* overexpression increase the leaf size, while the grf1/2/3 mutant showed a small leaf size (Kim et al., 2003). In addition, *BrGRF8* overexpression of Arabidopsis increased the leaf size by regulating cell proliferation (Wang et al., 2014). It has been reported that GRF regulates cell proliferation by the combination of GIF/AN3 (Horiguchi et al., 2005; Yuan et al., 2017). Moreover, the *miR396* also regulated the expression of *GRF* (Rodriguez et al., 2010).

Cell expansion is the second phase of leaf development: cell growth without division (Donnelly et al., 1999). An increase in cell size induced by cell expansion is essential to the complete growth and development of plant leaves (Breuer et al., 2010). Cytoplasmic accumulation and cell wall loose are vital factors affecting cell expansion (Czesnick and Lenhard, 2015). It has been reported that the expansion protein (EXP), xyloglucan endotransgluco sylase/hydrolase (XTH) and glycoside hydrolase (GH) are involved in the loose of the cell wall. The repression or overexpression of these related genes leads to changes in organ size (Goh et al., 2014; Lee et al., 2018; Hrmova et al., 2022). For example, the overexpression of *AtEXP10* promoted the increase of the cell volume and organ size, while the contrary result was observed when *AtEXP10* was repressed (Cho and Cosgrove, 2000). In addition, the disruption of *AtRPS13A (ARABIDOPSIS MUTANT OF CYTOPLASMIC RIBOSOMAL PROTEIN S13)* showed many enlarged cells and intercellular spaces in leaf blade (Ito et al., 2001). In contrary, the loss of proteases activity could inhibit the cell proliferation but not the cell volume, eventually led to the enlargement of leaves, flowers, seeds and other organs (Kurepa et al., 2009).Chinese cabbage is a typical leafy vegetable widely grown in Southeast Asian countries such as China, Japan and South Korea. Leaf size is a crucial trait affecting Chinese cabbage yield. So far, the molecular regulatory mechanisms of Chinese cabbage leaf size development remain unknown. A complete understanding of Chinese cabbage leaf development is critical to increasing the yield and editing the leaf shape of Chinese cabbage through genetic manipulation. In this study, two high-generation inbred lines of Chinese cabbage with significant differences in leaf size were used as materials. The molecular regulation mechanism that caused the leaf size difference was systematically explained through physiological, biochemical and transcriptomic analysis. This study provides help for the next step to improve the yield of Chinese cabbage by molecular breeding.

2 Materials and methods

2.1 Plant materials and physiological indexes

In this study, two high-generation inbred lines of Chinese cabbage, the large leaf size Y2 and the small leaf size Y7 were provided by our experiment field on August 20, 2020. We selected and marked the leaves with the same length (3 cm) from Y2 and Y7 on the 35th day after sowing. The length of ten marked leaves was

measured every four days eight times, which were recorded as S1-S8. Meantime, others marked leaves were selected and cut from 1.5 cm at the top of the petiole for sampling (Figure 1D). Finally, 3-5 leaves were mixed as one biological replicate.

To verify RNA-Seq results, we prepared the seedling for the RT-qPCR. Seeds of "Y2" were germinated in plastic pots (25 × 25 cm) containing soil and vermiculite at a volume ratio of 1:1 and grown in our experiment field. One-month-old seedlings with 5-8 leaves were selected, and the leaves were selected and numbered in the order L1-L8 from new to old leaves. Three leaves cut from 1.5 cm at the top of the petiole were mixed for sampling.

All the samples were frozen in liquid nitrogen immediately and stored at -80 °C for further research.

2.2 Scanning electron microscope observation and cell number statistics

The leaves of Y2 and Y7 in S8 were collected for scanning electron microscope (SEM) observation and cell number statistics. The part near the middle of the leaves was cut into 0.5 cm×0.5 cm of square pieces and adhered to the glutaraldehyde fixation solution (25% glutaraldehyde 1 ml, 0.2 M pH 7.4 phosphate buffer 5 ml, distilled water 4 ml) for more than 12 h. First, the adhered samples were washed with phosphate-buffered saline (PBS) 5 times for 20 minutes each time. Then, the samples were fixed in 1% osmic acid for 2 h and washed 3 times using PBS, 20 min each time. After that, the samples were dehydrated by gradient with 50%, 70%, 80%, 90% and absolute ethanol, 20 min each time. Finally, the samples after ethanol dehydration were subjected to carbon dioxide critical point drying and then observed and photographed by scanning electron microscope (HITACHI TM3030, Japan).

According to the statistical rule, each intact cell and stoma was counted as one, and those with fewer than one cell or stoma near the edge of the visual field were always counted as 0.5. Y2 and Y7 each had three replicates.

2.3 The measurement of plant hormones

The freeze leaves of Y2 and Y7 cut in S1, S4 and S8 were ground to power with liquid nitrogen. First, 1.0 g power was added into a glass test tube and 10 ml isopropanol/hydrochloric acid buffer, and 8 μL 1 μg/mL internal standards, vibrating at 4 °C for 30 min. Then, we added 20 ml methylene chloride and vibrated at 4 °C for 30 min. The mixture was centrifuged at 4 °C, 13000 rpm for 5 min. The lower phase was taken and blown to dry with nitrogen in the dark. The samples were dissolved in 400 μL methanol (0.1% formic acid) and filtered with 0.22 μm film to measure IAA, IBA, IPA, ZT, TZR, GA1, GA3, GA4, GA7 and ABA. Plant hormones' content was measured using HPLC-MS/MS (HPLC, Aglient1290; MS/MS: SCIEX-6500Qtrap). The standard curve was used to calculate the content. Three replicates of each assay were performed.

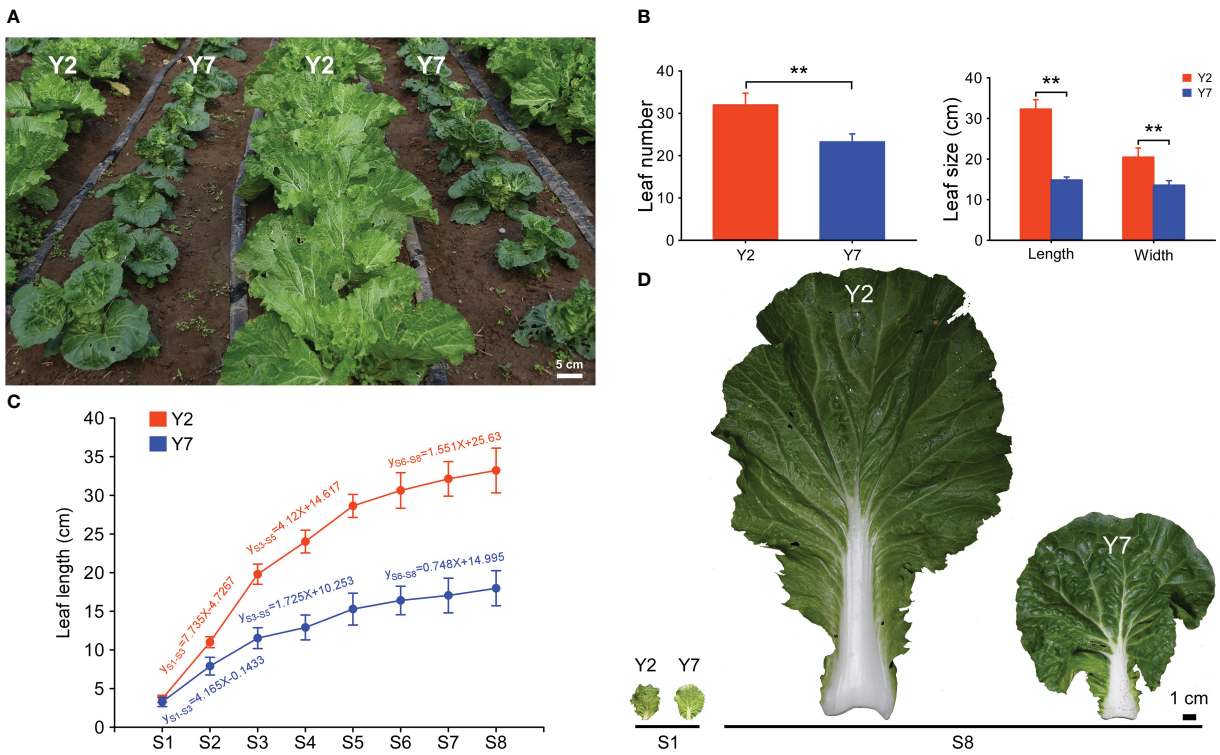


FIGURE 1
The growth of high-generation inbred lines Y2 and Y7. **(A)** Phenotypes of Y2 and Y7 in the field. **(B)** Statistics of leaf number and size 50 days after sowing. **(C)** Leaf length and growth rate of Y2 and Y7 from S1-S8. **(D)** The leaf of Y2 and Y7 at S1 and S8. * showed significant difference (Student's t-test, **P < 0.01).

2.4 Transcriptome analysis

According to the manual, the S1-S8 leaves of Y2 and Y7 were collected for total RNA extraction using TRIzol reagent (Invitrogen, Carlsbad, CA, USA). Three replicates for each sample were analysed. Barcoded cDNA libraries were constructed using NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA), and the 150bp paired-end reads were sequenced on NovaSeq 6000 platforms (Illumina). After the raw data were filtered, the sequencing error rate and the GC content distribution were checked, clean reads were obtained for subsequent analysis, and the mapped data were obtained by sequence alignment with the Chinese cabbage reference genome (Brara_Chiifu_V3.0, http://brassicadb.cn) using HISAT2. Use featureCounts to calculate the gene alignment. Fragments per kilobase of the transcript per million fragments mapped (FPKM) values were used to indicate transcript or gene expression levels. Principal component analysis (PCA) was conducted to evaluate the variation degree among the samples and groups. Weighted gene co-expression network analysis (WGCNA; PCC ≥ 0.8, minModuleSize = 30, cutHeight = 0.25) was performed on the transcriptome data to obtain co-expression gene modules and identify the network of genes regulating the leaf size of Chinese cabbage. The structural genes and transcription factors were organized into a connection network using Cytoscape software (Cytoscape 3.4.0). TBtool software was used to make the heatmap of gene expression. The RNA-seq data have been deposited in the NCBI Sequence Read Archive (NCBIvSRA) under accession number PRJNA895601.

2.5 RT-qPCR analysis

The total RNA of Y2 and Y7 leaves was extracted and used as a template, and a Takara Kit (PrimeScript 1st strand cDNA Synthesis Kit) was used to reverse-transcribe RNA into cDNA. Reactions were carried out using SYBR Green I Mix (QIAGEN) and ABI real-time quantitative PCR system. The analysis of each sample was repeated three times, and the $2^{-\Delta Ct}$ method was used for quantitative data analysis. The Actin gene of Chinese cabbage (BraActin) was used as an internal reference gene. In this study, all the primers (Qingdao WeiLai Biotechnology Co., Ltd.) are shown in "Supplemental File 1". Three biological replicates were performed in all the experiments in this study.

2.6 Cis-elements analysis

DNA sequences of 2000 bp in upstream regions of the start codon ("ATG") of cyclin genes were obtained from BRAD (http://brassicadb.cn) as promoters. PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) was used to predict cis-elements in promoter regions of cyclin genes (Supplemental Files 2, 3).

2.7 Statistical analysis

Three biological replicates were performed in all the experiments in this study. Statistical significance (Student's t-test) and Pearson correlation coefficients were analyzed by using SPSS v24.0 software (SPSS Inc., Chicago, IL, USA), and a difference was considered to be statistically significant when $P < 0.05$ or $P < 0.01$.

3 Results

3.1 The leaf phenotype of Y2 and Y7

The growth of high-generation inbred lines Y2 and Y7 is shown in Figure 1A. 50 days after sowing, the leaves more significant than 2 cm in length of Y2 were more than those of Y7 (Figure 1B). Furtherly, the size of the giant leaf of Y2, including length and width, was significantly larger than Y7 (Figure 1B). For the growth trend of Y2 and Y7, the leaves (length of 3 cm) were marked 35 days after sowing, and the length of marked leaves was measured every four days for a total of eight times. As shown in Figure 1C, the elongation of leaves in Y2 and Y7 decreased gradually from S1 to S8, and a high growth rate was shown in Y2 leaves. During the growth of leaves, three phases of leaf elongation were identified, both in Y2 and Y7, S1-S3, S3-S5 and S5-S8. In these three stages, the proportion of Y2 and Y7 leaves elongation to the final leaf length was 48.94% and 46.6% (S1-S3), 25.9% and 18.9% (S3-S5), 14.8% and 13.4% (S5-S8), respectively. It indicated that S1-S3 was the critical period of determining the leaf size, followed by S3-S5, while S5-S8 showed the weakest effect.

The SEM of the epidermal cell at S8 of Y2 and Y7 leaves was conducted (Figure 2A). The Epidermal number per mm^2 and the stomata number per mm^2 did not significantly differ between Y2 and Y7 (Figures 2B, C), indicating the same size of leaf epidermal cells. Based on the above, the difference in Y2 and Y7 leaf size was explained by the more vigorous cell proliferation of Y2 than Y7.

3.2 Differential accumulation of phytohormones between Y2 and Y7

The contents of auxins (IAA, IBA), cytokinins (IPA, ZT, TZR), gibberellins (GA1, GA3, GA4, GA7) and abscisic acid (ABA) in Y2 and Y7 at S1, S4 and S8 period were determined by GC-MS (Figure 3). A significantly higher IPA and lower ABA content were observed in Y2 at S1, S4 and S8. The higher IAA, IBA, ZT, TZR, GA1, GA3, GA4 and GA7 content were observed at S1 both in Y2 and Y7, indicating their vital roles in promoting the enlargement of leaf at S1. Even though, The IAA and ZT contents between Y2 and Y7 in S1 were no statistical difference, significantly lower contents in Y2 at S4 and S8 were observed. The content of IBA was similar at S1 in Y2 and Y7 and decreased as they grew. S4 showed a significantly higher IBA content in Y2, while S8 showed the opposite result.

3.3 Quality assessment of RNA-sequencing data

To study the molecular regulatory mechanisms of the leaf size difference between Y2 and Y7, the transcriptome analysis (RNA-

10.3389/fpls.2023.1183398

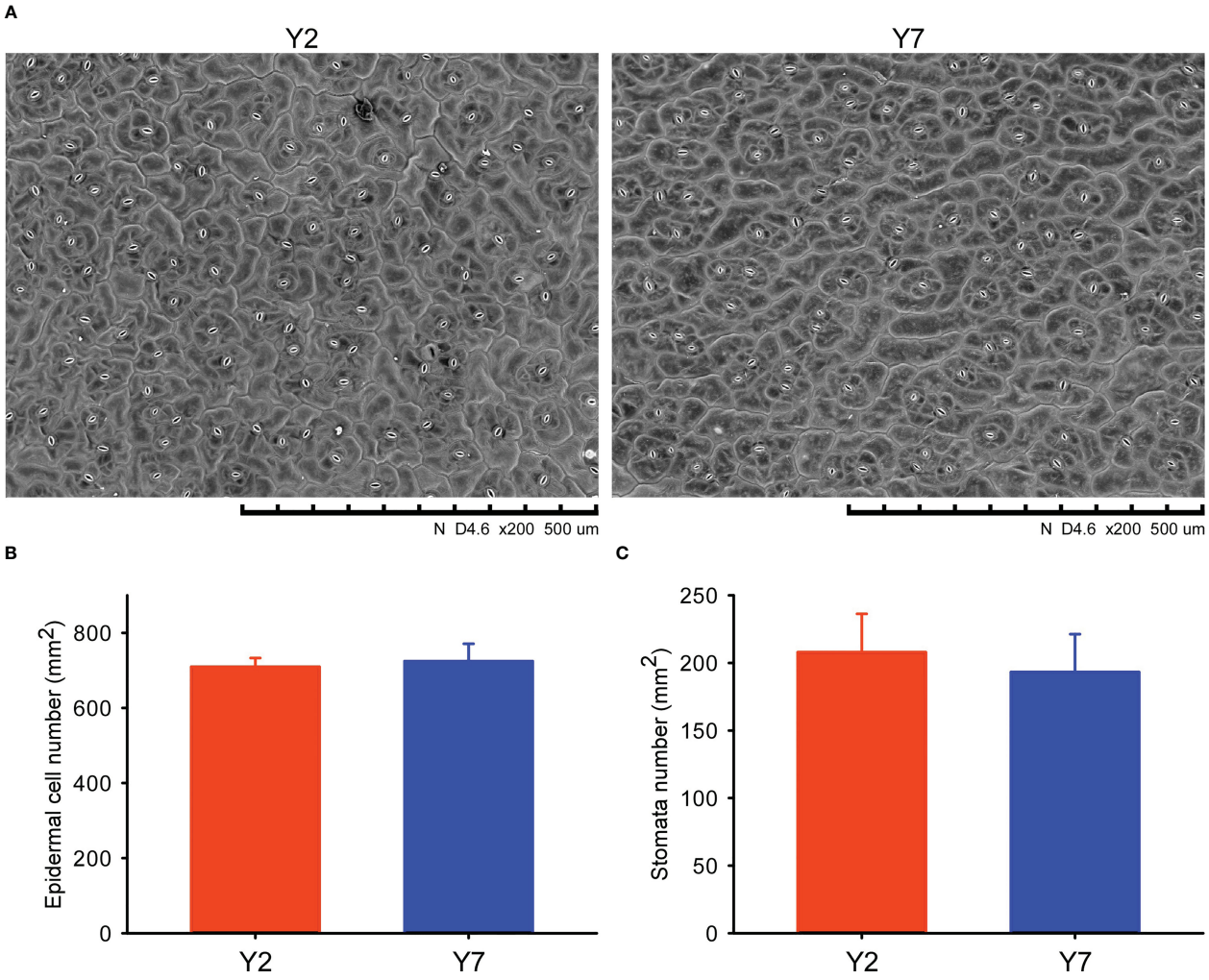


FIGURE 2
Scanning electron microscope (SEM) and statistical analyses of the Y2 and Y7 leaves. **(A)** SEM of Y2 and Y7 leaves. **(B)** Epidermal number per mm^2 of Y2 and Y7 leaves. **(C)** Stomata number per mm^2 of Y2 and Y7 leaves.

Seq) was conducted on Y2 and Y7 leaves from S1 to S8. The 48 transcriptome samples produced 2127.34 million clean data, more than 44.34 million per sample, with a percentage of Q20 and Q30 bases above 96.35 and 90.59%, respectively. Subsequently, all clean reads were compared and mapped to the Chinese cabbage reference genome sequence (http://www.brassicadb.cn/) by HISAT2 software. Transcriptional abundances were estimated using the fragments per kilobase of exon per million mapped reads (FRKM). The PCA score plot of FPKM showed that Y2 exhibited an apparent separation from Y7 in different stages, and three biological replicates of each stage were compactly gathered together (Figure 4A), indicating that the experiment was reproducible and reliable. This comparison indicated significant differences between Y2 and Y7 ($p \leq 0.05$). To verify the RNA-Seq results, we selected 13 genes with significant expression differences between Y2 and Y7. Then, the RT-qPCR investigated the expression patterns of these genes. The relative expression of these 13 genes agreed with the RNA-seq results, indicating consistency in the RNA-seq data and the qRT-PCR results (Figure 4B).

3.4 The identification of cyclins regulated the cell proliferation difference between Y2 and Y7

As mentioned above, the difference in leaf size between Y2 and Y7 is mainly due to cell division. Cyclins have been reported to involve to plant growth and development, especially Cyclin A, Cyclin B, and cyclin D play essential roles in the regulation of cell proliferation. Eighty cyclins belonging to ten subfamilies (A, B, C, D, H, L, T, U, J and SDS) were identified in Chinese cabbage. Subsequently, the phylogenetic tree of cyclins family proteins from Arabidopsis thaliana (At) and Chinese cabbage (Bra) was contrasted (Figure 5A), and 51 cyclins of Chinese cabbage belonging to A, B, and D classes were identified. Cyclin A, Cyclin B, and cyclin D of Chinese cabbage were mapped by the heatmap and clustered as their expression pattern. These cyclins were divided into three groups based on their accumulation in Y2 and Y7 at different developmental stages (Figure 5B). Cyclins in group I exhibited a high expression level in Y2 at S1 to S8, while

10.3389/fpls.2023.1183398

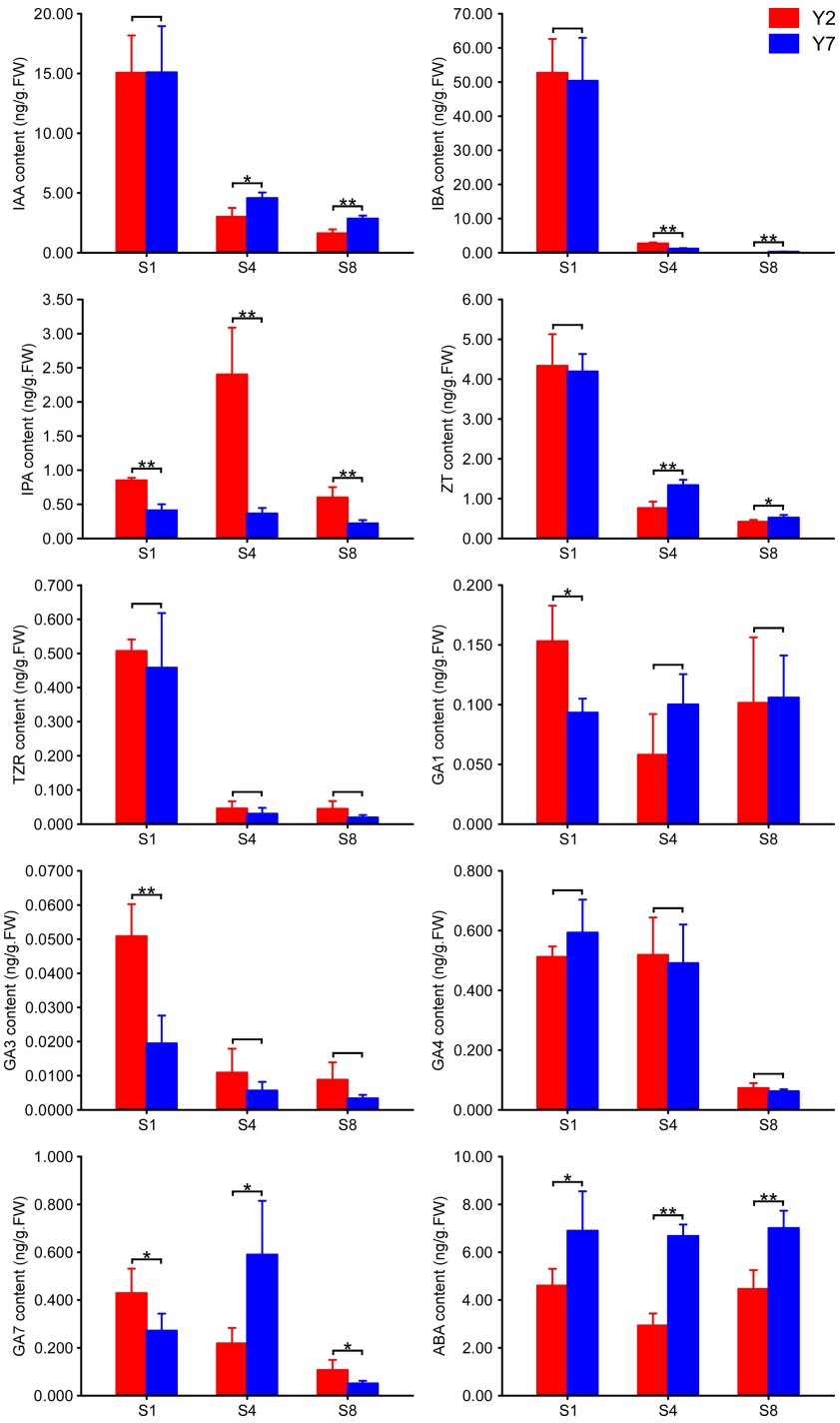


FIGURE 3
The content of plant hormones in Y2 and Y7 leaves at S1, S4 and S8. * showed significant difference (Student's t-test, *P < 0.05; **P < 0.01).

group II showed a high expression level in Y7. The cyclins in group III expressed highly at the earlier stage of leaf size development, namely, the stage of S1 and S2, then decreased gradually in Y2 and Y7 (Figure 5B). It indicates that the cyclins in group I were mainly responsible for the difference in leaf size between Y2 and Y7, and group III mainly maintained the primary state of leaf growth. Furtherly, based on the expression patterns of 51 cyclins of Chinese cabbage, the *BraA10g027420.3C* (group I), *BraA09g010980.3C* (group I), *BraA10g026850.3C* (group I), *BraA05g0030530.3C* (group I) and *BraA08g016120.3C* (group II) showed significantly high expression levels in Y2 than in that Y7, at S2-S8, S1-S2, S8, S6 and S4, respectively. Each of the five cyclins plays essential roles in different stages to regulate the leaf size difference between Y2 and Y7. According to Y2 and Y7, Chinese cabbage's growth rate, Y2 grew faster than Y7 throughout the growth period. S1-S3 showed the highest growth rate than other periods in Y2 and Y7, and S1-S3 were the critical period leading to the leaf size difference between Y2 and Y7. Fortunately, we found

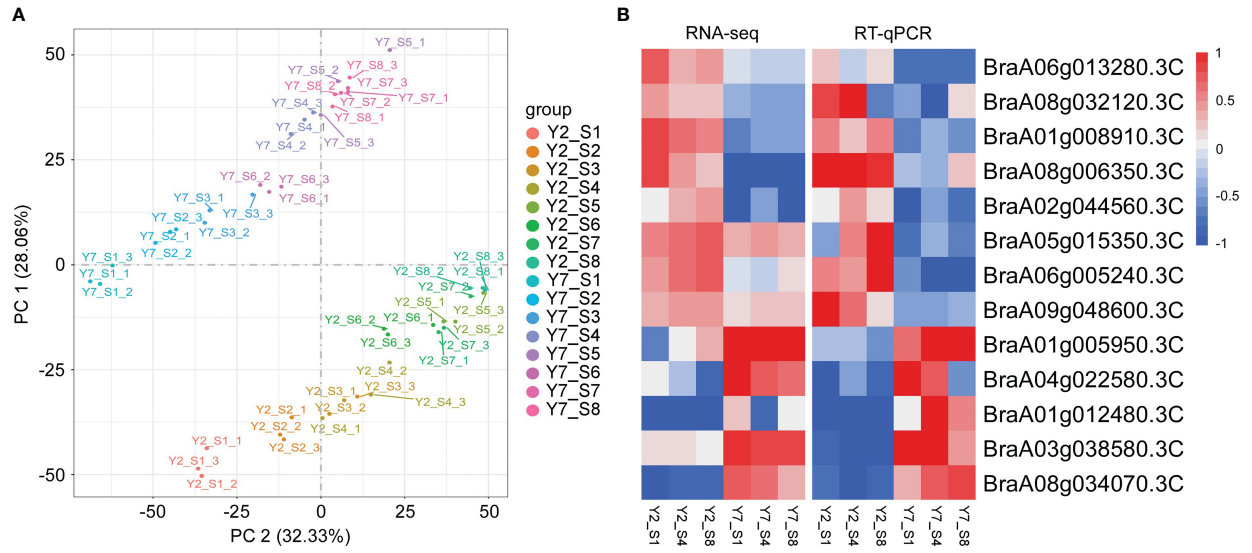


FIGURE 4
Transcriptome analysis of Y2 and Y7 from S1 to S8. **(A)** Principal component analysis (PCA) of RNA-Seq results. **(B)** The RNA-Seq and RT-qPCR results of genes were selected randomly. Each row represents the stage of Y2 and Y7. The red color indicates a positive correlation between the cluster and the sample. The blue color indicates a negative correlation.

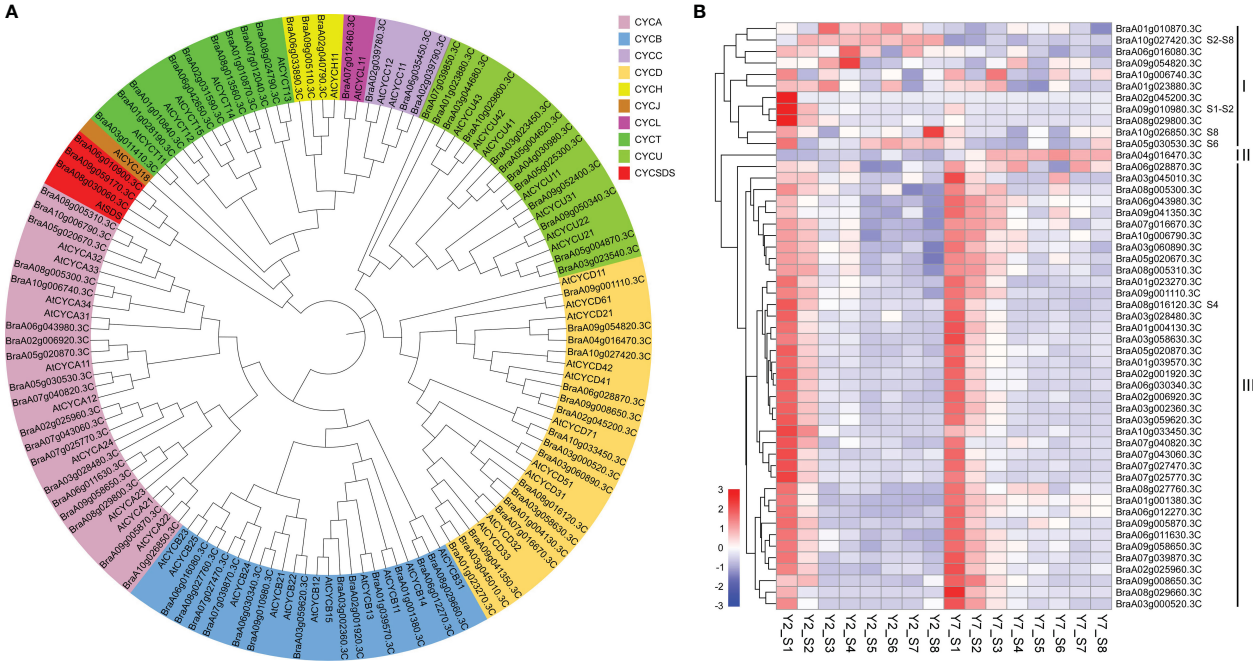


FIGURE 5
Phylogenetic analysis and expression patterns of the cyclins. **(A)** Phylogenetic tree of cyclin family proteins from Arabidopsis thaliana (At) and Chinese cabbage (Bra). The phylogenetic tree is generated by MEGA X software, according to the genetic distance model (Neighbor-joining tree). The same color in the phylogenetic tree indicates the same branch. **(B)** Heatmap representation of expression patterns in Y2 and Y7 from S1 to S8.

that the expression of *BraA09g010980.3C* was significantly higher in Y2 than in Y7 at S1-S2. Meantime, the expression of *BraA10g027420.3C* at S2-S8 showed significantly higher expression in Y2 than in Y7. *BraA10g027420.3C* and *BraA09g010980.3C* were the critical cyclin genes to regulate leaf size development.

3.5 Identification of critical transcription factors regulating leaf size of Chinese cabbage

Weighted gene co-expression network analysis (WGCNA) was used to identify the clusters of highly correlated genes. To construct

the regulation network between *BraA10g027420.3C* and *BraA09g010980.3C* and their regulators, the correlation between gene matrix of different modules and different stages of Y2 and Y7 was analyzed by WGCNA and the correlation and corresponding p-values were presented in a digital form at the intersection of modules and samples. As a result, 25406 differentially expressed genes (DEGs) between Y2 and Y7 were identified and grouped into 29 modules based on their expression patterns (Figure 6A). Interestingly, we found that the *BraA10g027420.3C* was in the "yellow", which was one module only positively correlated to the Y2 of S3-S8 (Figure 6B). On the other hand, *BraA09g010980.3C* was responsible for the leaf growth S1-S2 not located in any module.

Furtherly, to explore the critical transcription factors regulating the two structural cyclins genes and their regulatory roles, we identified transcription factors correlated to *BraA10g027420.3C* and *BraA09g010980.3C* based on the module analysis and expression pattern, respectively. Firstly, the regulator network between *BraA10g027420.3C* and correlated transcription factors was conducted by Cytoscape software. According to the weight value, the top ten transcription factors correlated to the *BraA10g027420.3C* were identified, including *BraA02gTFIIE*, *BraA06gMYB88*, *BraA05gMYB*, *BraA08gWRKY20*, *BraA06gWRKY51*, *BraA05gWRKY12*, *BraA05gbHLH2* (BASIC HELIX-LOOP-HELIX 2), *BraA04gAGL20* (AGAMOUS-LIKE 20), *BraA06gGATA11*, and *BraA08gGATA28* (Figure 7A). Based on the RNA-Seq data, except for *BraA08gGATA28*, the other nine transcription factors showed high expression levels in Y2 but low in Y7, which was in line with the structural gene of *BraA10g027420.3C* (Figure 7B). In addition, two transcription factors with the same expression pattern as *BraA09g010980.3C* were identified. The expression levels of *BraA09gMYB47* and *BraA09gbZIP* (BASIC REGION/LEUCINE ZIPPER MOTIF) at the S1 and S2 were significantly higher in Y2 than in Y7, while there was no statistic difference at other stages between Y2 and Y7 (Figure 7C).

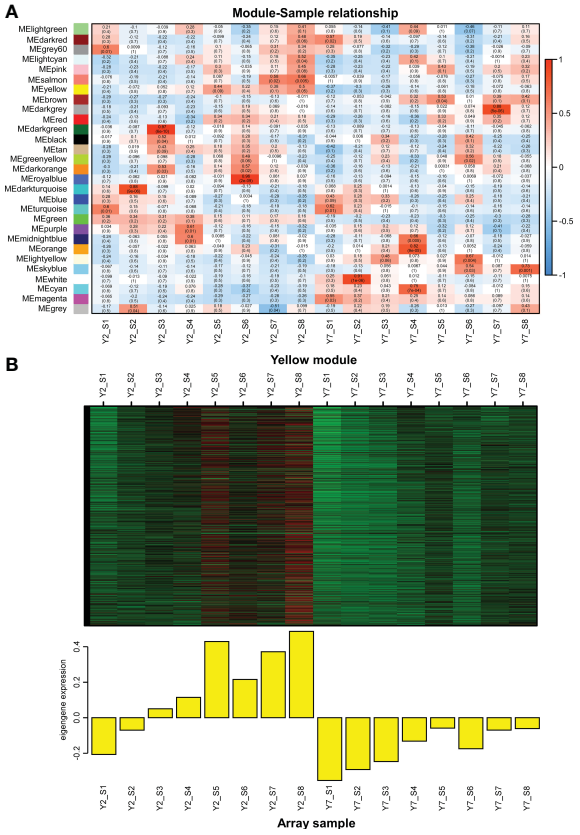


FIGURE 6
WGCNA of RNA-Seq data. **(A)** Module-Sample relationship. **(B)** The expression pattern of genes in the yellow module. The Pearson correlation coefficient (upper values) and the p-value (lower values) are indicated in the grid where each module and trait intersect. Each column corresponds to a module indicated by different colors. Each row represents the stage of Y2 and Y7. The red color indicates a positive correlation between the cluster and the sample. The blue color indicates a negative correlation.

3.6 Expression pattern and cis-elements analysis

To verify our results, RT-qPCR was used, and samples used for RT-qPCR verification were the leaves in different positions (L1-L8) of Y2 (Figure 8A). The expression of two cyclins, *BraA10g027420.3C* and *BraA09g010980.3C* and related transcription factors were confirmed (Figure 8B). For *BraA10g027420.3C*, the RT-qPCR analysis revealed the same expression trends as *BraA06gMYB88*. The expression pattern of *BraA09g010980.3C* was identical to the *BraA09gMYB47*. Subsequently, to confirm the regulation of MYB on *BraA10g027420.3C* and *BraA09g010980.3C*, we conducted the cis-elements analysis of these two essential cyclins. Interestingly, three and four MYB cis-elements were identified on the cyclin of *BraA10g027420.3C* and *BraA09g010980.3C*, respectively. The detail of these cis-elements is shown in Figure 8C. The cis-elements in the promoter of *BraA10g027420.3C* were MYB, MYB-like sequence and Myb-binding site, and the cis-elements in the promoter of *BraA10g027420.3C* were MBS, MAS-like, MYB, MYB recognition site and Myb (Figure 8C). These results confirmed the reliability of our above results.

4 Discussion

Leaf size is one of the essential traits crucial to the biomass of leaves and therefore affects Chinese cabbage yield. Therefore, leaf development has attracted the attention of botanists and biologists. The leaf development occurs from the formation of leaf primordia, followed by a period of rapid cell proliferation, and is directly related to the increase in cell number. Gradually cell proliferation becomes restricted until it stops proliferating, and then it begins to expand, with a dramatic increase in cell volume (Donnelly et al., 1999; Johnson and Lenhard, 2011). In our study, the cell division is responsible for the leaf size difference between Y2 (large leaf size) and Y7 (small leaf size). In a previous report, the leaf size difference is mainly due to the difference of cell number caused by the prolonged cell proliferation period (Powell and Lenhard, 2012).

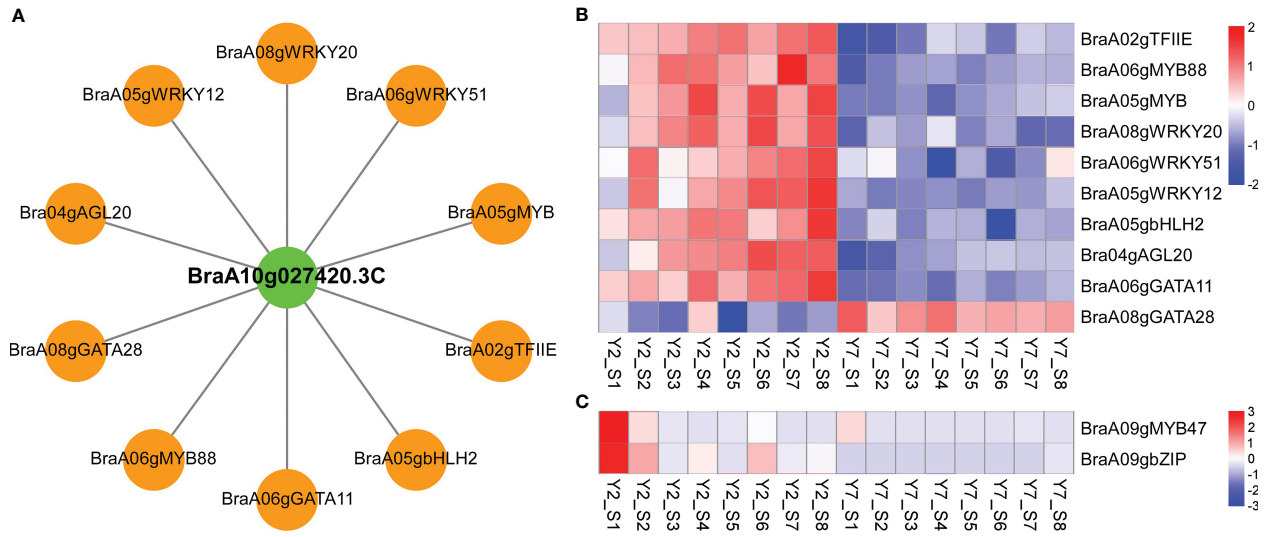


FIGURE 7
Co-expression of two cyclin genes and transcription factors involved the leaf development. **(A)** Connection network among top ten transcription factors and *BraA10g027420.3C* in Chinese cabbage. Green circles represent structural cyclin; orange circles represent transcription factors. **(B)** Transcriptome analysis of transcription factors co-expressed to *BraA10g027420.3C*. **(C)** Transcriptome analysis of transcription factors co-expressed to *BraA09g010980.3C*.

Therefore, consistent with our SEM results, the difference between Y2 and Y7 was probably caused by cell number (Figure 2A).

Phytohormones, as mobile signals, are essential regulators of leaf size by regulating cell differentiation and expansion. Cytokinins (CKs) are well-known for controlling the cell cycle (Rath et al., 2022). In our research, the IPA was the critical hormone regulating leaf size throughout the growth stage. Cytokinins regulate many cell cycle genes, such as B-type cyclins (*CYCB*), D-type cyclins (*CYCD*) and cell division cycle 2 (*cdc2*) (Kuluev et al., 2018). The CKs promote the growth of plant organs by stimulating cell proliferation, with CKs depletion or overproduction resulting in smaller or larger leaves (Werner et al., 2001; Bartrina et al., 2011). CKs and auxin interact with each in the cell cycle and leaf expansion providing positive feedback regulation of leaf growth (Davies, 2010). The crucial roles of GA1 and GA3 in regulating an earlier stage of leaf growth were confirmed in our study. GAs induced the *SAUR* (*SMALL AUXIN UP RNAs*) expression, and *SAUR* promoted cell expansion by activating plasma membrane H+-ATPase (Ren and Gray, 2015; Nagpal et al., 2022). The previous study also reported that GA could regulate an overlapping set of *SAUR* to maintain cell elongation by constructing the DELLA-ARF6-BZR1-PIF4 complex (Oh et al., 2014; van Mourik et al., 2017). It has been reported that the IAA and GA can activate the cell wall structural proteins and enzymes, such as EXP and XTH, to induce the loose of the cell wall and then the cell elongation (Kou et al., 2021). The auxin promotes organ growth by stimulating the expression of *ARGOS*, which prolongs the expression of *CYCD3* and *ANT* (Mizukami and Fischer, 2000; Hu et al., 2003). *EBP1* (ErbB-3 epidermal growth factor binding protein) could respond to auxin signals and stimulates the expression of *CyclinD3;1* (Horváth et al., 2006; Powell and Lenhard, 2012). The S1 showed a higher auxin level than S4 and S8 in Y2 and Y7, but there was no significant difference between Y2 and Y7 (Figure 3). This result indicated that the auxin should be mainly involved in the basic growth and development rather than the critical factor determining the size difference between Y2 and Y7 leaves. This coincides with reports that auxin mainly modulates cell and leaf expansion (Davies, 2010). ABA was negatively regulating the leaf size in the whole leaf growth stage. As reported in previous studies, the molecular basis of the antagonistic relationship between CK and ABA was unraveled. *SnRK* (Sucrose nonfermenting1-related kinases), as the essential positive regulators of the ABA signaling pathway, directly interacts with *ARR5* (phosphorylate type-A response regulator 5), a negative regulator of cytokinin signaling (Huang et al., 2018). Our study also demonstrated that GAs, CKs and ABA are essential in regulating the leaf size difference between Y2 and Y7. GA1 and GA3 probably play a significant role in the early stage of leaf growth, while GA and ABA act at S1-S8.

The cyclin genes are the primary regulator of plant cell cycle progression, which cooperated with cyclin-dependent kinases (Qi and Zhang, 2019). There are ten classes of cyclins in Arabidopsis consisting of approximately 49 cyclin genes has been reported, and only the classes of A, B and D were well known to regulate the cell cycle (Wang et al., 2004). The D-type is considered the regulator of the G1-to-S transition, the A-type plays a vital role in controlling the S-to-M phase, and B-type is mainly responsible for controlling the G2-to-M transition (Qi and Zhang, 2019). At different stages of the cell cycle, different CYC proteins bind to CDK kinase to form specific CYC-CDK complexes, which trigger the transition between G1/S and G2/M phases of cells, and control cell proliferation (Komaki and Sugimoto, 2012; Wood and Endicott, 2018). In our study, most cyclin genes showed a high expression level at S1 and

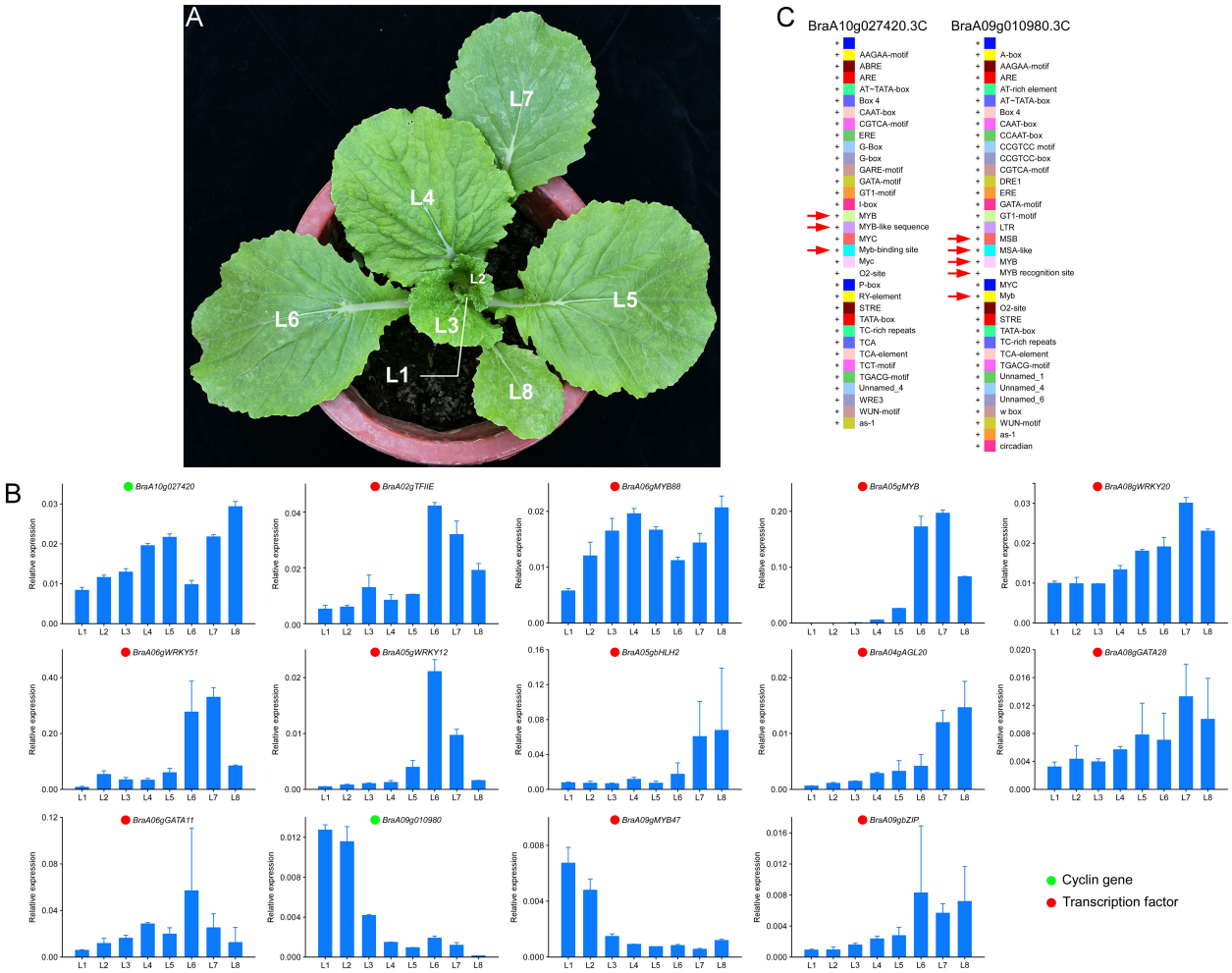


FIGURE 8
Validation of expression pattern and cis-elements analysis. **(A)** leaves positions of Y2. L1 is the youngest leaf, and L8 is the oldest leaf. **(B)** RT-qPCR validation of essential genes at different leaf positions of Y2. Green and cycles represent structural genes and transcription factors, respectively. **(C)** the cis-elements analysis of *BraA10g027420.3C* and *BraA09g010980.3C*.

S2, consistent with the highest growth vigorous proliferation of Y2 and Y7 at S1-S3. It indicated that most cyclins regulated the basic cell prefoliation of leaf in Y2 and Y7. Interestingly, we identified that the expression of *BraA09g010980.3C* (CYCB) does not belong to any module, significantly higher in Y2 than in Y7 at S1-S2, which probably the critical cyclin. In addition, *BraA10g027420.3C* (CYCD) was mainly responsible for the leaf growth at S3-S8 with a long growth duration of Y2.

Transcription factors regulate almost all major biological processes at the transcription level by binding to the cis-elements of target genes through the DNA-binding domain (DBD) (Velthuijs et al., 2021). Several transcription factors, such as *WRKY* (Yu et al., 2021), Zinc Finger of Arabidopsis thaliana (*ZAT*) (Fan et al., 2021), *MYB* (Okumura et al., 2021), *WUSCHEL-RELATED HOMEOBOX5* (*WOX5*) (Schwedersky et al., 2021), *bHLH* (Li et al., 2019) have been reported in the regulation of cell cycle, which was consistent to our study. Among these transcription factors, the expression pattern of MYBs were as consistent with the vital cyclin genes, *BraA10g027420.3C* and *BraA09g010980.3C*, thoroughly. Furtherly, the cis-elements analysis of these two cyclin genes confirmed our results that the MYB, MSA, Myb-binding site, MYB-like sequence, MBS, MAS-like, MYB recognition site and Myb cis-elements play essential roles in the regulation of cell cycle. As previously reported, *MYBs* play an important role in plant secondary metabolic regulation, hormone responses, cell differentiation and cycle regulation. In Arabidopsis, MYB3R binds to the promoters of the M phase-specific activator (MSA) of target genes to regulate the transcription of the G2/M phase-specific gene (Sumiya, 2021). In the G2/M phase, AtMYB3R1 and AtMYB3R4 active the CYCA2, CYCB1 and CYCB2 genes by recognizing and combing with MSA element (Nomoto et al., 2022). Furthermore, MYB3R1 and MYB3R4 positively regulate cytokinesis by activating *Knolle* transcription (Saito et al., 2015). In addition, *AtMYB88* encode closely related and atypical two-MYB-repeat proteins,

which, when mutated, result in excess divisions of stomata in contact (Xie et al., 2010). Thus, we speculated that the two essential cyclins promoted the cell cycle under the regulation of *MYBs* transcription factors in Chinese cabbage.

5 Conclusions

In conclusion, this study conducted a comprehensive analysis of the difference in leaf size between Y2 and Y7, and the regulatory mechanism. We confirmed that GA1 and GA3 mainly play a role in the early stage of leaf growth, while IPA and ABA play a vital role in the whole growth period of leaves in regulating the cell proliferation difference between Y2 and Y7. In addition, two essential cyclin genes that are involved in the regulation of leaf size differences between Y2 and Y7 were identified in this study. Further studies showed that the transcription factor of MYBs plays a vital role in the transcription and expression of the above two critical cyclin genes. Taken together, this study not only provided clues to understanding the molecular mechanism of leaf size regulation in Chinese cabbage but also established a foundation for improving the yield of Chinese cabbage by molecular biological methods in the future.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA895601.

Author contributions

FW and JG supervised and conceived this project. LW wrote the paper. SZ, YeZ and FW revised the paper. LW, YeZ and SZ performed the formal analysis. LW carried out qRT-PCR. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the Taishan Scholars Program of Shandong Province, China (tsqn201909167); Nature Foundation of Shandong Province (ZR2022QC113); Agricultural Science and Technology Innovation Project of SAAS (CXGC2022E08); Modern Agricultural Industrial Technology System Funding of Shandong Province, China (SDAIT-05); Agricultural Science and Technology Innovation Project of SAAS, China (CXGC2022D01); and the Prospect of Shandong Seed Project, China (2022LZGC008).

Acknowledgments

We thank JL, YZ, CL, LH, and HL for help with the performed sampling and preprocessing for transcriptome analyses. Great appreciations are given to the editor and reviewers' critical comments on the improvement of the manuscript.

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1183398/full#supplementary-material

SUPPLEMENTAL FILE 1
RT-qPCR primer.

SUPPLEMENTAL FILE 2 AND 3
Cis-elements in promoter regions of cyclin genes.

SUPPLEMENTAL FILE 4
RNA-Seq data.

SUPPLEMENTAL FILE 5
Metabolome data of plant hormones.

References

Bartrina, I., Otto, E., Strnad, M., Werner, T., and Schmülling, T. (2011). Cytokinin regulates the activity of reproductive meristems, flower organ size, ovule formation, and thus seed yield in arabidopsis thaliana. *Plant Cell* 23 (1), 69–80. doi: 10.1105/tpc.110.079079

Beemster, G. T. S., Fiorani, F., and Inzé, D. (2003). Cell cycle: the key to plant growth control? *Trends Plant Sci.* 8 (4), 154–158. doi: 10.1016/S1360-1385(03)00046-3

Boudolf, V. r., Lammens, T., Boruc, J., Van Leene, J., Van Den Daele, H., Maes, S., et al. (2009). CDKB1;1 forms a functional complex with CYCA2;3 to suppress endocycle onset. *Plant Physiol.* 150 (3), 1482–1493. doi: 10.1104/pp.109.140269

Bresso, E. G., Chorostecki, U., Rodriguez, R. E., Palatnik, J. F., and Schommer, C. (2018). Spatial control of gene expression by miR319-regulated TCP transcription factors in leaf development. *Plant Physiol.* 176 (2), 1694–1708. doi: 10.1104/pp.17.00823

Breuer, C., Ishida, T., and Sugimoto, K. (2010). Developmental control of endocycles and cell growth in plants. *Curr. Opin. Plant Biol.* 13 (6), 654–660. doi: 10.1016/j.pbi.2010.10.006

Cho, H. T., and Cosgrove, D. J. (2000). Altered expression of expansin modulates leaf growth and pedicel abscission in arabidopsis thaliana. *Proc. Natl. Acad. Sci. U.S.A.* 97 (17), 9783–9788. doi: 10.1073/pnas.160276997

Czesnick, H., and Lenhard, M. (2015). Size control in plants–lessons from leaves and flowers. *Cold Spring Harb. Perspect. Biol.* 7 (8), a019190. doi: 10.1101/cshperspect.a019190

Davies, P. J. A. (2010). *PLANT HORMONES: Biosynthesis, signal transduction, action*. Springer Dordrecht.

De Veylder, L. (2019). The discovery of plant d-type cyclins. *Plant Cell* 31 (6), 1194–1195. doi: 10.1105/tpc.19.00277

Ding, Q., Cui, B., Li, J., Li, H., Zhang, Y., Lv, X., et al. (2018). Ectopic expression of a brassica rapa AINTEGUMENTA gene (BrANT-1) increases organ size and stomatal density in arabidopsis. *Sci. Rep.* 8 (1), 10528. doi: 10.1038/s41598-018-28606-4

Donnelly, P. M., Bonetta, D., Tsukaya, H., Dengler, R. E., and Dengler, N. G. (1999). Cell cycling and cell enlargement in developing leaves of arabidopsis. *Dev. Biol.* 215 (2), 407–419. doi: 10.1006/dbio.1999.9443

Fan, Y., Yan, J., Lai, D., Yang, H., Xue, G., He, A., et al. (2021). Genome-wide identification, expression analysis, and functional study of the GRAS transcription factor family and its response to abiotic stress in sorghum [Sorghum bicolor (L.) moench]. *BMC Genomics* 22 (1), 509. doi: 10.1186/s12864-021-07848-z

Goh, H. H., Sloan, J., Malinowski, R., and Fleming, A. (2014). Variable expansin expression in arabidopsis leads to different growth responses. *J. Plant Physiol.* 171 (3-4), 329–339. doi: 10.1016/j.jplph.2013.09.009

Horiguchi, G., Kim, G. T., and Tsukaya, H. (2005). The transcription factor AtGRF5 and the transcription coactivator AN3 regulate cell proliferation in leaf primordia of arabidopsis thaliana. *Plant J.* 43 (1), 68–78. doi: 10.1111/j.1365-313X.2005.02429.x

Horváth, B. M., Magyar, Z., Zhang, Y., Hamburger, A. W., Bakó, L., Visser, R. G. F., et al. (2006). EBP1 regulates organ size through cell growth and proliferation in plants. *EMBO J.* 25 (20), 4909–4920. doi: 10.1038/sj.emboj.7601362

Hrmova, M., Stratilová, B., and Stratilová, E. (2022). Broad specific Xyloglucan: Xyloglucosyl transferases are formidable players in the re-modelling of plant cell wall structures. *Int. J. Mol. Sci.* 23 (3), pp. 1656. doi: 10.3390/ijms23031656.

Hu, Y., Xie, Q., and Chua, N. H. (2003). The arabidopsis auxin-inducible gene ARGOS controls lateral organ size. *Plant Cell* 15 (9), 1951–1961. doi: 10.1105/tpc.013557

Huang, X., Hou, L., Meng, J., You, H., Li, Z., Gong, Z., et al. (2018). The antagonistic action of abscisic acid and cytokinin signaling mediates drought stress response in arabidopsis. *Mol. Plant* 11 (7), 970–982. doi: 10.1016/j.molp.2018.05.001

Ito, T., Kim, G.-T., and Shinozaki, K. (2001). Disruption of an arabidopsis cytoplasmic ribosomal protein S13-homologous gene. *Plant J.* 22 (3), 257–264. doi: 10.1046/j.1365-313x.2000.00728.x

Johnson, K., and Lenhard, M. (2011). Genetic control of plant organ growth. *New Phytol.* 191 (2), 319–333. doi: 10.1111/j.1469-8137.2011.03737.x

Kalve, S., De Vos, D., and Beemster, G. T. (2014). Leaf development: a cellular perspective. *Front. Plant Sci.* 5, 362. doi: 10.3389/fpls.2014.00362

Kim, J. H., Choi, D., and Kende, H. (2003). The AtGRF family of putative transcription factors is involved in leaf and cotyledon growth in arabidopsis. *Plant J.* 36 (1), 94–104. doi: 10.1046/j.1365-313X.2003.01862.x

Komaki, S., and Sugimoto, K. (2012). Control of the plant cell cycle by developmental and environmental cues. *Plant Cell Physiol.* 53 (6), 953–964. doi: 10.1093/pcp/pcs070

Kou, E., Huang, X., Zhu, Y., Su, W., Liu, H., Sun, G., et al. (2021). Crosstalk between auxin and gibberellin during stalk elongation in flowering Chinese cabbage. *Sci. Rep.* 11 (1), 3976. doi: 10.1038/s41598-021-83519-z

Kuluev, B., Avalbaev, A., Nikonorov, Y., Ermoshin, A., Yuldashev, R., Akhiarova, G., et al. (2018). Effect of constitutive expression of arabidopsis CLAVATA3 on cell growth and possible role of cytokinins in leaf size control in transgenic tobacco plants. *J. Plant Physiol.* 231, 244–250. doi: 10.1016/j.jplph.2018.09.011

Kurepa, J., Wang, S., Li, Y., Zaitlin, D., Pierce, A. J., and Smalle, J. A. (2009). Loss of 26S proteasome function leads to increased cell size and decreased cell number in arabidopsis shoot organs. *Plant Physiol.* 150 (1), 178–189. doi: 10.1104/pp.109.135970

Lee, B. D., Apel, W. A., Sheridan, P. P., and DeVeaux, L. C. (2018). Glycoside hydrolase gene transcription by alicyclobacillus acidocaldarius during growth on wheat arabinoxylan and monosaccharides: a proposed xylan hydrolysis mechanism. *Biotechnol. Biofuels* 11, 110. doi: 10.1186/s13068-018-1110-3

Li, C., Potuschak, T., Colón-Carmona, A., Gutiérrez, R. A., and Doerner, P. (2005). Arabidopsis TCP20 links regulation of growth and cell division control pathways. *Proc. Natl. Acad. Sci. U.S.A.* 102 (36), 12978–12983. doi: 10.1073/pnas.0504039102

Li, T., Yang, S., Kang, X., Lei, W., Qiao, K., Zhang, D., et al. (2019). The bHLH transcription factor gene AtUPB1 regulates growth by mediating cell cycle progression in arabidopsis. *Biochem. Biophys. Res. Commun.* 518 (3), 565–572. doi: 10.1016/j.bbrc.2019.08.088

Mizukami, Y., and Fischer, R. L. (2000). Plant organ size control: AINTEGUMENTA regulates growth and cell numbers during organogenesis. *Proc. Natl. Acad. Sci. U.S.A.* 97 (2), 942–947. doi: 10.1073/pnas.97.2.942

Nag, A., King, S., and Jack, T. (2009). miR319a targeting of TCP4 is critical for petal growth and development in arabidopsis. *Proc. Natl. Acad. Sci. U.S.A.* 106 (52), 22534–22539. doi: 10.1073/pnas.0908718106

Nagpal, P., Reeves, P. H., Wong, J. H., Armengot, L., Chae, K., Rieveschl, N. B., et al. (2022). SAUR63 stimulates cell growth at the plasma membrane. *PloS Genet.* 18 (9), e1010375. doi: 10.1371/journal.pgen.1010375

Nomoto, Y., Takatsuka, H., Yamada, K., Suzuki, T., Suzuki, T., Huang, Y., et al. (2022). A hierarchical transcriptional network activates specific CDK inhibitors that regulate G2 to control cell size and number in arabidopsis. *Nat. Commun.* 13 (1), 1660. doi: 10.1038/s41467-022-29316-2

Oh, E., Zhu, J. Y., Bai, M. Y., Arenhart, R. A., Sun, Y., and Wang, Z. Y. (2014). Cell elongation is regulated through a central circuit of interacting transcription factors in the arabidopsis hypocotyl. *Elife* 3, e03031.1–19 doi: https://doi.org/10.7554/eLife.03031

Okumura, T., Nomoto, Y., Kobayashi, K., Suzuki, T., Takatsuka, H., and Ito, M. (2021). MYB3R-mediated active repression of cell cycle and growth under salt stress in arabidopsis thaliana. *J. Plant Res.* 134 (2), 261–277. doi: 10.1007/s10265-020-01250-8

Poethig, R. S. (1997). Leaf morphogenesis in flowering plants. *Plant Cell* 9 (7), 1077–1087. doi: 10.1105/tpc.9.7.1077

Powell, A. E., and Lenhard, M. (2012). Control of organ size in plants. *Curr Biol* 22 (9), R360–367. doi: 10.1016/j.cub.2012.02.010

Qi, F., and Zhang, F. (2019). Cell cycle regulation in the plant response to stress. *Front. Plant Sci.* 10, 1765. doi: 10.3389/fpls.2019.01765

Rath, M., Challa, K. R., Sarvepalli, K., and Nath, U. (2022). CINCINNATA-like TCP transcription factors in cell growth - an expanding portfolio. *Front. Plant Sci.* 13, 825341. doi: 10.3389/fpls.2022.825341

Ren, H., and Gray, W. M. (2015). SAUR proteins as effectors of hormonal and environmental signals in plant growth. *Mol. Plant* 8 (8), 1153–1164. doi: 10.1016/j.molp.2015.05.003

Rodriguez, R. E., Mecchia, M. A., Debernardi, J. M., Schommer, C., Weigel, D., and Palatnik, J. F. (2010). Control of cell proliferation in arabidopsis thaliana by microRNA miR396. *Development* 137 (1), 103–112. doi: 10.1242/dev.043067

Sablowski, R., and Dornelas, M.C. (2013). Interplay between cell growth and cell cycle in plants. *J. Exp. Bot.* 65 (10), 2703–2714. doi: 10.1093/jxb/ert354

Saito, T., Fujikawa, H., Haga, N., Suzuki, T., Machida, Y., and Ito, M. (2015). Genetic interaction between G2/M phase-specific transcription factor MYB3R4 and MAPKKK ANP3 for execution of cytokinesis in arabidopsis thaliana. *Plant Signal Behav.* 10 (3), e990817. doi: 10.4161/15592324.2014.990817

Scanlon, M. J. (2000). Developmental complexities of simple leaves. *Curr. Opin. Plant Biol.* 3 (1), 31–36. doi: 10.1016/S1369-5266(99)00040-0

Schwedersky, R. P., Saleme, M. L. S., Rocha, I. A., Montessoro, P. D. F., Hemerly, A. S., Eloy, N. B., et al. (2021). The anaphase promoting Complex/Cyclosome subunit 11 and its role in organ size and plant development. *Front. Plant Sci.* 12, 563760. doi: 10.3389/fpls.2021.563760

Sumiya, N. (2021). Cis-acting elements involved in the G2/M-phase-specific transcription of the cyclin b gene in the unicellular alga cyanidioschyzon merolae. *J. Plant Res.* 134 (6), 1301–1310. doi: 10.1007/s10265-021-01334-z

Takatsuka, H., Nomoto, Y., Araki, S., Machida, Y., and Ito, M. (2021). Identification of two tobacco genes encoding MYB3R proteins with repressor function and showing cell cycle-regulated transcript accumulation. *Plant Biotechnol. (Tokyo)* 38 (2), 269–275. doi: 10.5511/plantbiotechnology.21.0224a

Tsukaya, H. (2002). *Leaf development*. The Arabidopsis book, 1, e0072. The American Society of Plant Biologists. https://doi.org/10.1199/tab.0072.

Tsukaya, H. (2003). Organ shape and size: a lesson from studies of leaf morphogenesis. *Curr. Opin. Plant Biol.* 6 (1), 57–62. doi: 10.1016/S1369526602000055

Van Leene, J., Hollunder, J., Eeckhout, D., Persiau, G., Slijke, E.V. De, Stals, H., et al. (2010). Targeted interactomics reveals a complex core cell cycle machinery in arabidopsis thaliana. *Mol. Syst. Biol.* 6, 397. doi: 10.1038/msb.2010.53

van Mourik, H., Dijk, A.D.J.v., Stortenbeker, N., Angenent, G. C., and Bemer, M. (2017). Divergent regulation of arabidopsis SAUR genes: a focus on the SAUR10-clade. *BMC Plant Biol.* 17 (1), 245. doi: 10.1186/s12870-017-1210-4

Velthuijs, N., Meldal, B., Geessinck, Q., Porras, P., Medvedeva, Y., Zubritskiy, A., et al. (2021). Integration of transcription coregulator complexes with sequence-specific DNA-binding factor interactomes. *Biochim. Biophys. Acta Gene Regul. Mech.* 1864 (10), 194749. doi: 10.1016/j.bbagrm.2021.194749

Wang, G., Kong, H., Sun, Y., Zhang, X., Zhang, W., Altman, N., et al. (2004). Genome-wide analysis of the cyclin family in arabidopsis and comparative phylogenetic analysis of plant cyclin-like proteins. *Plant Physiol.* 135 (2), 1084–1099. doi: 10.1104/pp.104.040436

Wang, F., Qiu, N., Ding, Q., Li, J., Zhang, Y., Li, H., et al. (2014). Genome-wide identification and analysis of the growth-regulating factor family in Chinese cabbage (Brassica rapa l. ssp. pekinensis). *BMC Genomics* 15 (1), 807. doi: 10.1186/1471-2164-15-807

Wang, B., Zhou, X., Xu, F., and Gao, J. (2010). Ectopic expression of a Chinese cabbage BrARGOS gene in arabidopsis increases organ size. *Transgenic Res.* 19 (3), 461–472. doi: 10.1007/s11248-009-9324-6

Werner, T., Motyka, V., Strnad, M., and Schmülling, T. (2001). Regulation of plant growth by cytokinin. *Proc. Natl. Acad. Sci. U.S.A.* 98 (18), 10487–10492. doi: 10.1073/pnas.171304098

Wood, D. J., and Endicott, J. A. (2018). Structural insights into the functional diversity of the CDK-cyclin family. *Open Biol.* 8 (9), 2046–441. doi: 10.1098/rsob.180112

Xie, Z., Lee, E., Lucas, J. R., Morohashi, K., Li, D., Murray, J. A. H., et al. (2010). Regulation of cell proliferation in the stomatal lineage by the arabidopsis MYB FOUR LIPS via direct targeting of core cell cycle genes. *Plant Cell* 22 (7), 2306–2321. doi: 10.1105/tpc.110.074609

Yu, Y., Qi, Y., Xu, J., Dai, X., Chen, J., Dong, C. H., et al. (2021). Arabidopsis WRKY71 regulates ethylene-mediated leaf senescence by directly activating EIN2, ORE1 and ACS2 genes. *Plant J.* 107 (6), 1819–1836. doi: 10.1111/tpj.15433

Yuan, Q., Zhang, C. L., Zhao, T. T., and Xu, X. Y. (2017). Research advances of GRF transcription factor in plant. *Genomics Appl. Biol.* 36 (8), 3145–3151.

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 26 April 2023
DOI 10.3389/fpls.2023.1117069

OPEN ACCESS

EDITED BY
Lin Xi,
University of Hohenheim, Germany

REVIEWED BY
Jiancan Feng,
Henan Agricultural University, China
Deguo Han,
Northeast Agricultural University, China

*CORRESPONDENCE
Steve van Nocker
vannocke@msu.edu

†PRESENT ADDRESSES
Songwen Zhang,
Department of Pharmacology,
University of Washington, Seattle, WA,
United States
Christopher Gottschalk,
Agricultural Research Service, U.S.
Department of Agriculture, Appalachian
Fruit Research Station, Kearneysville, WV,
United States

SPECIALTY SECTION
This article was submitted to
Functional and Applied Plant Genomics,
a section of the journal
Frontiers in Plant Science

RECEIVED 06 December 2022
ACCEPTED 15 March 2023
PUBLISHED 26 April 2023

CITATION
Zhang S, Gottschalk C and van Nocker S
(2023) Conservation and divergence of
expression of *GA2-oxidase* homeologs in
apple (*Malus* x *domestica* Borkh.).
Front. Plant Sci. 14:1117069.
doi: 10.3389/fpls.2023.1117069

Conservation and divergence of expression of *GA2-oxidase* homeologs in apple (*Malus* x *domestica* Borkh.)

Songwen Zhang†, Christopher Gottschalk†
and Steve van Nocker*

Department of Horticulture and Graduate Program in Plant Breeding, Genetics and Biotechnology, Michigan State University, East Lansing, MI, United States

In domesticated apple (*Malus* x *domestica* Borkh.) and other woody perennials, floral initiation can be repressed by gibberellins (GAs). The associated mechanism is a major unanswered question in plant physiology, and understanding organismal aspects of GA signaling in apple has important commercial applications. In plants, the major mechanism for elimination of GAs and resetting of GA signaling is through catabolism by GA2-oxidases (GA2ox). We found that the GA2ox gene family in apple comprises 16 genes representing eight, clearly defined homeologous pairs, which were named as *MdGA2ox1A/1B* to *MdGA2ox8A/8B*. Expression of the genes was analyzed in the various structures of the spur, where flowers are initiated, as well as in various structures of seedlings over one diurnal cycle and in response to water-deficit and salt stress. Among the results, we found that *MdGA2ox2A/2B* dominated expression in the shoot apex and were strongly upregulated in the apex after treatment with exogenous GA_3, suggesting potential involvement in repression of flowering. Several *MdGA2ox* genes also showed preferential expression in the leaf petiole, fruit pedicel, and the seed coat of developing seeds, potentially representing mechanisms to limit diffusion of GAs across these structures. In all contexts studied, we documented both concerted and distinct expression of individual homeologs. This work introduces an accessible woody plant model for studies of GA signaling, *GA2ox* gene regulation, and conservation/divergence of expression of homeologous genes, and should find application in development of new cultivars of apple and other tree fruits.

KEYWORDS

gibberellin, GA2-oxidase, apple, floral induction, gene duplication, gene expression

1 Introduction

Gibberellins (GAs) are a class of phytohormones found in all vascular plants. To date, a total of 136 GA forms have been identified, among which GA_1, GA_3, GA_4 and GA_7 are the well-studied bioactive forms in plants (reviewed in Hedden, 2020). The early steps of GA biosynthesis involve the production of *ent*-kaurene from trans-geranylgeranyl diphosphate (GGPP) by *ent*-copalyl diphosphate synthase (CPS) and *ent*-kaurene synthase (KS), and the subsequent generation of the initial GA form, GA_{12}, by *ent*-kaurene oxidase (KO) and *ent*-kaurenoic acid oxidase (KAO). The later steps involve sequential oxidation of non-bioactive, 20-carbon (C20) GA precursors to generate active C19 GAs, accomplished by two groups of 2-oxoglutarate-dependent dioxygenases (2-ODDs), GA20-oxidase (GA20ox) and GA3-oxidase (GA3ox).

In angiosperms and gymnosperms, both C20 and C19 GAs can be catabolized by a distinct group of 2-ODDs, the GA2-oxidases (GA2ox). In *Arabidopsis thaliana* (Arabidopsis), where the genetics of GA biosynthesis have been most extensively studied, GA2ox is thought to limit the accumulation of bioactive GAs, both through diversion of early intermediates from the pathway and through catabolism of bioactive forms. Although alternative pathways for inactivation of GAs have been described (Schneider and Schliemann, 1994; Zhu et al., 2006; Varbanova et al., 2007), GA2ox catabolism appears to be the predominant route for the elimination of bioactive GAs and resetting of GA signaling. In plant tissues, appropriate levels of GAs are maintained through a homeostatic mechanism of biosynthesis and catabolism, which involves the GA-dependent repression (*GA20ox*, *GA3ox*) and activation (*GA2ox*) of gene expression.

GA2ox enzymes (gibberellin 2-β-dioxygenases; EC 1.14.11.13) comprise three distinct classes (Lee and Zeevaart, 2005; Serrani et al., 2007). Class I and II GA2ox specifically target C19 GAs for hydroxylation at C-2 (Rieu et al., 2008). Class III enzymes target the early intermediate C20-GAs, GA_{12} and GA_{53} (Schomburg et al., 2003). Distinction between Class I and Class II enzymes has been based on amino acid sequence and catalytic activity (Thomas et al., 1999; Ubeda-Tomás et al., 2006; Serrani et al., 2007). It is estimated that C19-GA2ox evolved prior to the establishment of gymnosperms followed by the appearance of C20-GA2ox (Class III) and subclassification of C19-GA2ox into Classes I and II before the emergence of angiosperms, with gene copy number within each class expanding in further duplication events (Huang et al., 2015; Yoshida et al., 2020; Hernández-García et al., 2021). In Arabidopsis, Class I enzymes include GA2ox1, GA2ox2, and GA2ox3, whereas Class II enzymes include GA2ox4 and GA2ox6 (*GA2ox5* is a pseudogene). Class III comprises GA2ox7, GA2ox8 and the recently characterized GA2ox9 and GA2ox10 (Lange et al., 2020).

By their ability to reduce or eliminate bioactive GAs, the GA2ox enzymes have potential to govern the domains of bioactive GAs, and thus may have a paramount role in influencing patterns of GA trafficking and GA-dependent processes. Previous research has identified diverse roles for *GA2ox* genes in various plants. For example, in both rice and Arabidopsis, localized expression of specific *GA2ox* genes at the base of the meristem isolates the meristem from surrounding GA-rich tissues (Sakamoto et al., 2001; Jasinski et al., 2005). In Arabidopsis, touch-induced expression of the *AtGA2ox7* gene mediates GA-dependent thigmomorphogenesis (Lange and Lange, 2015), while induction of the same gene under salt stress represses growth, potentially to allow for better survival (Magome et al., 2008). Thus, understanding the genomic complement of the *GA2ox* genes, and when and where the individual genes are expressed, should give new insight into functions of GAs in woody plant growth, development and physiology, including stress responses.

In apple (*Malus x domestica* Borkh.), one of the most widely grown and economically important temperate tree fruits, GAs influence numerous aspects of development and physiology that condition production traits important for yield, fruit quality and sustainability of production. A notable example is floral induction. In apple, as well as many other tree fruit species, exogenous GAs generally have a repressive effect on floral induction (Tromp, 1982; Southwick et al., 1995; Bertelsen and Tustin, 2002; Goldberg-Moeller et al., 2013; Zhang et al., 2016; Zhang et al., 2019). In addition, in many important apple cultivars, GAs produced in the seeds of developing fruit have been thought to repress floral initiation on adjacent bourse shoots, leading to reduced flowering and fruiting the following season, a major production problem termed alternate or biennial bearing (Luckwill et al., 1969). Despite its fundamental and applied importance, relatively little is known about the components of GA signaling in apple and their relationships with flowering.

Previously, we identified four *MdGA2ox* genes, now named as *MdGA2ox1A*, *-1B*, *-2A* and *-2B*, that were rapidly and persistently induced in the apple shoot apex in response to exogenous GA_{4+7} as well as by the presence of fruits during the anticipated period of floral induction (Zhang et al., 2019; Gottschalk et al., 2021). These results reflect a feedforward GA dynamic and prompted us to initiate the development of *MdGA2ox* genes as models to study early events in GA signaling and the potential roles of *GA2ox* genes in growth and development of apple and other perennial tree species. In the current study, we identified canonical GA2ox-like genes in the apple genome, documented their genomic organization and homeology, and assessed potential functional redundancy or divergence by identifying their developmental, diurnal, and stress-related expression profiles. This study provides a biological and evolutionary perspective into GA2ox functions in apple with an emphasis on floral induction.

2 Materials and methods

2.1 Genome-wide census of apple *GA2ox* genes

Homology-based identification of *MdGA2ox* genes among all genes previously annotated in the GDDH13 reference genome used the BLAST (Basic Local Alignment Search Tool) algorithm (Altschul et al., 1990) and the GDDH13 protein (https://iris.angers.inra.fr/gddh13/downloads/GDDH13_1-1_prot.fasta), or mRNA (https://iris.angers.inra.fr/gddh13/downloads/GDDH13_1-1_mrna.fasta) reference datasets. The query sequences comprised all 14 peptide

sequences cataloged as "Gibberellin 2-beta-dioxygenase" in ExPASy (https://enzyme.expasy.org/EC/1.14.11.13): 7 GA2ox sequences from Arabidopsis, 4 from rice, 2 from pea, and 1 from bean (UniProtKB/Swiss-Prot accessions: G2OX1_ARATH, Q8LEA2; G2OX2_ARATH, Q9XFR9; G2OX3_ARATH, O64692; G2OX4_ARATH, Q9C7Z1; G2OX6_ARATH, Q9FZ21; G2OX7_ARATH, Q9C6I4; G2OX8_ARATH, O49561; G2OX1_ORYSJ, Q5W726; G2OX2_ORYSJ, Q5ZA21; G2OX3_ORYSJ, Q8S0S6; G2OX6_ORYSJ, Q7XP65; G2OX1_PEA, Q9SQ80; G2OX2_PEA, Q9XHM5; G2OX_PHACN, Q9XG83) (Supplementary File 1). At an Expect (e)-value cutoff of 1E-12, this resulted in the identification of 185 GDDH13 gene names. The designated open reading frame translation from each gene model was retrieved from the GDDH13 protein database (https://iris.angers.inra.fr/gddh13/downloads/GDDH13_1-1_prot.fasta) (Supplementary File 2) and used as a query to reciprocally interrogate a dataset of Arabidopsis open reading frame translations (TAIR10_pep_20101214; TAIR10; http://arabidopsis.org) with e<1E-12. This resulted in 118 Arabidopsis protein names (Supplementary File 3).

HMMer (http://hmmer.org) was implemented with the hmmbuild module to create a profile based on the multiple alignment of the 14 GA2ox index sequences. The profile was then used as a query with hmmsearch and an e-value of 0.01 for both whole sequence and best match domain. This approach resulted in identification of 233 apple genes. Protein sequences of these 233 homologs were retrieved from the reference genome (GDDH13_1-1_prot.fasta) using shell scripts. These sequences were scanned for domains in the Pfam database (El-Gebali et al., 2019; https://pfam.xfam.org).

As HMMer analysis did not result in identification of additional homologs, only the first set of sequences along with 14 query sequences were aligned separately, using MUSCLE (Edgar, 2004) and then analyzed for phylogenetic relationship using MEGAX (version 10.1.8) with the phylogenetic tree shown in Figures 1A, B generated using the maximum likelihood method and a bootstrap of 1000 replicates. All positions with less than 80% site coverage were eliminated (partial deletion option). There were a total of 306 positions in the final dataset.

2.2 Gene, transcript and protein model analyses

Gene models were based on the GDDH13 annotation (Daccord et al., 2017) and curated with the 'Gala' phased diploid genome (Sun et al., 2020); http://bioinfo.bti.cornell.edu/apple_genome/) as well as transcriptome datasets derived from the 'Gala' shoot apex (Zhang et al., 2019; NCBI Sequence Read Archive PRJNA299491). Gene models were visualized using IGV (version 2.8.13; Robinson et al., 2011) and SnapGene® software (version 5.3.2; from Insightful Science; available at snapgene.com). As the 'Gala' reference genome is more consistent with our transcriptome datasets in general, the sequences of the consensus gene models were used for protein structure analyses and to build the refined phylogenetic tree shown in Figure 1B. Loci with inconsistent gene models were curated based on the reads from the transcriptome datasets. Putative amino acid sequences were predicted based on the open reading frame sequences using ORF Finder (https://www.bioinformatics.org/sms2/orf_find.htm; Stothard, 2000). Protein structures from novel transcripts and curated gene models were analyzed using SMART (Simple Molecular Architecture Research Tool, http://smart.embl-heidelberg.de/; Letunic and Bork, 2018).

2.3 Plant materials and growth conditions

'Gala' apple trees ('Brookfield Gala' grafted on 'Pajam 2' rootstocks) used to detect the expression of *MdGA2ox* genes in response to GA and in spur structures were located at the Michigan State University Research Center in Clarksville, Michigan (42° 52'24"N, 85°15'30"W) and were managed in accordance with standard commercial practices for disease, insect and weed

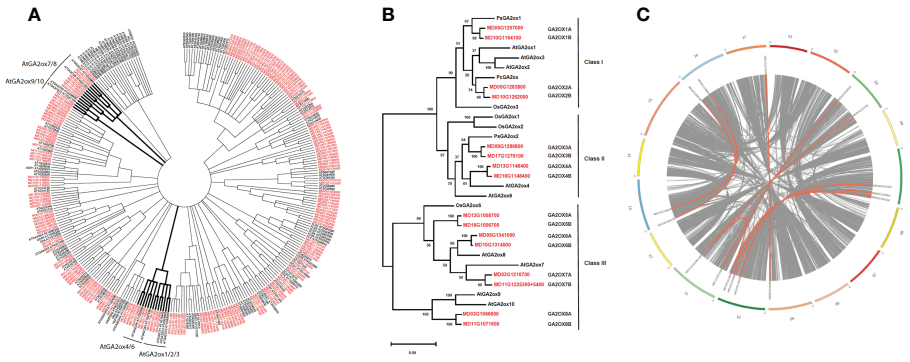


FIGURE 1
Phylogenetic analyses of GA2ox-related protein sequences. **(A)** Phylogenetic analyses of GA2ox-related protein sequences identified by BLASTp. Apple and Arabidopsis sequences are shown in red and black font, respectively. Clades containing canonical GA2ox representatives from Arabidopsis are shown in bold. A bootstrap cutoff value of 70 was used to define branches. **(B)** Phylogenetic tree of apple GA2ox protein sequences (red font) and characterized GA2ox proteins from Arabidopsis, rice, pea and bean. **(C)** Circos plot showing chromosomal position of the *MdGA2ox* genes with respect to homeologous chromosomes, which are designated with similar color. Homeologous genes are connected with an orange line.

control. The date of full bloom was defined as the date on which the maximum numbers of flowers were at anthesis. Three trees were sprayed with GA_3 at a concentration of 400 parts per million (ppm) at 30 days after full bloom (DAFB); three control trees were sprayed with water; each tree served as a biological replicate. Regulaid (0.1%) (Kalo, Overland Park, Kansas) was used as surfactant for each spray. For gene expression in spur structures, there were three biological replicates with each consisting of 15 spurs from three trees (5 spurs/tree). Fruiting spurs were dissected into 11 structures including shoot apex, bourse leaf, petiole of the bourse leaf, base of bourse leaf petiole, spur leaf, petiole of spur leaf, base of spur leaf petiole, pedicel, base of pedicel, immature seeds and fruit (minus seeds; longitudinally cut with skin, ~3 mm thick). Transverse sections (~2 mm thick) of the fully expanded bourse leaf and spur leaf that were most adjacent to the shoot apex and the largest fruit, respectively, were dissected. Non-fruiting spurs were dissected into 4 structures, including shoot apex, bourse leaf, petiole of bourse leaf and base of bourse leaf petiole. Samples were collected at 40 DAFB and placed immediately into liquid nitrogen and stored at -80°C until use. For seed structures, fruit was collected at 37 DAFB and embryo, endosperm and seed coat were dissected from about 150 immature seeds per biological replicate. Three trees served as three biological replicates.

For experiments with seedlings, seeds were obtained from fruits of open-pollinated 'Gala' trees and subjected to stratification at 4°C for three weeks in moist vermiculite (fruits were pre-stored in the cold for two months). Germinating seeds were transferred into square pots (size: 10 cm X 10 cm X 10 cm) containing artificial soil mix (Sungro Professional Growing Mix, Metro-Mix® 852). Plants were maintained in a controlled environment chamber under 16h-light/8h-dark photoperiods. Lighting was supplied with white fluorescent lights (Philips 32-watt, model: F32T8/TL741, average light intensity: 154 $\mu mol \cdot m^{-2} \cdot s^{-1}$). The temperature was held at 25°C during the light period and 18°C during the dark period. The relative humidity ranged from 25% to 28% in both the light and dark periods. Experiments to evaluate gene expression in various seedling structures, as well as diurnal expression, and changes in expression in response to water loss and salt exposure were carried out three weeks after transplanting, when plants had developed 9-11 true leaves and were ~20 cm tall. For analysis of developmental expression, structures were excised using a razor blade between 5 h and 7 h after the onset of illumination. To evaluate diurnal changes in expression, the first fully expanded leaf and 3-5 mm shoot apex were collected. Each sample comprised pooled tissue from 10-15 seedlings, and all experiments utilized three biological replicates sampled within 5 min. Dissected tissue was immediately frozen in liquid N_2.

2.4 Leaf water loss experiment

To evaluate gene expression in the leaf associated with water loss, the first fully expanded leaf was excised at the base of the petiole using a razor blade and placed on a paper towel at ambient conditions. Specifically, a total of 54 plants were randomly assigned into two groups (air drying and control). Each group consisted of eight subgroups. Each subgroup had three replicates with each replicate consisting of a total of three leaves taken from three individual plants. The leaf was weighed immediately following the cut. The "air drying" group of leaves were left on the bench in the same growth room with no change in the environmental conditions. For the control, leaves were kept hydrated by putting the petioles in water. Eight subgroups of leaves in both "air drying" and control were weighed again and then collected into liquid nitrogen for RNA extraction at 10 min, 30 min, 2 h, and 6 h. The beginning of the treatment was 5 h to 6 h after the onset of illumination.

2.5 Salt treatment

To examine the expression of *MdGA2ox* genes in response to salt, ~500 seedlings were well-watered and fertilized two days before the experiment. About 2 h prior to treatments, the plants were watered accordingly to ensure that the soil in each pot was fully saturated with water, and excessive water in the bottom tray was drained before treatments. The plants were then irrigated with 100 mL of 100 mM NaCl or 100 mL of water (control). Excess salt solution or water in the bottom tray was drained immediately following the treatment. Young leaves with a diameter of 0.5-1 cm (one leaf per seedling) were excised at 10 min, 30 min, 2 h, 6 h, 1 d, 2 d, and 6 d after treatment, respectively, and quickly collected into liquid nitrogen. There were three biological replicates at each time point in two treatment groups (salt or control) with each replicate containing 10-15 young leaves.

2.6 RNA extraction and RT-PCR

Total RNA was extracted using the method of Gasic et al. (2004) with the exception that spermine was substituted for spermidine in the extraction buffer, followed by a final 'clean-up' step using a commercial kit (RNeasy Mini; Qiagen). RNA was treated with DNase I (Qiagen) to remove genomic DNA. A small amount of RNA was checked on a Nanodrop spectrophotometer and an agarose gel for quantity, quality, and integrity. About 1 ug of RNA was then reverse-transcribed to cDNA using the High-Capacity Reverse Transcriptase Kit (ThermoFisher). Taqman primer and probe sets were designed according to the curated *MdGA2ox* transcript sequences in our 'Gala' dataset. The primer/probe design rules and quantitative RT-PCR conditions were as described in our previous study (Zhang et al., 2019).

2.7 Cluster and correlation analysis

Clustering and correlation analysis was conducted in RStudio (version 1.1.463) using packages Hmisc and corrplot, respectively. Euclidean distance followed by implementation of the Ward 2 algorithm was used for clustering, while Pearson correlation coefficients were used for correlation analysis.

3 Results

3.1 Census and genomic organization of apple *GA2ox* genes

As a first step to comprehensively identify canonical *GA2ox* genes in apple, we carried out a census of GA2ox genes as annotated for the genome of a doubled-haploid derivative of cv. 'Golden Delicious' ('GDDH13'; Daccord et al., 2017). The protein BLAST sequence homology search tool (Altschul et al., 1990) was used with 14 protein sequences from Arabidopsis, *Oryza sativa* (rice), *Pisum sativum* (pea) or *Phaseolus coccineus* (runner bean) cataloged as "Gibberellin 2-beta-dioxygenase" in the ExPASy Enzyme database as queries (Supplementary File 1). All of the 185 apple sequences showing significant homology (Supplementary File 2) were reciprocally used as queries to identify more distantly related sequences (118) from Arabidopsis (Supplementary File 3). We additionally used a Hidden-Markov-Model-based approach (HMMer; Johnson et al., 2010) with sequence motif queries derived from the conserved domains of the 14 cataloged GA2ox proteins. However, HMMer analysis did not result in identification of additional homologs (Supplementary File 4). All identified sequences from both apple and Arabidopsis were subjected to phylogenetic analyses (Figure 1A).

We found that the known Arabidopsis GA2ox sequences were represented by three clades, one containing AtGA2ox1-AtGA2ox4 and AtGA2ox6 along with eight apple sequences, one containing AtGA2ox7 and AtGA2ox8 and seven apple sequences, and one containing AtGA2ox9 and AtGA2ox10 along with two apple sequences. The AtGA2ox1/2/3/4/6 clade was further defined by two subclades, one containing AtGA2ox1-3 and four apple sequences, and the other containing AtGA2ox4 and AtGA2ox6 and four apple sequences (Figure 1A). The 17 apple sequences included in these three clades were also those that showed the strongest homology with each of the individual GA2ox query sequences, as expected. The next highest scoring apple sequences identified with BLAST were MD13G1170400, whose closest Arabidopsis homolog is LBO1 (LATERAL BRANCHING OXIDOREDUCTASE 1), and MD07G1299900 and MD01G1228800, both most closely related to Arabidopsis DMR6 (DOWNY MILDEW RESISTANCE 6) (Figure 1A). These collective results suggested that the GDDH13 reference genome annotation included 17 canonical GA2ox sequences.

To evaluate this annotation of the GDDH13 genome, transcript models for the 17 genes were generated using data from our RNA-seq-based profiling of the shoot apex of 'Gala' (Zhang et al., 2019) and mapping to both GDDH13 and a recently released, phased-diploid 'Gala' genome sequence (Sun et al., 2020) (Supplementary Files 5 and Supplementary File 6, respectively). GDDH13 and Gala annotations were generally consistent with our results, with a few exceptions (Figure 2). First, the 'Gala' genome supports a more extensive and distinct first exon of *MD09G1286800*. Second, the 'Gala' genome and our transcriptome data indicate that the transcribed region of *MD16G1148400* is more extensive and contains a third exon. Third, the 'Gala' genome annotates three exons in the transcribed region of *MD05G1341000*, rather than four as annotated in the GDDH13 genome. Fourth, the 'Gala' genome supports a more extensive transcribed region as well as a distinct third exon for *MD03G1210700*. Finally, we identified a single transcript model including both *MD11G1225300* and *MD11G1225400*. We verified that these two distinctly annotated loci generate a common transcript with an atypically long (~5.9 kbp) intron using a Taqman PCR primer-probe set (Figure 2 and Supplementary File 7). Additional transcript models were identified in the transcriptome datasets for several of the MdGA2ox genes. Most of these represented alternatively spliced forms with premature termination codons within the open reading frames (Figure 2).

Canonical GA2ox proteins comprise two highly conserved peptide sequences: an amino-terminal segment found within proteins with 2-oxoglutarate/Fe(II)-dependent dioxygenase activity (DIOX_N) and a carboxyl-terminal segment that define members of the 2OG-Fe(II) oxygenase superfamily (2OG-FEII_Oxy). A total of 20 transcript models, representing 16 genes, could encode a protein containing both domains (Figure 2 and Supplementary File 4). Two genes, *MD05G1207000* and *MD13G1008700*, produced detectable levels of apparently unspliced transcripts in which premature translation termination would lead to loss of the 2OG-FEII_Oxy domain. The DIOX_N domain was absent in the transcript model for *MD09G1286800* annotated in the GDDH13 genome, however, both domains were present according to the gene model annotated in the 'Gala' genome. The tandem reference genes *MD11G1225300* and *MD11G1225400* encode a DIOX_N domain and 2OG-FEII_Oxy domain, respectively, further supporting that these loci represent one gene. The 2OG-FEII_Oxy domain comprises the catalytic core that interacts with 2-oxoglutarate (including the conserved residues His, Asp, and His) and Fe^{2+} (Arg, Ser) (Huang et al., 2015). All of the major gene models contain these conserved residues at the expected positions (Supplementary File 8).

Based on this data, we concluded that there are a total of 16 canonical *GA2ox*-like genes in these apple genomes. A refined phylogenetic tree of apple *GA2ox* genes and selected relatives from other plants is shown in Figure 1B. This phylogeny suggested that these genes represent eight pairs of duplicated genes, and this idea is consistent with their annotated chromosomal locations within the apple genome and syntenic relationship among apple chromosomes (Figure 1C) (Daccord et al., 2017). These apple *GA2ox* genes were named based on their phylogenetic relationship and genomic organization. Four (*MdGA2ox1A/1B* and *-2A/2B*) are clustered in Class I, four (*MdGA2ox3A/3B* and *-4A/4B*) in Class II, and eight (*MdGA2ox5A/5B*, *-6A/6B*, *-7A/7B*, and *-8A/8B*) in Class III (Figure 1B). The characterization of the transcriptional activity of the genes described below omitted *MdGA2ox8A/8B*, the apple homologs of Arabidopsis *GA2ox9* and *GA2ox10*, as these Arabidopsis genes were not recognized as *GA2ox* genes when our experiments were initiated.

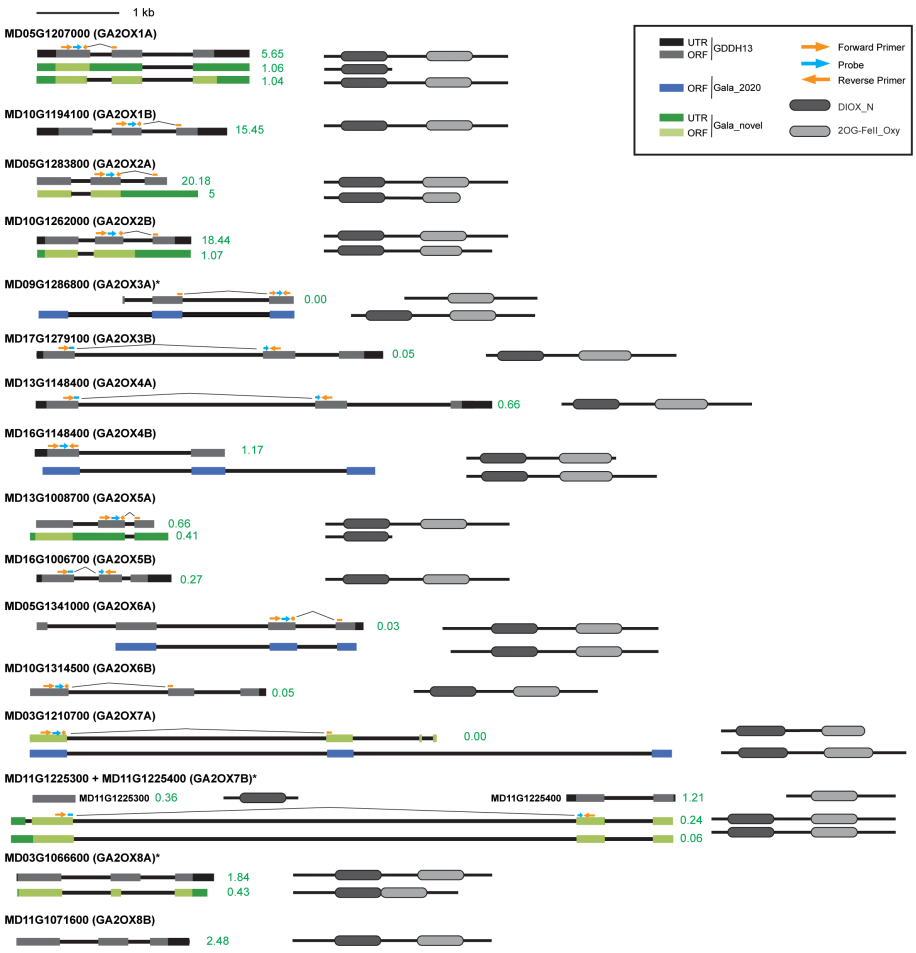


FIGURE 2
MdGA2ox gene, transcript, and protein models. Primary gene models as annotated for the GDDH13 genome are shown at left in black/gray. The gene models annotated in the 'Gala' genome are mostly consistent with the 'GDDH13' genome, except for four loci (*MdGA2ox3A*, *-4B*, *-6A* and *-7A*) where 'Gala' models are depicted in blue. Additional transcript models from 'Gala' based on our data are shown in green. Untranslated regions (UTRs) (black or dark green) and exons (gray, light green or blue) are indicated as boxes. A scale bar for transcript models is given at the top. Positions of Taqman primers (orange) and probes (light blue) are shown as arrows. Estimates of expression levels as FPKM (Fragments Per Kilobase of transcript per Million mapped reads) for each transcript in the 'Gala' shoot apex are shown in green next to the transcript models.

3.2 GA feedforward regulation of apple *GA2ox* genes

Our previous finding that the expression of at least four *MdGA2ox* genes was rapidly increased in the apple shoot apex after exposure to GA_{4+7} suggests a feed-forward mechanism. In the present study, we examined the transcriptional responses of the *GA2ox* genes in the shoot apex, two days following a foliar-applied commercial formulation of GA_3 at 30 days after full bloom (DAFB) (Figure 3). Most of the genes, including the four previously identified as GA-responsive (*MdGA2ox1A/1B*, and *-2A/2B*), were expressed to higher levels in the GA-treated plants. This upregulation by GA at this stage was most striking for the Class I/II *MdGA2ox1A/1B*, *-2A/2B*, and *-4A/4B* (Figure 3). Four genes (*MdGA2ox3A/3B*, *-6A* and *-7A*) did not show obvious upregulation by the applied GA.

3.3 Developmental expression pattern of apple *GA2ox* genes

To evaluate developmental regulation of the *MdGA2ox* genes, their expression was monitored within 11 structures of fruiting spurs at 40 DAFB (Figure 4). Spurs are condensed shoots comprising structures initiated during the previous season (spur leaves, flowers/fruit) and structures initiated in the current season (bourse shoot and floral primordia (Figure 4A). Unlike many cultivars, 'Gala' typically forms flowers on bourse shoots irrespective of the presence of developing fruit (Zhang et al., 2019), and so the bourse shoot apices analyzed here were assumed to be committed to flowering.

Nearly all of the structures analyzed showed expression of multiple *MdGA2ox* genes, with one or a few genes predominating (Figure 4B). The petiole of both spur and bourse leaves, especially

10.3389/fpls.2023.1117069

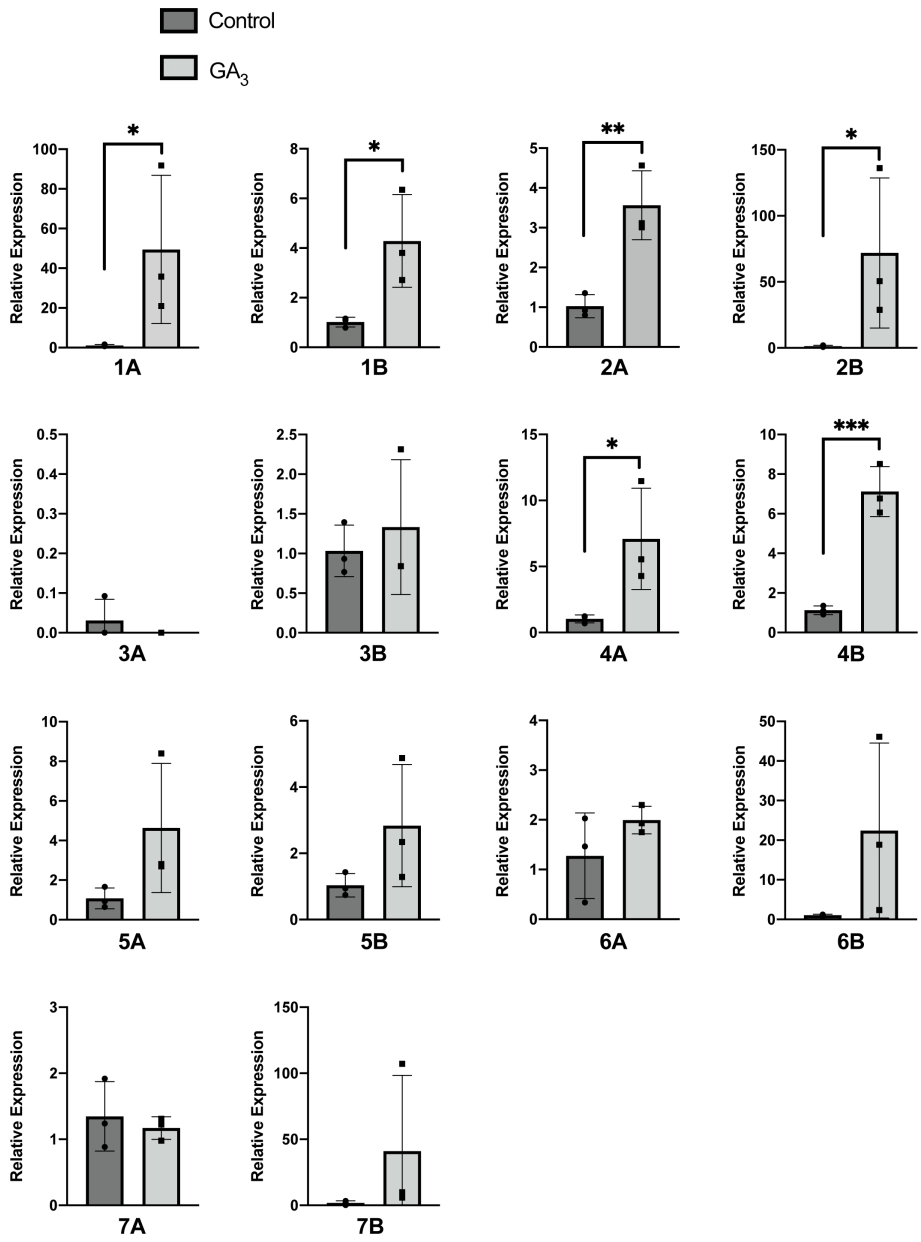


FIGURE 3
Expression of *MdGA2ox* genes in the apple shoot apex in response to applied GA_3. 'Gala' trees were subjected to a foliar application of either 400 ppm of GA_3 or water (control) at 30 DAFB, and the shoot apices were collected 2 d after treatment. Black dots represent individual values from three biological replicates. Relative expression was calculated based on the expression ratio between *GA2ox* and a reference *MdACTIN* gene. An error bar indicates standard deviation among the three biological replicates, while asterisks denote statistical significance (*, $p < 0.05$; **, $p < 0.01$, ***, $p < 0.001$).

the dissected petiole base, showed the greatest number of distinct genes expressed, whereas in the fruit (whole fruit minus seed) only one (*MdGA2ox5B*) was appreciably expressed (Figure 4B). *MdGA2ox2A/2B* and *-5A* dominated expression in the shoot apex; *-1B* and *-6A/6B* in the bourse and spur leaf; *-6A* and *-7B* in the bourse leaf petiole, *-1B* in the base of the spur leaf petiole, *-4B* in the pedicel, and *-1B* and *-4A/4B* in the immature seed. Each *GA2ox* gene also showed strong structure-specific expression (Figure 4B): *MdGA2ox1A* was preferentially expressed in the petiole, especially the base; *-1B* in the petiole base and immature seed; *-2A/2B*, *-3B*, and *-5A/5B* in the shoot apex; *-3A* in the fruit; *-4A* and *-4B* in the immature seed; and *-7B* in the base of the spur leaf petiole. For the three genes that showed strongest expression in the immature seed - *MdGA2ox1B*, *-4A*, and *-4B*, - expression was further analyzed within the seed coat, endosperm and embryo (Figure 4C) at 33 DAFB, when the embryo was patterned but prior to enlargement and desiccation. This revealed that the expression of these three genes was largely confined to the seed coat (Figure 4C).

It has been hypothesized that developing fruit comprise a source of GAs that can repress floral initiation on the proximal bourse shoot. To assess the potential effect of developing fruit, and

10.3389/fpls.2023.1117069

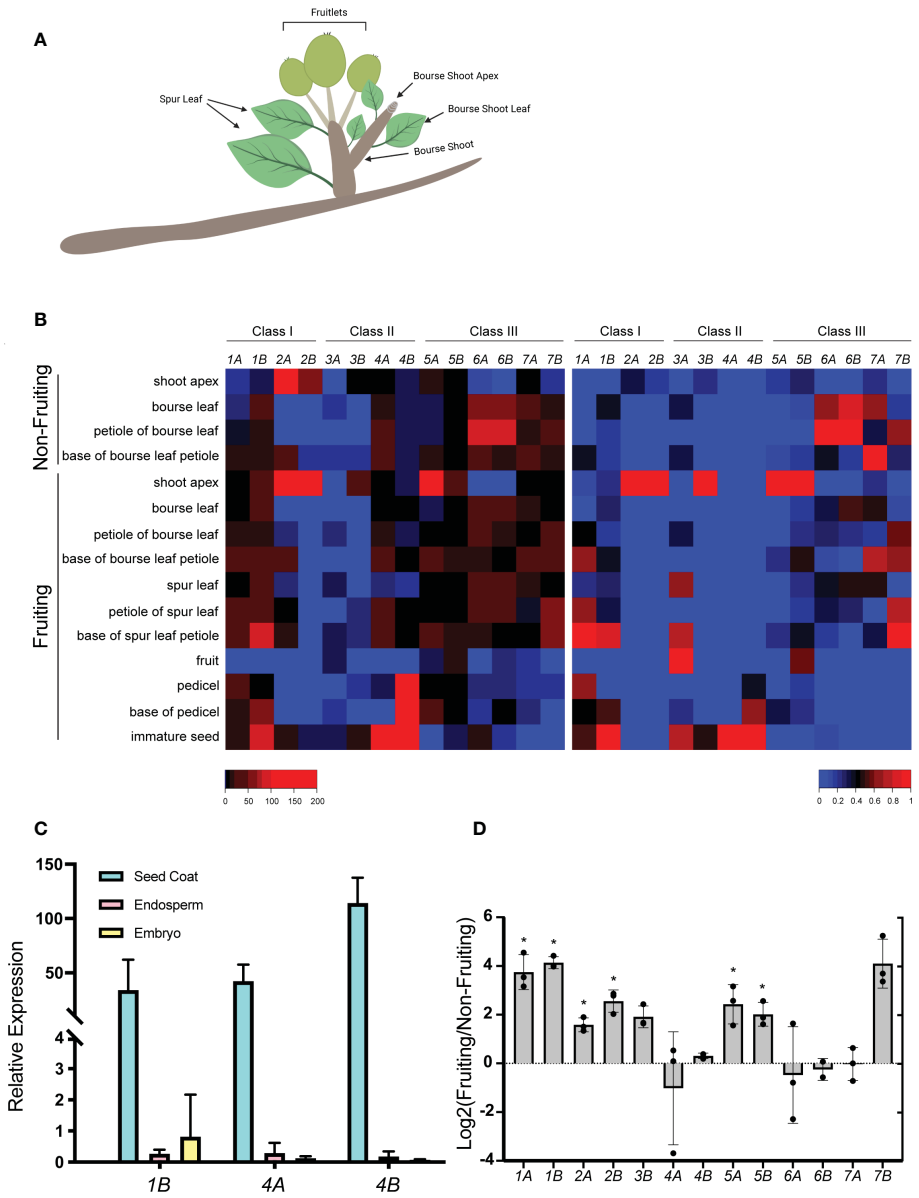


FIGURE 4
Expression profiles of *MdGA2ox* genes in spur and seed structures. **(A)**. Illustration of a fruiting apple spur at the time of analysis. **(B)**. Expression of *MdGA2ox* genes in structures dissected from non-fruiting or fruiting spurs. The heat map on the left represents expression values derived from quantitative PCR and relative to an apple *ACTIN* gene. The heat map on the right represents expression values relative to the strongest expression for each gene among structures. Color keys at the bottom indicate expression levels from low/least (blue) to high/greatest (red). **(C)**. Relative expression of *MdGA2ox1B*, *-4A*, and *-4B* in the seed coat, endosperm, and embryo of the developing seed. **(D)**. Relative *MdGA2ox* expression in fruiting and non-fruiting shoot apex. Log_2(Fruiting/Non-fruiting) represents the fold difference of relative expression in the shoot apex of fruiting spurs versus that of non-fruiting spurs. Black dots indicate individual values of three biological replicates, and error bars represent standard deviation among replicates. Asterisks indicate statistical significance (t-test, $p < 0.05$).

potentially GAs, on the expression of the 14 *MdGA2ox* genes, their expression was also monitored within non-fruiting spur structures (Figure 4B). We noted several cases in which individual *MdGA2ox* genes were expressed differentially in the same structure between fruiting and non-fruiting spurs. Most strikingly, in the shoot apex, six genes, comprising three homeologous pairs (*MdGA2ox1A/1B*, *-2A/-2B*, and *-5A/5B*), were expressed to significantly higher levels in the presence of fruit (Figure 4D). In contrast, *MdGA2ox6A* and *-6B* were expressed highly in the bourse leaf and its petiole in non-fruiting spurs, but not in fruiting spurs (Figure 4B). It is worth noting that *MdGA2ox1A/1B* and *-2A/-2B* were apparently responsive to both GA and fruit, whereas *MdGA2ox4A/-4B* were responsive to GA but not fruit, and *MdGA2ox5A/-5B* were responsive to fruit but not GA (Figures 3, Figure 4). The relative expression levels of these *MdGA2ox* genes estimated from qRT-PCR correlated well with direct counting of expressed-gene fragments, at least for the shoot apex, where RNA-seq data was available (Figure 2).

Expression of *MdGA2ox* genes was also evaluated in various structures dissected from rapidly growing seedlings that were ~6 weeks old and had produced 9-11 true leaves. These structures included the shoot apex (containing leaf primordia and leaves <5 mm in length), young structures taken from the apical section of the seedling (leaf, petiole, base of the petiole, stem node and internode) and the same structures from the older, central section of the seedling (leaf, petiole, base of the petiole, stem node and internode). In addition, stipules pooled from the younger and older sections were analyzed. Similar to the spur, nearly all of the structures analyzed showed expression of multiple *MdGA2ox* genes, with one or a few genes predominating (Figure 5). In the petiole and its base of older leaves, as well as the internode of the older stem, and stipule, the majority of the genes were strongly expressed, whereas in the shoot apex only *MdGA2ox2A* was strongly expressed. *MdGA2ox1B*, *-5A/B* and *-6B* dominated expression in the leaf, *MdGA2ox2A* and *-4A* in the base of the petiole, and *MdGA2ox2A* in the node. Also similar to the spur, each *MdGA2ox* gene also showed strong structure-specific expression (Figure 5). *MdGA2ox1A* and *-3B* were preferentially expressed in the petiole of older leaves, *-1B*, *-5A* and *-5B* in the young leaf, *-2A*, *-4A/B* and *-6A* in the base of the older leaf petiole; *-2B* in the older stem node, *-3A/B* in the young node and internode; *-3B* also in the older petiole, *-6B* and *-7A* in the leaf and in the internode of older stem; and *-7B* in the shoot apex.

Expression levels of individual genes were compared between the structures from the growing, apical section of the seedlings and the fully expanded, more basal section. In the leaf, nearly all genes were expressed to similar levels in younger and older structures. In contrast, in the petiole, base of the petiole, node, and internode, several genes were expressed to higher levels in the older structures: *MdGA2ox1A*, *-2A*, *-3B*, *-4A/B* and *-6A* in the petiole; *-1B*, *-2A*, *-4A/B*, and *-6A* in the petiole base; *-2B*, *-3B*, *-5B* and *-6B* in the node; and *-1A/B*, *-2A*, *-4A*, *-5A/B*, and *-6B* in the internode. Three of the genes, *MdGA2ox1A*, *-2B*, and *-4A*, were expressed almost exclusively only in the older structures (Figure 5).

3.4 Diurnal changes in expression

To better interpret the developmental expression patterns and guide further analyses of gene expression over time, diurnal expression of each of the 14 genes was examined over one diurnal cycle. Expression was evaluated every 3 h in the shoot apex and most apical fully expanded leaf of ~6-wk-old seedlings maintained under 16-h light, 8-h dark photoperiods. Nearly all of the genes showed apparently patterned changes in expression over the cycle in both structures (Figure 6). However, we were unable to identify a pattern that was consistently shared among the genes. *MdGA2ox1A* and *-1B* showed a convincing decrease in expression at the beginning of the light period, whereas expression of *MdGA2ox5A/5B* and *-6A* decreased gradually throughout the light period. *MdGA2ox1B*, *-2A/2B*, *-6B* and *-7A* showed gradual increase in expression during the light period. Five genes (*MdGA2ox3A/3B*, *-5A/5B*, and *-7B*) showed expression increases at the advent of, or during, the dark period, whereas *MdGA2ox1B*, *-2A*, *-6B*, and *-7A* showed a decrease in expression. These seemed to be structure specific; for example, the strong dark-associated increase in *MdGA2ox3A* expression was observed only in the apex, whereas that of *-5A* and *-5B* was seen only in the leaf (Figure 6).

Consistent with the seedling developmental experiments documented in Figure 5, each gene generally showed clear preferential expression for the apex or leaf. For example, *MdGA2ox2A/2B*, *-3A/3B*, and *-7B* had much higher expression levels in the shoot apex than that in the first fully expanded leaf (corresponding to the older leaf in Figure 5), whereas *MdGA2ox1B*, *-5A*, *-6B*, and *-7A* were more highly expressed in the leaf. However, interestingly, three genes clearly changed their preferential expression between these structures during the cycle. *MdGA2ox1B* was preferentially expressed in the shoot apex at the beginning of the light period, but in the leaf during the remains of the cycle. *MdGA2ox5A/5B* were also expressed preferentially in the

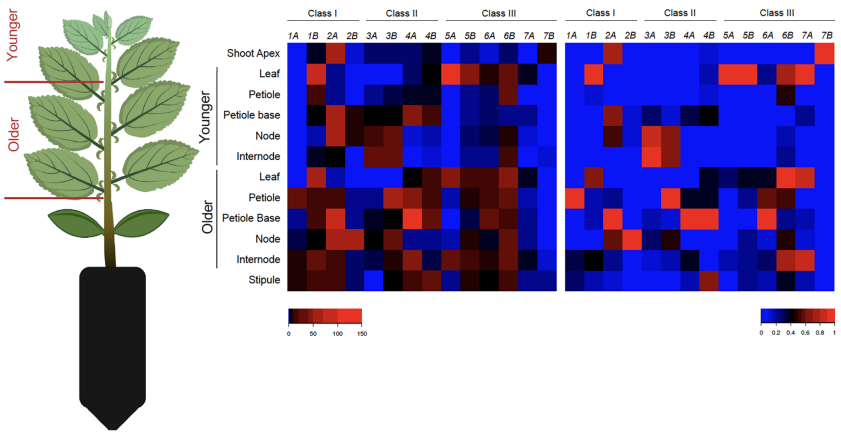


FIGURE 5
Expression profiles of *MdGA2ox* genes in apple seedling structures. Seedlings were partitioned into two general sections, Younger, comprising all structures apical to the most apical fully expanded leaf and Older, comprising this leaf and all basipetal structures. Younger and older stipules were pooled. The heat map on the left represents expression values derived from quantitative PCR and relative to an apple *ACTIN* gene. The heat map on the right represents expression values relative to the strongest expression for each gene among structures. Color keys at the bottom indicate expression levels from low/least (blue) to high/greatest (red).

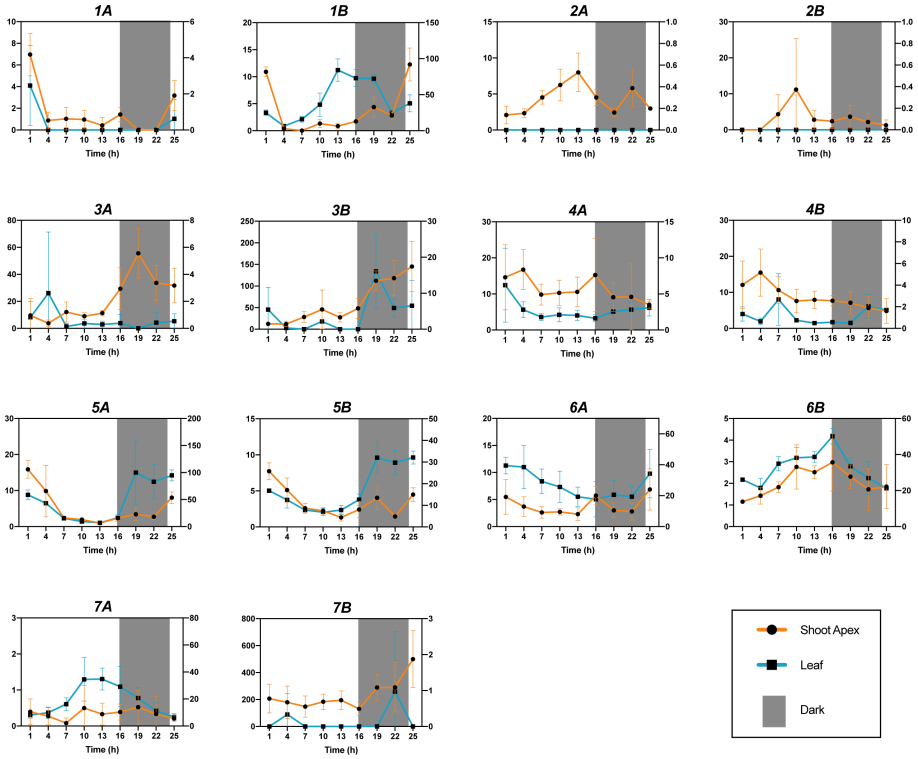


FIGURE 6
Diurnal expression of *MdGA2ox* genes in the shoot apex and first fully expanded leaf in seedlings. The shoot apex and first fully expanded leaf were excised from rapidly growing seedlings maintained under 16-h light/8-hr dark photoperiods. The y axis represents relative expression. Relative expression in the shoot apex and leaf is represented with orange and blue lines, respectively. The grey box in each graph indicates the period of darkness.

apex early in the cycle, but in the leaf during the dark period (Figure 6).

We found only limited similarity between the diurnal expression patterns of homeologous genes. For example, *MdGA2ox1B* but not *-1A* showed a clear increase during the light period, while *MdGA2ox6A/6B* showed nearly reciprocal expression patterns throughout the cycle. Additionally, *MdGA2ox7A* but not *-7B* showed a peak of expression midway through the light period (Figure 6).

Some expression results were not consistent with simple diurnal oscillation. For example, in the shoot apex both *MdGA2ox3B* and *-7B* were relatively highly expressed at 25h in contrast with 1h, even though these times are diurnally equivalent (Figure 6). One possibility to explain this is that the experiment captured a snapshot of increasing expression of these genes during this general stage of seedling development.

3.5 Expression in the leaf blade in response to leaf removal

GA2ox genes have been identified as important components limiting GA-associated growth during abiotic stress. We evaluated expression of the 14 *MdGA2ox* genes in the blade of the first fully expanded leaf of ~6-wk-old seedlings under conditions of rapid water loss. In this experiment, the leaves were excised and transferred to a paper towel, and allowed to air-dry for up to 6 hr. As a control, leaves were excised and immediately floated in water, with the cut end of the petiole immersed to maintain vascular continuity. Gene expression was assayed at 10, 30, 120, and 360 min after excision. Determination of water content in the samples (Figure 7A) indicated that water was lost from the air-drying leaves at least as early as the 10-min time point and further throughout the first 4 h of the treatment.

Expression of three genes, *MdGA2ox2A/2B* and *MdGA2ox7B*, was not reliably detected in leaf blades in either the treated or control samples, at any of the four time points. Of the remaining 11 genes, 7 showed increased expression associated with drying (Figure 7B), and all of these showed significant induction within 2 h after the beginning of the treatment (Figure 7B). This increase was most striking (~200-fold) for *MdGA2ox1A/1B*. In contrast, and unexpectedly, four genes, *MdGA2ox3A/3B*, *-4A* and *-4B*, were expressed to higher levels in the control samples than in the drying samples after 2-6 h. This pattern could not be explained by potential disruption of diurnal cycling by excision, as none of the three genes showed substantial diurnal changes in expression during the six-hour period of the experiment (Figure 6).

Expression of the 14 genes was also evaluated in response to salt exposure. Six-week-old seedlings growing in soil were subjected to daily watering with 100 mM NaCl, and gene expression was measured in young leaves at 10 min, 30 min, 2 h, 6 h, 1 d, 2 d, and 6 d after the beginning of the treatment. This NaCl

10.3389/fpls.2023.1117069

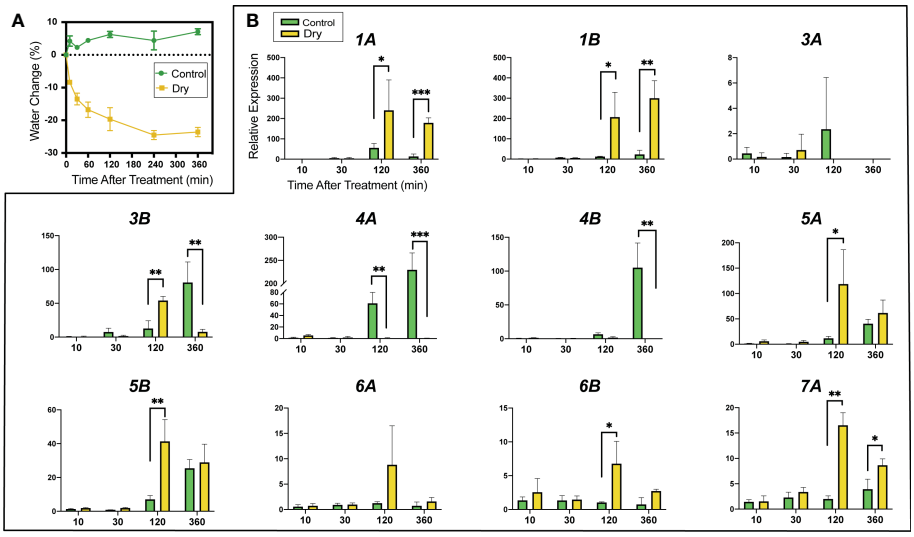


FIGURE 7
Expression of *MdGA2ox* genes in the leaf blade after leaf excision and induced water loss. The first fully expanded leaf was excised from rapidly growing seedlings and either left on a bench to air dry or kept hydrated with the tip of petiole immersed in water. **(A)**. Changes of water content in the control (green line) and drying (yellow line) leaves. Leaf water change is defined as the difference between the after-treatment weight and the initial weight. **(B)**. Relative gene expression in the control (green bar) and drying (yellow bar) leaves over time. Expression was analyzed at 10 min, 30 min, 120 min (2 h) and 360 min (6 h) after excision. Asterisks denote statistical significance (*, $p < 0.05$; **, $p < 0.01$, ***, $p < 0.001$).

concentration was chosen because it was the minimum concentration that noticeably reduced growth in seedlings over a one-week period. Compared with the results of leaf excision, the expression response to salt was subtle (Figure 8). *MdGA2ox1A* and *-1B* were induced as early as 10 min after treatment, whereas *MdGA2ox6B* was downregulated at 10 min but upregulated at 30 min and 6 h after treatment. None of these three genes were differentially expressed at later time points (Figure 8). Unlike *MdGA2ox1A*, *-1B* or *-6B*, *MdGA2ox4A* and *-4B* were significantly downregulated only at later time points, ~2-6 d after treatment (Figure 8).

3.6 Cluster and correlation analysis

To provide an unbiased view of the degree of relationship in expression pattern exhibited by these seven *MdGA2ox* homeologous pairs and individual homeologs, we carried out a clustering analysis of the developmental, diurnal and stress-related experiments described above. *MdGA2ox4A/4B* showed generalized expression characteristics easily distinguishable from the remaining MdGA2ox genes, while several non-homeologous genes fell into the same clade (Figure 9A). For homeologs, the strongest correlation was identified between *MdGA2ox2A/2B*, and strong correlation was also found for *MdGA2ox6A/6B*, *-4A/4B* and *-1A/1B*. In contrast, expression of *MdGA2ox7A/7B* was slightly anticorrelated (Figure 9B).

4 Discussion

Studies to date suggest that GA2ox proteins are key determinants of GA accumulation in plant tissues - Class I/II GA2ox by catabolizing bioactive C19 GAs, and Class III GA2ox by diverting early C20 intermediates from the pathway. Based on our results from two high quality genome sequences and exhaustive transcriptional profiling, coupled with phylogenic analyses, apple contains four members of Class I (MdGA2ox1A/1B, MdGA2ox2A/2B), four of Class II (MdGA2ox3A/3B, MdGA2ox4A/4B), and eight of Class III (MdGA2ox5A/5B, MdGA2ox6A/6B, MdGA2ox7A/7B, MdGA2ox8A/8B). In the absence of post-translational regulation, the expression pattern of the corresponding genes documented here should provide important clues to GA function at the whole-plant level. However, paradoxically, strong expression of *GA2ox* genes within specific domains could represent foci of high GA signaling (for those genes that are GA-responsive), or low bioactive GA concentrations (as a result of GA2ox activity). Obviously, future work defining the substrate specificity of the 16 apple GA2ox, as well as direct measurements of specific GA forms within the shoot apex, will be required to support further hypotheses.

The primary motivation for this study was to identify those *GA2ox* genes that are potentially most important for mediating GA levels in the shoot apical meristem. This is a necessary first step to fully understand how bioactive GAs can repress differentiation of the shoot apex into a floral fate in apple and other woody perennial plants. In the apex of non-fruiting spurs, which typically are undergoing floral initiation, both the direct counting of RNA-seq reads and QRT-PCR analysis suggested that expression was dominated by *MdGA2ox2A/2B*, with at least four additional *MdGA2ox* genes also appreciably expressed. Thus, *MdGA2ox2A/2B* may play a paramount role in governing GA levels in the shoot apex. Interestingly, in the analogous spurs in which developing fruit was present, the expression of several *MdGA2ox* genes was markedly increased, most notably *MdGA2ox1A/1B* and *MdGA2ox5A/5B*. This could reflect increased GA signaling in the

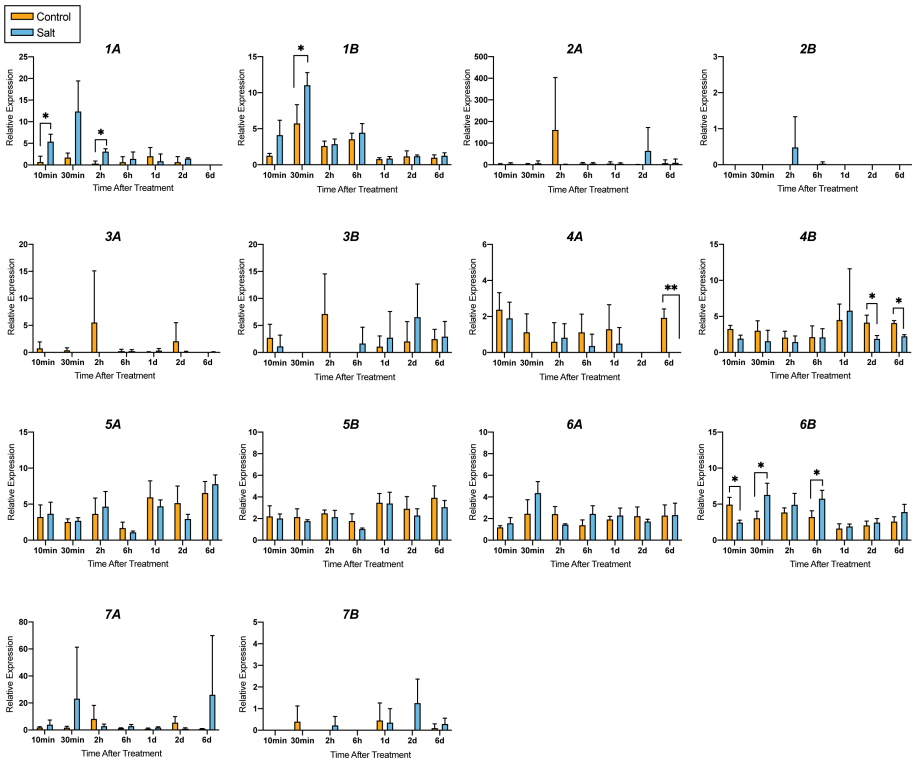


FIGURE 8
Relative expression of *MdGA2ox* genes in the first fully expanded leaves under salt treatment. Rapidly growing 'Gala' apple seedlings were either treated with 100 ml of water (in orange) or with 100 mL of 100 mM salt (NaCl) (in blue). Gene expression was detected at 10 min, 30 min, 2 h, 6 h, 1 d, 2 d, and 6 d after treatment. Asterisks represent statistical significance (*, $p < 0.05$; **, $p < 0.01$).

apex, although not all of these four genes were significantly induced by exogenous GA_3. Alternatively, some or all of these genes may be induced through a GA-independent pathway.

It is more intuitive to consider domains of *GA2ox* expression as barriers for the diffusion of GAs from source tissues. A model for this is the localized expression of *OsGA2ox1*, and *AtGA2ox2/AtGA2ox4* at the base of the meristem in the rice and Arabidopsis shoot apex, respectively, which is thought to preclude bioactive leaf-produced GAs from disrupting meristem organization (Sakamoto et al., 2001; Jasinski et al., 2005). In the same way, domains of expression of Class III *GA2ox* could set up a barrier for diffusion of C20 GAs. This idea is supported by the finding that GA_{12} is the major mobile GA signal over long distances in plant tissues (Regnault et al., 2015).

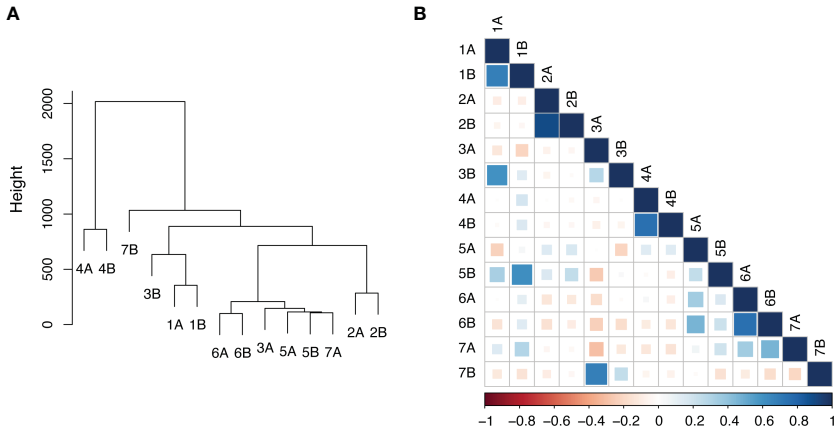


FIGURE 9
Cluster and correlation analysis of apple *GA2ox* genes. **(A)** The dendrogram depicts the relationship among the 14 *MdGA2ox* genes based on their expression. The y axis represents Euclidean distance ('Height'). **(B)** Expression correlation. The heat map represents the correlation values between *MdGA2ox* genes (dark blue, strongly positively correlated; dark red, strongly negatively correlated).

In this context it is notable that our data shows strong *GA2ox* gene expression in the fruit pedicel and leaf petiole. It has long been considered that GAs produced in the seeds of developing fruit diffuse to the shoot meristem to repress floral initiation. In contrast to this idea, we found that several *GA2ox* genes, especially *MdGA2ox4B*, are expressed in the fruit pedicel, suggesting that bioactive GAs produced in the seeds remain isolated. This idea is also strongly supported by the highly localized expression of *MdGA2ox1B -4A*, and *-4B* in the developing seed coat. If expression in the seed coat and/or pedicel is sufficient to metabolize all diffusing seed-produced bioactive GAs, then an established tenet of biennial bearing in apple must be re-examined. In the same way, the observed strong expression of *MdGA2ox6A/6B* in the bourse and spur leaves, including the petiole, may limit GA diffusion from leaves. Supporting this idea in the seedling, several *MdGA2ox* genes expressed in the petiole and petiole base of older leaves were relatively silent in the same structures of younger leaves, suggesting decreased control of GA diffusion. This would be anticipated if leaf-produced GAs were required to assist in promoting internode elongation in the young shoot.

The observation that nearly all of the *MdGA2ox* genes showed at least some indication of diurnal change in expression in rapidly growing seedlings was anticipated, given the pervasive diurnal regulation of genes in plants and the important role for GAs in growth. We hypothesized that bioactive GAs promoting growth of young tissues would be catabolized near or after the advent of darkness, and this would be reflected by increasing *GA2ox* gene expression at the end of the light period. For the leaf, where this should be most apparent, four genes fit this criterion: both *MdGA2ox1B* and *MdGA2ox6B* showed increasing expression throughout most of the light period, whereas *MdGA2ox5A/5B* showed strong upregulation at the start of the dark period. Notably, *GA2ox* expression in the young leaf was dominated by these same four genes. Other diurnal patterns are difficult to explain at the whole-structure level, but these could reflect unanticipated operations of GA signaling occurring in limited cell types.

One mechanism by which plants can respond to abiotic stress is the rapid depletion of bioactive GA levels, at least partially through deployment of GA2ox activity, which helps to limit further growth (reviewed in Colebrook et al., 2014). Consistent with previous findings in Arabidopsis (Achard et al., 2008; Magome et al., 2008) we found that subsets of the *MdGA2ox* genes were induced in the leaf blade under conditions of acute water deficit and salt exposure. In excised leaves allowed to air-dry, transcriptional responses were observed in the blade as early as 2 h after excision, and 7 of the 11 genes expressed in the leaf blade showed the expected stress induction. The striking induction of *MdGA2ox1A/1B* was surprising given the already strong expression of these two genes in the blade and petiole, respectively, and suggest that these two genes have an important role in limiting GA-promoted growth.

The unexpected finding that four of the genes were apparently induced when leaves were excised and placed in water is interesting but difficult to explain. Possibly it reflects a response to abrupt loss of negative pressure in the xylem, systemic response to wounding, response to touch (Lange and Lange, 2015) or disconnection from signaling molecules generated outside of the leaf, such as abscisic acid (ABA). These responses may be dysfunctional in the drying leaf due to the dehydration stress. The degree of this response (>100-200-fold induction) suggests that the underlying mechanism is biologically important and worth exploring.

The observation that only a subset of the genes responded significantly to salt, and that these responses were relatively subtle, was surprising, given that the salt concentration used was sufficient to arrest growth. For these genes, increase in expression was seen as early as 10-30 min after watering, suggesting that it reflects a response to decreased osmotic pressure caused by the salt water, rather than by ion toxicity. This is consistent with the observed rapid response of these three genes to leaf drying.

The number of *GA2ox* genes identified in the apple genome is greater than that identified for Arabidopsis (10; Rieu et al., 2008; Lange et al., 2020), maize (13; Li et al., 2021), tomato (11; Chen et al., 2016), grapevine (11; He et al., 2019), poplar (11; Gou et al., 2011), and peach (7; Cheng et al., 2021), and likely reflects not only the depth of our census but also gene family expansion. Neither direct RNA-seq counting for the genes nor the QRT-PCR analyses gave strong indication that individual homeologs had lost all function through silencing of expression. This homeology and retention of homeologs is consistent with the relatively recent whole-genome duplication in apple (Velasco et al., 2010). Retainment of duplicated *GA2ox* gene pairs after whole-genome duplication has been reported in many other flowering plant species and can reflect not only the evolutionary history of the species, but also the utility of the individual homeologs, which is augmented by diversification in function through expression.

Using correlation and clustering analysis based on a small snapshot of expression data, we found that expression of most of the homoeologous pairs was correlated, as expected, indicating potential functional redundancy in specific structures. However, some striking differences in expression patterns between individual homeologs were seen. In the spur, *MdGA2ox4A* but not *-4B* is strongly expressed in the leaf petiole and its base; whereas in the fruit pedicel and its base, *MdGA2ox4A* was essentially silent and *MdGA2ox4B* is strongly expressed. In addition, *MdGA2ox3A* is preferentially expressed in fruit, where *MdGA2ox3B* is silent, while *MdGA2ox3B* is preferentially expressed in the shoot apex, where *MdGA2ox3A* is silent. A final example in the spur is the preferential expression of *MdGA2ox7A* and *-7B* in the base of the bourse or spur leaf petiole, respectively, whereas their homeologs are preferentially expressed in other structures. In the seedling, *MdGA2ox2A* but not *-2B* is strongly expressed in the apex, whereas *MdGA2ox2B* is preferentially expressed in the older node, and *MdGA2ox1B* but not *-1A* is strongly expressed in leaf. Future work is needed to further define divergence in expression patterns among homeologs, and link this with sequence variation within regulatory regions of the genes that may govern this.

Our expression analyses did not include *MdGA2ox8A* and *-8B*, the homeologs of Arabidopsis *AtGA2ox9* and *AtGA2ox10*, because the Arabidopsis genes were recognized as *GA2ox* only after our study was initiated (Lange et al., 2020), and because the genes were found in a distinct clade from the known GA2ox proteins. In Arabidopsis, *AtGA2ox9* and *AtGA2ox10* are required for full freezing tolerance

and to limit seed number in the developing silique, respectively (Lange et al., 2020). As potential Class III GA2ox enzymes, restricted expression of *MdGA2ox8A* and *-8B* may contribute to limiting diffusion of GA12 and/or other C20 GAs.

This study provides a large amount of preliminary data with both fundamental and practical utility. The apparently simple genomic organization, comprising eight pairs of duplicated genes at homeologous positions, provides an accessible model to study gene expression diversification in a perennial woody plant species with recent whole-genome duplication. Apple seedlings can be grown quickly in controlled environment chambers and greenhouses, and apple seeds are easily obtained from fruit in markets. Many traits influenced by GAs have important production implications in tree fruit crops, including shoot elongation, plant stature, flowering, juvenility, and fruit size, and this work should greatly facilitate further studies to characterize mechanisms of GA signaling in these crops. Finally, the genes identified here are attractive candidates for genetic loci identified as influencing these traits, and the work should enable development of cultivars better suited to emerging production conditions, including those related to high-density management, mechanization, and resilience to climate change.

A limitation of this study is that expression of the genes was analyzed in bulk in entire structures (e.g., petiole, leaf, etc.), which are expected to comprise many distinct cell types, each of which may express specific *GA2ox* genes in a distinct pattern. Understanding the tissue-level expression patterns, and how these are distinct from or overlap with known routes of GA trafficking, is especially important. Future research should extend this genomic/transcriptional study through high-resolution approaches, including *in situ* RNA hybridization with gene-specific probes, and through single-cell RNA sequencing.

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

SZ and SVN designed the research. SZ, CG and SVN performed the research and analyzed the data; SZ drafted the manuscript. SVN composed the final manuscript copy. All authors contributed to the article and approved the submitted version.

Funding

This work was funded by USDA-NIFA-HATCH (project MICL02648) and MSU AgBioResearch (project GREEEN GR18-068).

Acknowledgments

We thank Dr. Patrick Edger and Dr. Courtney Hollender in the Department of Horticulture, and Dr. Jianping Hu in the Department of Plant Biology at Michigan State University for helpful discussions. We also thank our undergraduate assistant, Lauren Larson, for her help with sample collection and RNA extraction.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1117069/full#supplementary-material

SUPPLEMENTARY FILE 1
Protein sequences from Arabidopsis, rice, pea and bean cataloged as "Gibberellin 2-beta-dioxygenase" in the ExPASy Enzyme Database (fasta formatted file).

SUPPLEMENTARY FILE 2
Protein sequences in the 'GDDH13' genome annotation with high homology to the canonical GA2ox protein sequences cataloged in the ExPASy Enzyme Database as shown in Supplementary File 1 (fasta formatted file).

SUPPLEMENTARY FILE 3
Arabidopsis protein sequences with high homology to the GDDH13 sequences shown in Supplementary File 2 (fasta formatted file).

SUPPLEMENTARY FILE 4
HMMer protein structure analysis results.

SUPPLEMENTARY FILE 5
Sequences of transcript isoforms at each *MdGA2ox* locus derived from the 'Gala' shoot apex based on the 'GDDH13' genome annotation (fasta formatted file).

SUPPLEMENTARY FILE 6
Coding sequences of apple *GA2ox* gene models as annotated in the 'Gala' reference genome (fasta formatted file).

SUPPLEMENTARY FILE 7
TaqMan™ gene expression assays for apple *GA2ox* and *ACTIN* genes.

SUPPLEMENTARY FILE 8

Sequence alignment of the 16 GA2ox protein sequences from apple and other plants as shown in . Several sequences (MdGA2ox3A, -4B, -6A, -7A and -7B) were curated based on our transcriptome data and the 'Gala' genome. Conserved active site amino acids of the 2OG-FEII_Oxy domain catalytic core that interact with 2-oxoglutarate and Fe^{2+} are highlighted in the consensus sequence in yellow and red, respectively, and denoted under the sequence with an asterisk.

References

Achard, P., Renou, J.-P., Berthomé, R., Harberd, N. P., and Genschik, P. (2008). Plant DELLAs restrain growth and promote survival of adversity by reducing the levels of reactive oxygen species. *Curr. Biol.* 18, 656–660. doi: 10.1016/j.cub.2008.04.034

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Bertelsen, M. G., and Tustin, D. S. (2002). Suppression of flower bud formation in light cropping rrees of 'Pacific rose' apple using gibberellin sprays. *J. Hortic. Sci. Biotech.* 77, 753–757. doi: 10.1080/14620316.2002.11511568

Chen, S., Wang, X., Zhang, L., Lin, S., Liu, D., Wang, Q., et al. (2016). Identification and characterization of tomato gibberellin 2-oxidases (GA2oxs) and effects of fruit-specific SlGA2ox1 overexpression on fruit and seed growth and development. *Hortic. Res.* 3, 1–9. doi: 10.1038/hortres.2016.59

Cheng, J., Ma, J., Zheng, X., Lv, H., Zhang, M., Tan, B., et al. (2021). Functional analysis of the gibberellin 2-oxidase gene family in peach. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.619158

Colebrook, E. H., Thomas, S. G., Phillips, A. L., and Hedden, P. (2014). The role of gibberellin signalling in plant responses to abiotic stress. *J. Exp. Biol.* 217, 67–75. doi: 10.1242/jeb.089938

Daccord, N., Celton, J.-M., Linsmith, G., Becker, C., Choisne, N., Schijlen, E., et al. (2017). High-quality *de novo* assembly of the apple genome and methylome dynamics of early fruit development. *Nat. Genet.* 49, 1099–1106. doi: 10.1038/ng.3886

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

El-Gebali, S., Mistry, J., Bateman, A., Eddy, S. R., Luciani, A., Potter, S. C., et al. (2019). The pfam protein families database in 2019. *Nucleic Acids Res.* 47 (D1), D427–D432. doi: 10.1093/nar/gky995

Gasic, K., Hernandez, A., and Korban, S. S. (2004). RNA Extraction from different apple tissues rich in polyphenols and polysaccharides for cDNA library construction. *Plant Mol. Biol. Rep.* 22, 437–438. doi: 10.1007/BF02772687

Goldberg-Moeller, R., Shalom, L., Shlizerman, L., Samuels, S., Zur, N., Ophir, R., et al. (2013). Effects of gibberellin treatment during flowering induction period on global gene expression and the transcription of flowering-control genes in citrus buds. *Plant Sci.* 198, 46–57. doi: 10.1016/j.plantsci.2012.09.012

Gottschalk, C., Zhang, S., Schwallier, P., Rogers, R., Bukovac, M. J., and van Nocker, S. (2021). Genetic mechanisms associated with floral initiation and the repressive effect of fruit on flowering in apple (Malus X domestica borkh). *PloS One* 16, e0245487. doi: 10.1371/journal.pone.0245487

Gou, J., Ma, C., Kadmiel, M., Gai, Y., Strauss, S., Jiang, X., et al. (2011). Tissue-specific expression of populus C19 GA2-oxidases differentially regulate above- and below-ground biomass growth through control of bioactive GA concentrations. *New Phytol.* 192, 626–639. doi: 10.1111/j.1469-8137.2011.03837.x

He, H., Liang, G., Lu, S., Wang, P., Liu, T., Ma, Z., et al. (2019). Genome-wide identification and expression analysis of GA2ox, GA3ox, and GA20ox are related to gibberellin oxidase genes in grape (Vitis vinifera l). *Genes* 10, 680. doi: 10.3390/genes10090680

Hedden, P. (2020). The current status of research on gibberellin biosynthesis. *Plant Cell Physiol.* 61, 1832–1849. doi: 10.1093/pcp/pcaa092

Hernández-García, J., Briones-Moreno, A., and Blázquez, M. A. (2021). Origin and evolution of gibberellin signaling and metabolism in plants. *Semin. Cell Dev. Biol.* 109, 46–54. doi: 10.1016/j.semcdb.2020.04.009

Huang, Y., Wang, X., Ge, S., and Rao, G. Y. (2015). Divergence and adaptive evolution of the gibberellin oxidase genes in plants. *BMC Evol. Biol.* 15, 207. doi: 10.1016/j.semcdb.2020.04.009

Jasinski, S., Piazza, P., Craft, J., Hay, A., Woolley, L., Rieu, I., et al. (2005). KNOX action in arabidopsis is mediated by coordinate regulation of cytokinin and gibberellin activities. *Curr. Biol.* 15, 1560–1565. doi: 10.1016/j.cub.2005.07.023

Johnson, L. S., Eddy, S. R., and Portugaly, E. (2010). Hidden Markov model speed heuristic and iterative HMM search procedure. *BMC Bioinf.* 11, 431. doi: 10.1186/1471-2105-11-431

Lange, T., Krämer, C., and Lange, M. J. P. (2020). The class III gibberellin 2-oxidases AtGA2ox9 and AtGA2ox10 contribute to cold stress tolerance and fertility. *Plant Physiol.* 184, 478–486. doi: 10.1104/pp.20.00594

Lange, M. J. P., and Lange, T. (2015). Touch-induced changes in arabidopsis morphology dependent on gibberellin breakdown. *Nat. Plants* 1, 14025. doi: 10.1038/nplants.2014.25

Lee, D. J., and Zeevaart, J. A. D. (2005). Molecular cloning of GA 2-oxidase3 from spinach and its ectopic expression in nicotiana sylvestris. *Plant Physiol.* 138, 243–254. doi: 10.1104/pp.104.056499

Letunic, I., and Bork, P. (2018). 20 years of the SMART protein domain annotation resource. *Nucleic Acids Res.* 46 (D1), D493–D496. doi: 10.1093/nar/gkx922

Li, Y., Shan, X., Jiang, Z., Zhao, L., and Jin, F. (2021). Genome-wide identification and expression analysis of the GA2ox gene family in maize (*Zea mays* l .) under various abiotic stress conditions. *Plant Physiol. Biochem.* 166, 621–633. doi: 10.1016/j.plaphy.2021.06.043

Luckwill, L. C., Weaver, P., and Macmillan, J. (1969). Gibberellins and other growth hormones in apple seeds. *J. Hortic. Sci.* 44, 413–424. doi: 10.1080/00221589.1969.11514325

Magome, H., Yamaguchi, S., Hanada, A., Kamiya, Y., and Oda, K. (2008). The DDF1 transcriptional activator upregulates expression of a gibberellin-deactivating gene, GA2ox7, under high-salinity stress in arabidopsis. *Plant J.* 56, 613–626. doi: 10.1111/j.1365-313X.2008.03627.x

Regnault, T., Daviere, J.-M., Wild, M., Sakvarelidze-Ahard, L., Heintz, D., Carrera Bergua, E., et al. (2015). The gibberellin precursor GA_{12} acts as a long-distance growth signal in arabidopsis. *Nat. Plants* 1, 15073. doi: 10.1038/nplants.2015.73

Rieu, I., Eriksson, S., Powers, S. J., Gong, F., Griffiths, J., Woolley, L., et al. (2008). Genetic analysis reveals that C19-GA 2-oxidation is a major gibberellin inactivation pathway in arabidopsis. *Plant Cell* 20, 2420–2436. doi: 10.1105/tpc.108.058818

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

Sakamoto, T., Kobayashi, M., Itoh, H., Tagiri, A., Kayano, T., Tanaka, H., et al. (2001). Expression of a gibberellin 2-oxidase gene around the shoot apex is related to phase transition in rice. *Plant Physiol.* 125, 1508–1516. doi: 10.1104/pp.125.3.1508

Schneider, G., and Schliemann, W. (1994). Gibberellin conjugates: An overview. *Plant Growth Regul.* 15, 247–260. doi: 10.1007/BF00029898

Schomburg, F. M., Bizzell, C. M., Lee, D. J., Zeevaart, J. A. D., and Amasino, R. M. (2003). Overexpression of a novel class of gibberellin 2-oxidases decreases gibberellin levels and creates dwarf plants. *Plant Cell* 15, 151–163. doi: 10.1105/tpc.005975

Serrani, J. C., Sanjuán, R., Ruiz-Rivero, O., Fos, M., and García-Martínez, J. L. (2007). Gibberellin regulation of fruit set and growth in tomato. *Plant Physiol.* 145, 246–257. doi: 10.1104/pp.107.098335

Southwick, S. M., Weis, K. G., and Yeager, Y. T. (1995). Controlling cropping in 'Loadel' cling peach using gibberellin: Effects on flower density, fruit distribution, fruit firmness, fruit thinning, and yield. *J. Am. Soc. Hortic. Sci.* 120, 1087–1095. doi: 10.21273/JASHS.120.6.1087

Stothard, P. (2000). The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *Biotechniques* 28, 1102–1104. doi: 10.2144/00286ir01

Sun, X., Jiao, C., Schwaninger, H., Chao, C. T., Ma, Y., Duan, N., et al. (2020). Phased diploid genome assemblies and pan-genomes provide insights into the genetic history of apple domestication. *Nat. Genet.* 52, 1423–1432. doi: 10.1038/s41588-020-00723-9

Thomas, S. G., Phillips, A. L., and Hedden, P. (1999). Molecular cloning and functional expression of gibberellin 2-oxidases, multifunctional enzymes involved in gibberellin deactivation. *P Natl. Acad. Sci. U.S.A.* 96, 4698–4703. doi: 10.1073/pnas.96.8.469

Tromp, J. (1982). Flower-bud formation in apple as affected by various gibberellins. *J. Hortic. Sci.* 57, 277–282. doi: 10.1080/00221589.1982.11515053

Ubeda-Tomás, S., García-Martínez, J. L., and López-Díaz, I. (2006). Molecular, biochemical and physiological characterization of gibberellin biosynthesis and catabolism genes from nerium oleander. *J. Plant Growth Regul.* 25, 52–68. doi: 10.1007/s00344-005-0087-x

Varbanova, M., Yamaguchi, S., Yang, Y., McKelvey, K., Hanada, A., Borochov, R., et al. (2007). Methylation of gibberellins by arabidopsis GAMT1 and GAMT2. *Plant Cell* 19, 32–45. doi: 10.1105/tpc.106.044602

Velasco, R., Zharkikh, A., Affourtit, J., Dhingra, A., Cestaro, A., Kalyanaraman, A., et al. (2010). The genome of the domesticated apple (*Malus* × *domestica* borkh.). *Nat. Genet.* 42, 833–839. doi: 10.1038/ng.654

Yoshida, H., Takehara, S., Mori, M., Ordonio, R. L., and Matsuoka, M. (2020). Evolution of GA metabolic enzymes in land plants. *Plant Cell Physiol.* 61, 1919–1934. doi: 10.1038/ng.654

Zhang, S., Gottschalk, C., and van Nocker, S. (2019). Genetic mechanisms in the repression of flowering by gibberellins in apple (*Malus X domestica* borkh.). *BMC Genomics* 20, 747. doi: 10.1186/s12864-019-6090-6

Zhang, S., Zhang, D., Fan, S., Du, L., Shen, Y., Xing, L., et al. (2016). Effect of exogenous GA3 and its inhibitor paclobutrazol on floral formation, endogenous hormones, and flowering-associated genes in 'Fuji' apple (Malus domestica borkh.). *Plant Physiol. Biochem.* 107, 178–186. doi: 10.1016/j.plaphy.2016.06.005

Zhu, Y., Nomura, T., Xu, Y., Zhang, Y., Peng, Y., Mao, B., et al. (2006). *ELONGATED UPPERMOST INTERNODE* encodes a cytochrome P450 monooxygenase that epoxidizes gibberellins in a novel deactivation reaction in rice. *Plant Cell* 18, 442–456. doi: 10.1105/tpc.105.038455

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 31 May 2023
DOI 10.3389/fpls.2023.1198930

OPEN ACCESS

EDITED BY
Chenxia Cheng,
Qingdao Agricultural University, China

REVIEWED BY
Songtao Jiu,
Shanghai Jiao Tong University, China
Deguo Han,
Northeast Agricultural University, China

*CORRESPONDENCE
Xiaorong Wang
wangxr@sicau.edu.cn

†These authors have contributed equally to this work

RECEIVED 02 April 2023
ACCEPTED 08 May 2023
PUBLISHED 31 May 2023

CITATION
He W, Luo L, Xie R, Chai J, Wang H, Wang Y, Chen Q, Wu Z, Yang S, Li M, Lin Y, Zhang Y, Luo Y, Zhang Y, Tang H and Wang X (2023) Transcriptome sequencing analyses uncover mechanisms of citrus rootstock seedlings under waterlogging stress.
Front. Plant Sci. 14:1198930.
doi: 10.3389/fpls.2023.1198930

Transcriptome sequencing analyses uncover mechanisms of citrus rootstock seedlings under waterlogging stress

Wen He†, Liang Luo†, Rui Xie, Jiufeng Chai, Hao Wang, Yan Wang, Qing Chen, Zhiwei Wu, Shaofeng Yang, Mengyao Li, Yuanxiu Lin, Yunting Zhang, Ya Luo, Yong Zhang, Haoru Tang and Xiaorong Wang*

College of Horticulture, Sichuan Agricultural University, Chengdu, Sichuan, China

Citrus plants are sensitive to waterlogging, which can cause yield reduction. Their production heavily depends on the rootstock being used for grafting of scion cultivars, and the rootstock is the first organ to be affected by waterlogging stress. However, the underlying molecular mechanisms of waterlogging stress tolerance remain elusive. In this study we investigated the stress response of two waterlogging-tolerant citrus varieties (*Citrus junos* Sieb ex Tanaka cv. Pujiang Xiangcheng and Ziyang Xiangcheng), and one waterlogging-sensitive variety (red tangerine) at the morphological, physiological, and genetic levels in leaf and root tissues of partially submerged plants. The results showed that waterlogging stress significantly decreased the SPAD value and root length but did not obviously affect the stem length and new root numbers. The malondialdehyde (MDA) content and the enzyme activities of superoxide dismutase (SOD), guaiacol peroxidase (POD), and catalase (CAT) were enhanced in the roots. The RNA-seq analysis revealed that the differentially expressed genes (DEGs) were mainly linked to 'cutin, suberine, and wax biosynthesis', 'diterpenoid biosynthesis', and 'glycerophospholipid metabolism' in the leaves, whereas were linked to 'flavonoid biosynthesis', 'biosynthesis of secondary metabolites and metabolic pathways' in the roots. Finally, we developed a working model based on our results to elucidate the molecular basis of waterlogging-responsive in citrus. Therefore, our data obtained in this study provided valuable genetic resources that will facilitate the breeding of citrus varieties with improved waterlogging tolerance.

KEYWORDS

flooding, transcriptomic, citrus rootstock, *Citrus junos* Sieb ex Tanaka cv. Pujiang Xiangcheng, differentially expressed genes

Introduction

Citrus is one of the most economically and socially important evergreen fruit tree species globally. Waterlogging is a serious impediment to citrus production because it has significantly affected their yield, quality, as well as geographical distribution (García-Sánchez et al., 2007; Sampaio et al., 2021). Exposure to waterlogging may damage the root and could decrease the chlorophyll and cause leaf senescence (Pan et al., 2021; Ren et al., 2023; Tahjib-Ul-Arif et al., 2023). Hypoxia brought by the stress is becoming one of the crucial stress factors negatively affecting citrus production (Xie et al., 2021). Citrus is cultivated by grafting, with the rootstock being the organ that initially senses and responds to low oxygen conditions, thereby being seriously damaged by soil hypoxia (Ghaffari et al., 2021; Xie et al., 2021). Although changes in protein, gene expression, and metabolite levels have been studied under hypoxic stress, scarcely attention has been paid for the molecular responses to waterlogging stress in citrus.

Plants can temporarily maintain energy production during hypoxia caused by waterlogging (Gong et al., 2020; Pan et al., 2021). However, the anoxic metabolism in root forms toxic substances, including organic acids, excessive ethanol, and aldehydes, along with increased formation of reactive oxygen species (ROS) (Cheng et al., 2016). These inhibit the root function and plant growth, thus eventually causing cell death and leaf senescence (Teoh et al., 2022). During this complex biological process, numerous genes are activated, which are important for plant survival (Voesenek and Bailey Serres, 2015). For instance, transcription factors (TFs) are thought to control many waterlogging-responsive genes by directly binding to the appropriate stress-responsive cis-acting elements in the promoter regions (Christianson et al., 2010). The ERF transcription factors directly regulate waterlogging stress response, e.g., *PhERF2* expression was found to be upregulated in petunia (Yin et al., 2019), while *TaERFVII.1* expression alleviated the negative effects of waterlogging stress in wheat (Wei et al., 2019). Additionally, transcriptomic analysis of roots of kiwifruit (Zhang et al., 2015), citrus (Xie et al., 2021), and wheat (Shen et al., 2020) showed that numerous genes involved in photosynthesis, hormone production and signaling pathways, and ROS generation or scavenging etc. are expressed in response to waterlogging stress.

Plants can develop multiple adaptation strategies against waterlogging stress, including adventitious root formation, petiole elongation, and secondary aerenchyma development (Hattori et al., 2009; Bailey-Serres et al., 2012; Tahjib-Ul-Arif et al., 2023). Gene expression and signal transduction are also critical for plant survival under waterlogging stress (Liu et al., 2021). Currently, studies are being conducted on how citrus rootstocks respond to flooding stress, such as Xie et al. (2021) studied the transcript levels of rootstock and identified 232 hypoxia-responsive genes. However, the difference in transcript levels between the aboveground and belowground parts in citrus rootstock remain unclear. In this study, we performed a transcriptome profiling, and validated the morphological, physiological and biochemical to investigated the stress response of two waterlogging-tolerant citrus varieties, i.e., a new citrus rootstock called *Citrus junos* Sieb ex Tanaka cv. Pujiang Xiangcheng (Fu et al., 2017) and *C. junos* Ziyang Xiangcheng, along with one waterlogging-sensitive variety red tangerine to waterlogging stress. The objectives of this paper were to: (1) explore how waterlogging affects the plant growth of citrus rootstock cultivars with different waterlogging tolerances, and (2) investigate the underlying physiological and molecular mechanisms of the differences in waterlogging tolerance between the three citrus rootstock cultivars.

Material and methods

Plant materials and treatments

In this study, the cultivars named *C. junos* Sieb. Ex Tanaka cv. 'Pujiang Xiangcheng' (abbreviated Pj), *C. junos* 'Ziyang Xiangcheng' (abbreviated Zy) and red tangerine (*Citrus reticulata* Blanco, abbreviated Rt) were grown in an experimental orchard at Sichuan Citrus Germplasm Repository in Chengdu, China were used as materials. The mature fruits were harvested from the same tree, and the seeds were collected. Isolated seeds were surface sterilized using 0.5 M NaOH. The sterilized seedlings were then selected based on a uniform size and subsequently grown in wetted soil-containing pots in a growth chamber. All seedlings were cultured for approximately six months with normal watering and fertilizing regimen.

Three biological replicates (32 seedlings per replicate) were randomly selected for each treatment. Pots were then filled with water until the water level was ~1 cm above the soil level. Water was maintained at this level for the entire duration of the treatment. For control treatments, plants were watered every three days. On day 15 and 35 of the experiment, the seedlings with at least three fully expanded leaves were taken, and their SPAD value, shoot height, root length, and root numbers were measured.

Determination of the physiological and biochemical indexes

The activities of antioxidant enzymes, including superoxide dismutase (SOD, EC 1.15.1.1), guaiacol peroxidase (POD, EC 1.11.1.7) and catalase (CAT, EC 1.11.1.6) were determined according to Shahid et al. (2019). Soluble sugar and protein were analyzed from leaves according to the previously described method (Zaher-Ara et al., 2016). The hydrogen peroxide (H_2O_2) and malondialdehyde (MDA) contents were determined by a spectrophotofluorometric method (Muñoz et al., 2008).

RNA-seq, data processing, and expression analysis

Based on the research conducted by Xie et al. (2021), and daily field observations, all three rootstocks showed symptoms of waterlogging stress after processing treatment for 35 days. Then, the leaves and roots of Pj, Zy, and Rt were collected and stored at

-80 °C. Small and equal amount of leaves from five plants were pooled together as one replicate. Three independent replicates were subjected to RNA extraction as described previously (He et al., 2018; He et al., 2022). Clean reads were submitted to the China National GeneBank DataBase (CNGBdb) database sequence read archive, under the project: CNP0004172. Before sequence assembly, the adapter sequences and low-quality reads were removed from the raw data (Supplementary Table S1). TopHat2 (Kim et al., 2013) was used to map the clean reads to the *C. junos* reference genome (unpublished). The number of fragments per kilobase of transcript per million mapped reads (FPKM) was calculated using the RSEM tool (Li and Dewey, 2011). The average FPKM values of the three replicates were taken as the expression level of genes in each sample. The sets of DEGs (differentially expressed genes) were identified using the eBays function in the limma package with |log2 (foldchange)| ≥ 1.0 and the adjusted *p*-value < 0.05.

A Weighted Gene Co-expression Network Analysis (WGCNA) was applied to evaluate the gene expression (Langfelder and Horvath, 2008). The flashClust toolkit (R language) was used to perform cluster analysis on samples and set appropriate thresholds (Müllner, 2013). Heatmaps were created with the heatmap2 function in the R environment.

qRT-PCR analysis

To verify the authenticity of the transcriptional data, 13 key TFs identified in the leaf or root were selected for further qRT-PCR validation. Total RNA was extracted and purified using the EasyPure® Plant RNA Kit (TransGen Biotech Co., Ltd., Beijing, China), as per the manufacturer's protocol. The qRT-PCR mixture included 1 µL cDNA template, 5 µL of 2×*TransStart*® Green qPCR SuperMix (TransGen Biotech Co., Ltd., Beijing, China), and 0.5 µL each of reverse and forward primer. The PCR program was as described previously (He et al., 2018). The gene-specific primers used are shown in Supplementary Table S2. Three biological and three technical replicates were adopted. The $2^{-\triangle\triangle Ct}$ method was utilized to calculate the relative gene expression, with the *EF-1α* gene being used as the internal control (He et al., 2021).

Statistical analysis

The collected data was prepared with Microsoft Excel. Significant differences between grafted combinations were analyzed by Tukey's method, while Pearson correlation analysis and principal component analysis (PCA) were performed using SPSS 20.0 software (IBM, USA). Figures were drawn in the GraphPad Prism software (v. 7.04).

Results

Effects of waterlogging stress on phenotypic and physiological index

As shown in Figure 1A, waterlogging stress-induced root injury was observed in all the citrus rootstocks at 35 Days After Treatment (DAT). Additionally, the SPAD value of leaves, stem length, new root numbers, lateral root numbers, primary root length, and lateral root length were quantified post waterlogging stress or under control conditions (Figure 1B). The results indicated that Pj exhibited significant difference in its SPAD value, stem length, new root numbers, and lateral root length as compared with Zy and Rt (Figure 1B). Based on those data, we demonstrated that Pj was highly tolerant to waterlogging stress as compared with Zy and Rt.

Effects of waterlogging stress on physiological and biochemical parameters

Waterlogging stress significantly affected the MDA content and SOD activity in the roots of Zy and Rt (Figure 2A). The treatment also induced noticeable changes in the POD activity in the roots of Pj and Rt, and the CAT activity in the roots of Pj and Zy, respectively (Figure 2A). However, the pattern of the effect of waterlogging stress on the soluble sugar and soluble protein contents in the roots of all rootstocks remained consistent with the control (Figure 2A).

Waterlogging stress significantly affected the accumulation of soluble sugar and soluble protein contents in the leaves of Pj (Figure 2B). Furthermore, the MDA contents of both Pj and Zy at 15 DAT and 35 DAT were higher than those in the control (Figure 2B). Although there were changes in physiological parameters in leaves of Rt, none were significant (Figure 2B).

Analysis of differentially expressed genes

To further understand the underlying molecular mechanisms of waterlogging tolerance, we performed RNA sequencing of the root and leaf to identify the waterlogging stress-responsive DEGs (Supplementary Table S1). At 35 DAT, there were 624, 882, and 1058 upregulated DEGs and 984, 960, and 986 downregulated DEGs in the leaves of Pj, Zy and Rt, respectively. Additionally, there were 629, 1358, and 1506 upregulated DEGs, and 971, 800, and 894 downregulated DEGs in the roots of Pj, Zy and Rt, respectively (Supplementary Table S3). After performing KEGG (Kyoto Encyclopedia of Genes and Genomes) analysis on these DEGs, we found that the enriched pathways were mainly photosynthesis-related in the leaves and secondary metabolite biosynthesis-related in the roots (Supplementary Figure S1).

To understand the difference in gene expression between *C. junos* and *C. reticulata* under waterlogging stress, we identified the DEGs and subjected them to KEGG enrichment analysis, and found 154 DEGs and 147 DEGs in the leaves and roots, respectively (Figures 3A, B). The KEGG analysis results showed that most DEGs in the leaves are those participated in the process of 'cutin, suberine, and wax biosynthesis', 'diterpenoid biosynthesis', 'glycerophospholipid metabolism' and 'plant hormone signal transduction' ($p < 0.05$, Figure 3C). Furthermore, the KEGG analysis results showed that most genes in the root participated in the 'flavonoid biosynthesis', 'biosynthesis of secondary metabolites', 'metabolic pathways',

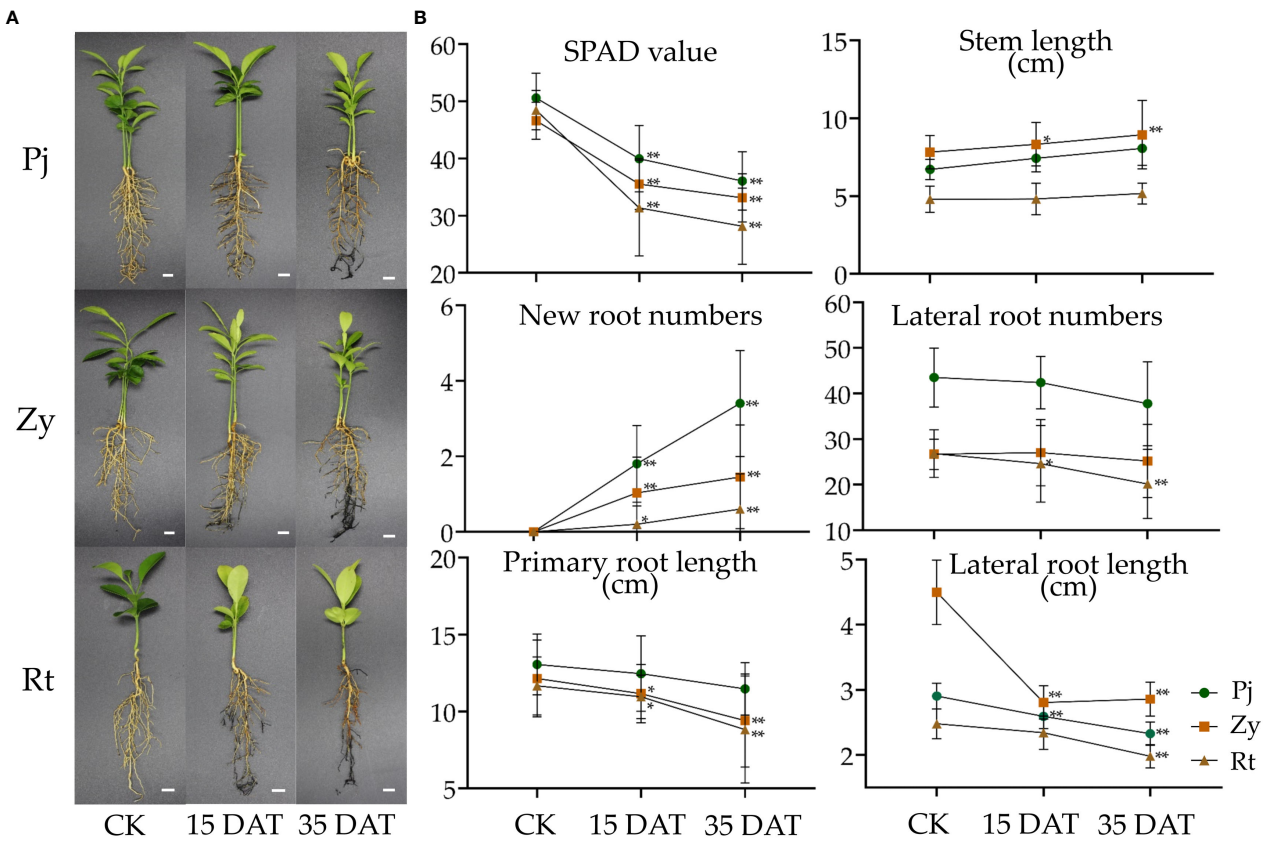


FIGURE 1
Phenotypic and physiological analysis of citrus rootstocks in response to waterlogging stress. **(A)** The phenotype of three citrus rootstocks under waterlogging stress. **(B)** The physiological analysis of citrus rootstocks in response to waterlogging stress. The bars represent the standard errors of the means ($n \geq 3$). The asterisks indicate that the values are significantly different (* for $p < 0.05$ and ** for $p < 0.01$). Bars = 1 cm. Pj, Pujiang Xiangcheng (*Citrus junos* [Sieb.] Tanaka); Zy, Ziyang Xiangcheng (*C. junos* [Sieb.] Tanaka); Rt, Red tangerine (*Citrus reticulata* Blanco); DAT, Days after Treatment.

'brassinosteroid biosynthesis', 'phenylpropanoid biosynthesis', 'stilbenoid, diarylheptanoid and gingerol biosynthesis', 'steroid biosynthesis', 'circadian rhythm – plant', and 'carbon fixation in photosynthetic organisms' ($p < 0.05$, Figure 3D).

DEGs involved in ROS generation/scavenging and hormonal pathways

Typically, waterlogging causes ROS accumulation, which is mainly attributed to the coordinated expression of ROS generation/scavenging-related genes. In this study, epoxide hydrolase (gene_07329) genes were significantly upregulated by waterlogging stress in Pj and Zy, as compared with Rt (Figure 4). Among those DEGs in the leaves, five genes were enriched in 'plant hormone signal transduction' pathway, including one auxin/indole-3-acetic Acid (Aux/IAA, gene_01190), three cytokinine related genes (one histidine-containing phosphotransfer protein, gene_21726; two two-component response regulators, gene_05220 and gene_24622) and one ABA-responsive element binding factor (ABF, gene_14446) (Figure 4). However, there were no ROS-related DEGs in the root. But two hormone signal transduction related DEGs, including ABSCISIC ACID-INSENSITIVE 5-like protein (ABI, gene_30411) and Aux/IAA (gene_19419), were identified in the roots (Figure 4). Therefore, ABF was upregulated, while Aux/IAA was downregulated in all the rootstocks post waterlogging treatment for 35 days.

Differentially expressed transcription factors

In the shoot, six TFs, including an ethylene response factor (AP2/ERF, gene_00508), a *bHLH* (basic helix-loop-helix, gene_18799), a *HSF* (Heat Stress Transcription Factor, gene_30482) and three unkown TFs (gene_18195, gene_16249 and gene_30857) were differentially expressed between *C. junos* and *C. reticulata* (Supplementary Table S4). In the root, 11 differentially expressed TFs between genotypes were identified, including two HSFs, one *EREBP* (ethylene-responsive element binding protein), three WRKY transcription factor and five unkonw TFs (gene_04814, gene_25615, gene_11186, gene_16249 and gene_09850) (Supplementary Table S5). Overall, 16 of 17 TFs were upregulated in treated sample compared with controls

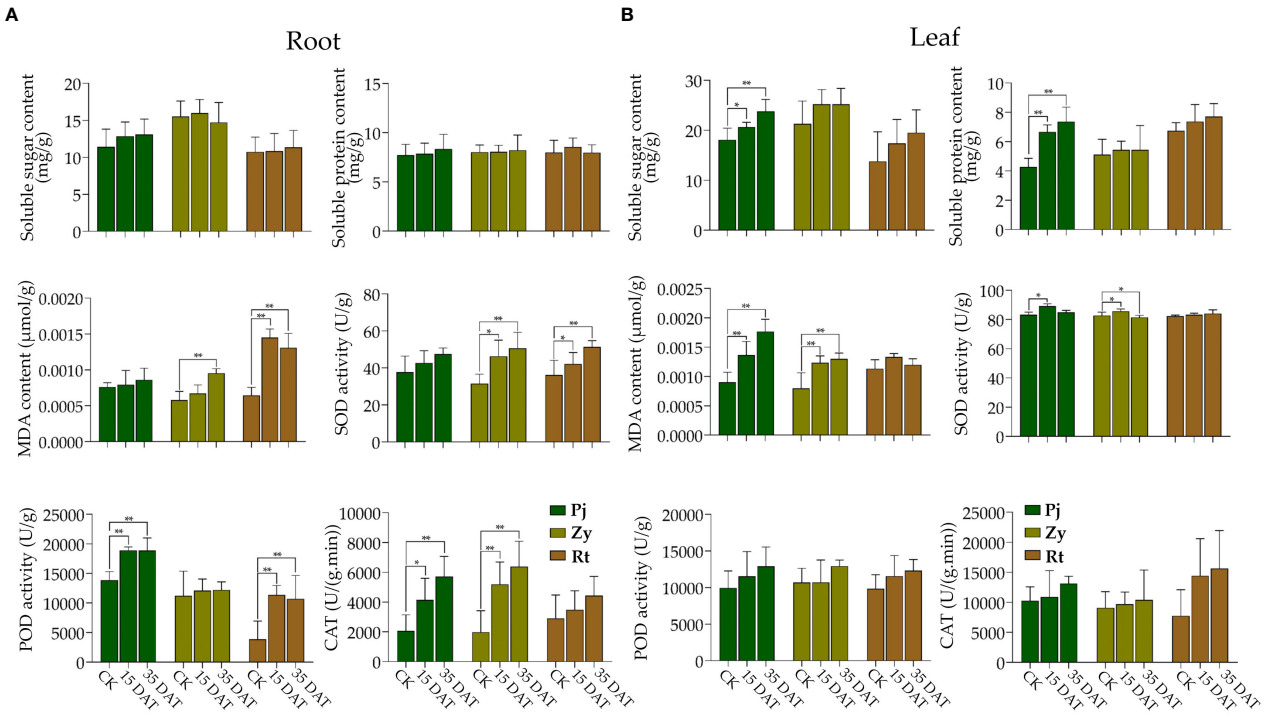


FIGURE 2
Effects on physiological and biochemical parameters of citrus rootstocks under waterlogging stress. **(A, B)** indicate the physiological and biochemical parameters of the root and leaf, respectively. The asterisks indicate that the values are significantly different (** for $p < 0.01$ and * for $p < 0.05$).

(Figure 5A). It is worthy to note that *EREBP* (gene_02956) only expressed in the roots of Pj and Zy, and gene_30857 only expressed in the leaves of Zy after treatment (Figure 5A). All the differentially expressed TFs were tissue-specific, except gene_16249 and gene_30482 (Figures 5B, C). In details, gene_16249 and gene_30482 were upregulated in three rootstocks (Figure 5C).

Weighted gene co-expression network analysis

To identify the co-expression modules correlated to waterlogging tolerance and the hub genes involved in their transcriptional regulatory networks, we carried out a weighted gene co-expression network analysis (WGCNA). Our results revealed 30 and 37 co-expressed modules in the leaf and the root, after gathering from 28 to 4168 and from 44 to 2132 genes, respectively (Figures 6A, B and Supplementary Figure S2). Additionally, we also analyzed the correlation between these modules with seven metabolites (soluble sugar content, soluble protein content, MDA content, SOD activity, POD activity, CAT and H_2O_2 content). The results showed that the 'salmon' module was significantly positively correlated with the H_2O_2 content in the leaf (r = 0.83), whereas the 'skyblue' module was significantly positively correlated with the POD activity in the root (r = 0.79). Therefore, this result suggested that the 441 and 206 genes in the 'salmon' and 'skyblue' modules, respectively, may play important roles in the waterlogging tolerance of citrus (Supplementary Tables S6, S7). Furthermore, we utilized the KME (eigengene connectivity) value to determine the hub genes in the 'salmon' module in the leaf and the 'skyblue' module in the root. This analysis allowed us to identify gene_23569 (NADH dehydrogenase (ubiquinone) 1 alpha subcomplex subunit 5) as a hub gene within the 'salmon' module, and gene_13979 (function is currently unknown) as a hub gene within the 'skyblue' module. These hub genes may play important roles in regulating the expression patterns of other genes within their respective modules.

We also performed KEGG enrichment analysis for genes in the 'salmon' and 'skyblue' modules. The results showed that genes in the 'salmon' module participated in the 'proteasome', 'oxidative phosphorylation' and 'carbon fixation in photosynthetic organisms' pathways (top three pathways of KEGG enrichment) (Figure 6C). Furthermore, the genes in the 'skyblue' module participated in the 'phenylpropanoid biosynthesis', 'tropane, piperidine, and pyridine alkaloid biosynthesis' and 'circadian rhythm – plant' pathways (top three pathways of KEGG enrichment) (Figure 6D).

Validation of candidate genes by qRT-PCR analysis

To validate the accuracy of the RNA-Seq expression patterns, we have chosen 13 key candidate TFs for qRT-PCR validation. The results showed that the qRT-PCR expression levels were generally consistent with RNA-Seq data, with a good positive correlation (R^2 = 91.92) (Supplementary Figure S3), thereby confirming the transcriptome data reliability in the present study. Among those genes, *EREBP* (gene_02956) was only expressed in the roots of Pj

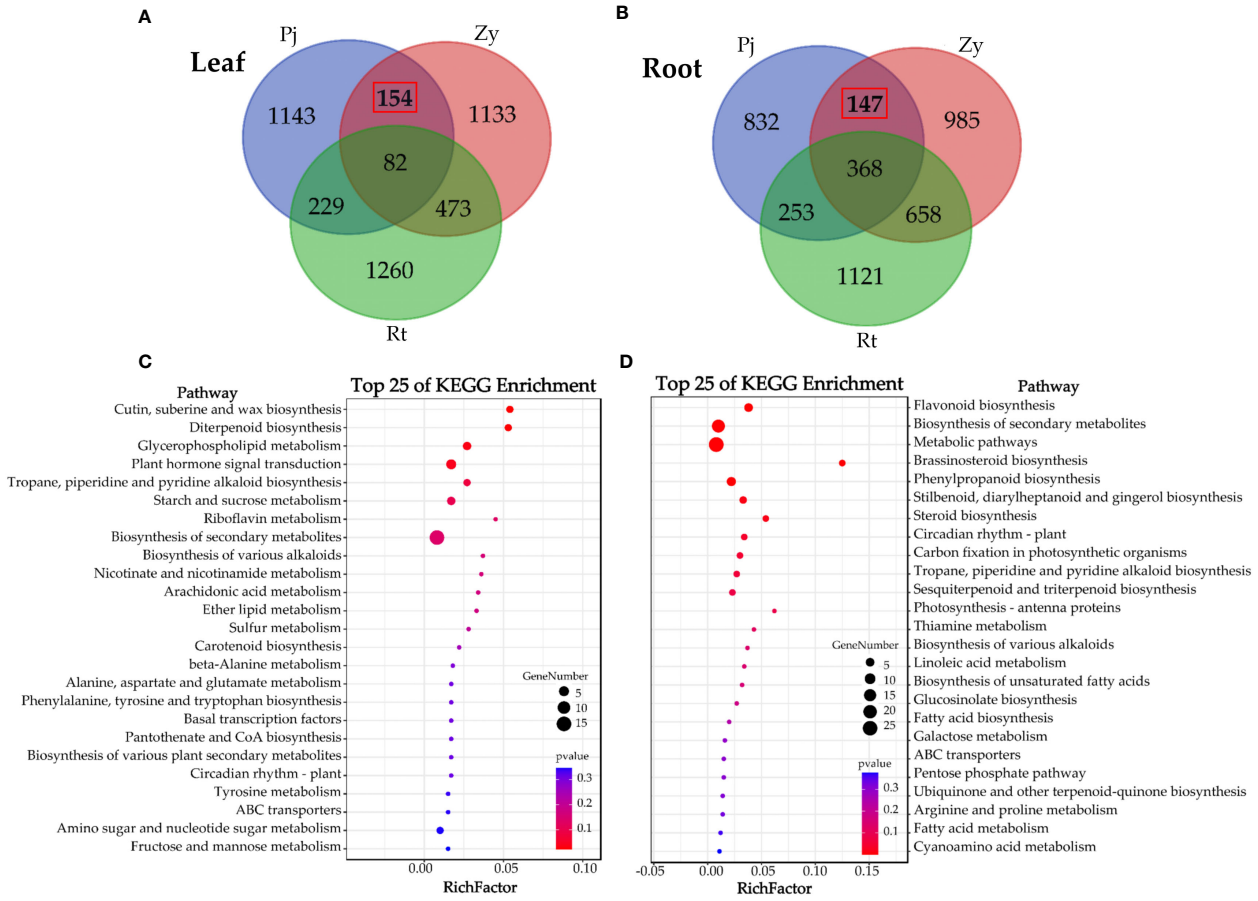


FIGURE 3
Analysis of the DEGs in the leaves and root of three rootstocks. **(A, C)** show the Venn diagram analyses of DEGs in leaves and KEGG enrichment analysis of 154 DEGs, respectively. **(B, D)** shows the Venn diagram analyses of DEGs in root and KEGG enrichment analysis of 147 DEGs, respectively.

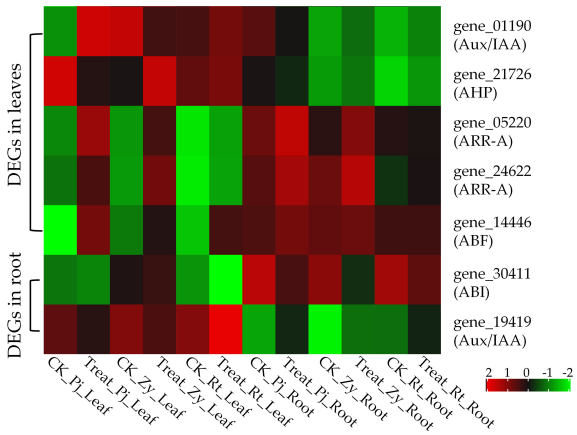


FIGURE 4
Heatmap of DEGs involved in ROS generation/scavenging and hormones. Heat map showing expression of the DEGs based on $\log_{10}^{(FPKM+0.01)}$ values, the same as below.

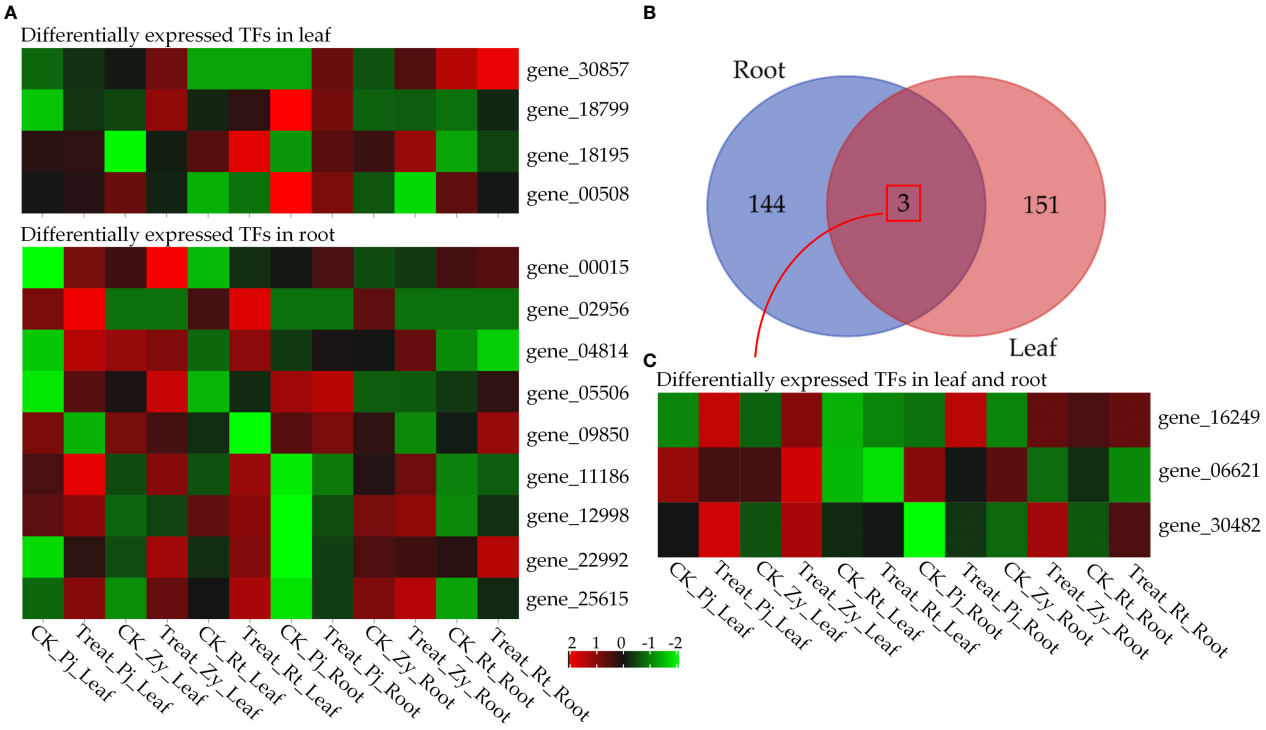


FIGURE 5
Differentially expressed transcription factors. **(A)** Heatmap of tissue-specific transcription factors. **(B)** Venn diagram analyses of DEGs in leaf and root. **(C)** Heatmap of differentially expressed transcription factors in both leaf and root.

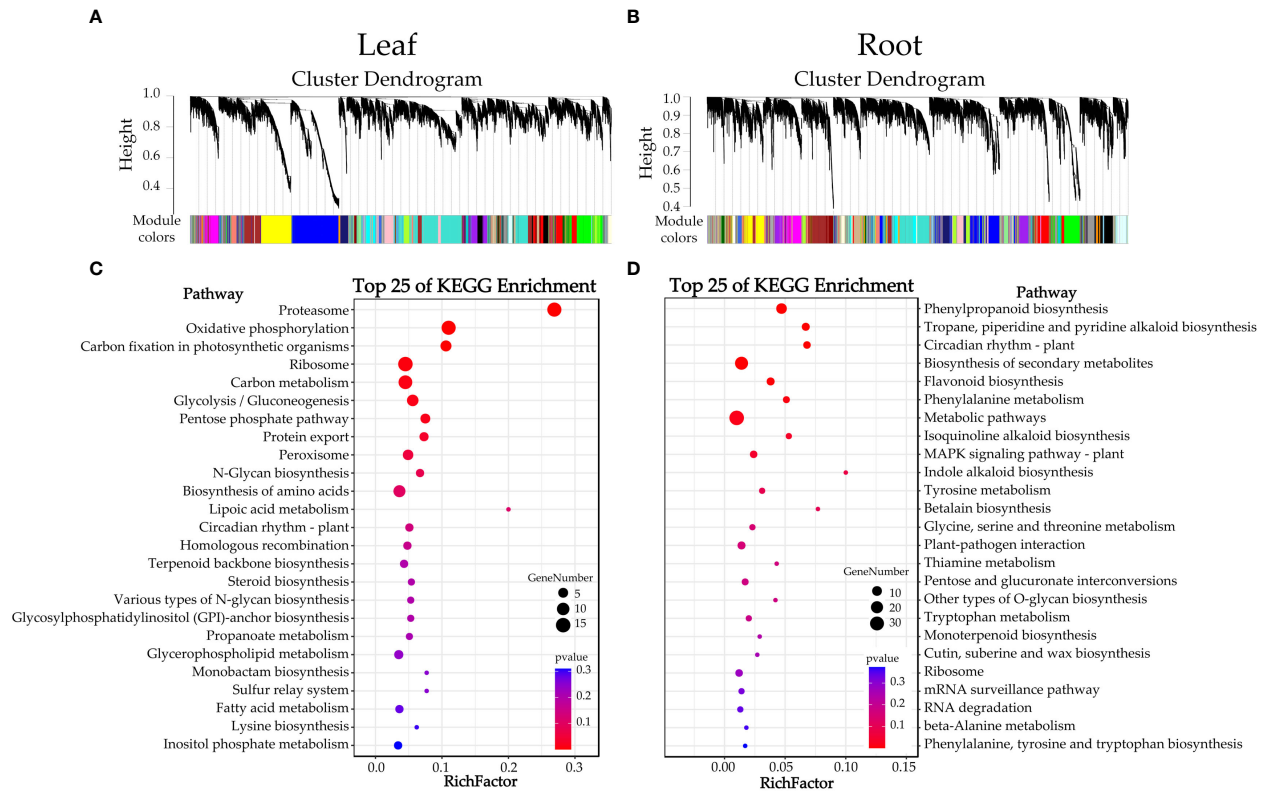


FIGURE 6
Clusters dendrograms and KEGG enrichment analysis. **(A, B)** indicate the merged clusters and dendrograms, **(C, D)** indicate the KEGG enrichment analysis of genes in the 'salmon' and 'skyblue' modules, respectively.

and Zy, while gene_30857 only expressed in the leaves of Zy after treatment (Figure 7).

Discussion

Physiological and morphological response to waterlogging stress

Field observations of waterlogging responses in citrus have shown that the roots of Rt suffered greater damage than those of Pj and Zy (Figure 1A). In this study, the leaf chlorophyll content detected by SPAD in Pj showed the smallest reduction as compared to those of the other two rootstocks. Additionally, waterlogging stress slightly increased the stem length (Figure 1B), which is consistent with previous results in soybean (Kim et al., 2015). Waterlogging stress also affected plant growth and development, thereby causing morphological developments, including aerenchyma formation, adventitious root formation, and shoot elongation (Qi et al., 2019). Here, all rootstock seedlings under waterlogging stress showed increased adventitious root formation, with Pj being significantly higher than in other rootstocks (Figure 1B). Stress triggers diverse plant responses, including physiological and metabolic process (Zhang, H. et al., 2020; Zhang, H.M. et al., 2022). A series of antioxidant enzymes generated following ROS accumulation in plant, helped eliminate the excess ROS. In this study, the soluble sugar, soluble protein, and MDA contents increased in the leaf, while the POD and CAT activities increased in roots of stress tolerant citrus, respectively (Figure 2). Therefore, these might keep ROS at a low level in Pj.

Gene transcription changes in citrus under waterlogging stress

The waterlogging stress-responsive molecular and physiological events have been reported in diverse plant species (Pan et al., 2021). Recent studies focusing on citrus rootstocks have revealed that hypoxia stress can lead to a significant decrease in mineral element contents and identification of 232 hypoxia-responsive genes (Xie et al., 2021). However, there is still limited knowledge about the underlying waterlogging stress-responsive molecular mechanisms in citrus under waterlogging stress. In this study, there were 154 DEGs in the leaves and 147 DEGs in the root after comparing the gene expression of three rootstocks. Typically, a range of energy-related and oxygen-consuming metabolic pathways were differentially regulated to mitigate the harmful effects of waterlogging stress (Blokhina and Fagerstedt, 2010). In this study, numerous genes related to 'carbon fixation in photosynthetic organisms' in the root (Figure 3D).

Hormone has long been known to be involved in waterlogging stress responses. Several genes associated with auxin and the ABA

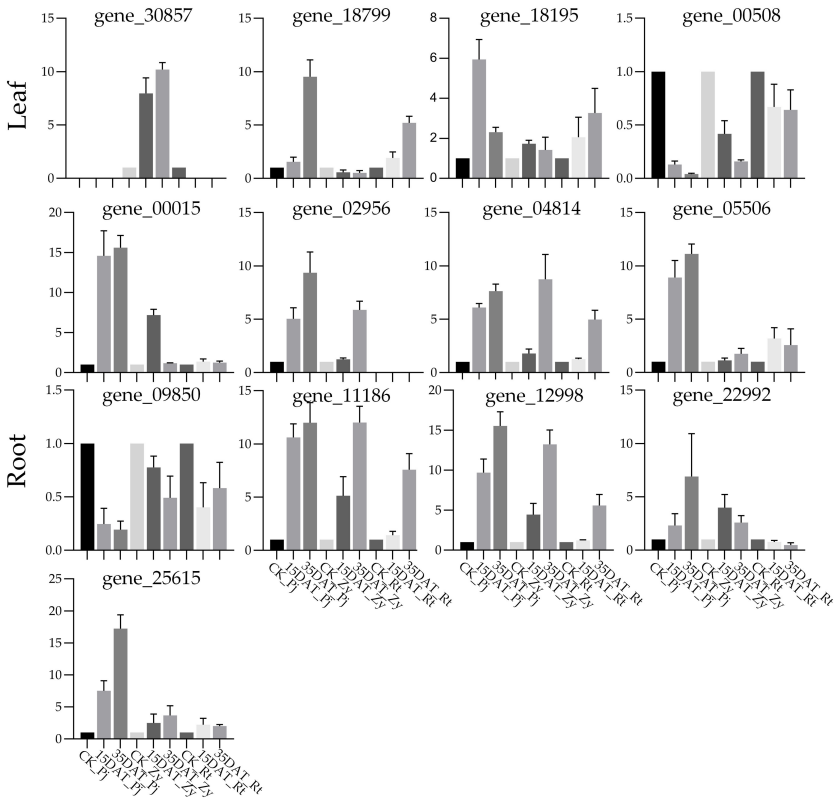


FIGURE 7
Verification of the expression of eight candidate genes' in different rootstocks. The error bars with standard deviations are calculated from three biological replicates.

pathway were found to have a significantly expressed in response to waterlogging stress. ABA has been regarded as being closely related to water stress (Zhang, Q. et al., 2022). We identified one ABSCISIC ACID-INSENSITIVE 5-like protein (ABI, gene_30411) in the root, which confers hypersensitivity to ABA and sugar (Brocard et al., 2002). Additionally, one ABA-responsive element binding factor (ABF, gene_14446) was differentially expression in leaf (Figure 4A). ABA is able to integrate auxin signaling to modulate plant performance under different stress conditions (Emenecker and Strader, 2020). There were two different expression Auxin/Indole-3-Acetic Acid genes in the leaf (gene_01190) and root (gene_19419), respectively. Auxin/IAA is induced by auxin and together with *SAUR* and *GH* family genes, and they are collectively referred to as early auxin-induced genes (Yu et al., 2022). Moreover, we identified three cytokinin-related genes (one histidine-containing phosphotransfer protein, gene_21726; two two-component response regulators, gene_05220 and gene_24622) in the leaf, which have been previously associated with plant growth and stress tolerance (Merewitz et al., 2010).

Transcription factors involved in waterlogging stress

Several types of TFs have been shown to be involved in abiotic stresses. Two specific TFs, namely *MxWRKY64* and *MxbHLH18*, which are extracted from *Malus xiaojinensis*, have been found to play an effective role in enhancing salt tolerance (Han et al., 2021; Liang et al., 2022). Moreover, four other TFs, *MbMYB4* and *MbMYB108* derived from *Malus baccata*, and *ERF9* and *ERF108* derived from *Poncirus trifoliata* have been proven useful in elevating cold tolerance (Khan et al., 2021; Yao et al., 2022a; Yao et al., 2022b; Zhang, Y. et al., 2022). Through various research results, it has been increasingly apparent that such TFs can play a significant role in improving plant resilience against waterlogging stress. *MaRAP2-4* showed enhanced waterlogging and subsequent oxidative stress tolerance (Phukan et al., 2018). RNA-seq results also revealed the involvement of WRKY, MYB, bHLH, NAC, ERFs, DOF, HD-ZIP and DBP in hypoxic stress (Xie et al., 2021; Wang et al., 2023). In this study, 16 TFs, including AP2/ERFs, bHLHs, HSFs and WRKYs, were found to be associated with waterlogging stress. However, due to the limited data available on these waterlogging-responsive TFs, their functions need to be intensively dissected in the future.

Underlying mechanisms of waterlogging tolerance in citrus

To elucidate the molecular basis of waterlogging tolerance in *C. junos*, we developed a working model based on our results (Figure 8). According to this model, the transcription factors AP2/ERF, bHLH and HSF may modulate the expression of DEGs involved in auxin, cytokinine and ABA pathways in leaves (Zhao et al., 2020; Li et al., 2022; Li et al., 2023). This may lead to physiological and biochemical adjustments, such as enhanced sugar content, protein content and MDA content. Similarly, the transcription factors HSFs, EREBP and WRKYs may regulate the expression of DEGs in roots (Zhang, C. et al., 2020; Liu et al., 2021). These differentially expressed TFs may alter metabolic processes, such as elevating CAT activity and POD activity. These changes may ultimately confer waterlogging tolerance to citrus plants.

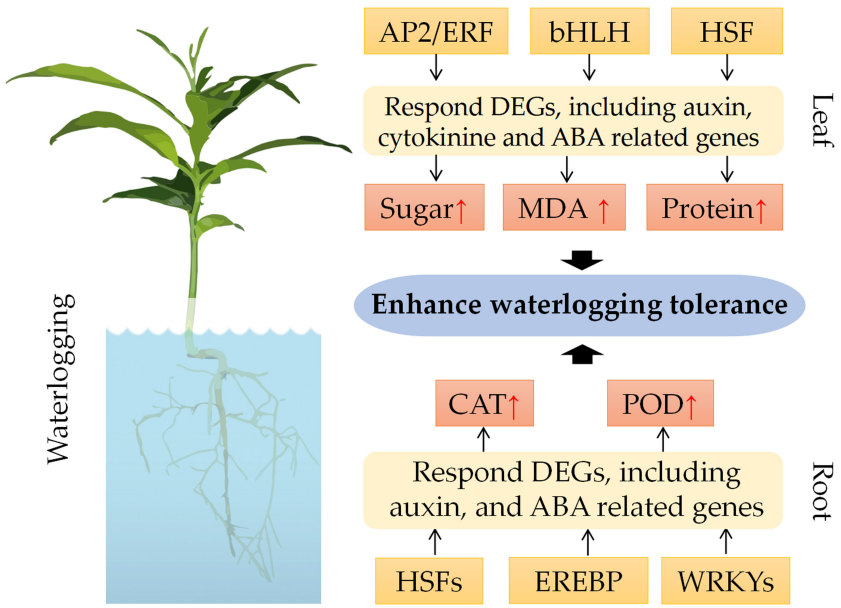


FIGURE 8
Proposed model for the waterlogging tolerance in *Citrus junos*.

Conclusions

Waterlogging stress negatively affected citrus growth, which decreased the SPAD value and root number. Under waterlogging stress, the soluble sugar, soluble protein, and MDA contents in the leaf increased, while the CAT and POD activities increased in the root. RNA-seq analysis identified 154 DEGs in the leaves and 147 DEGs in the root. Processes like 'cutin, suberine and wax biosynthesis', 'diterpenoid biosynthesis', and 'glycerophospholipid metabolism' were enriched in the leaf, while those linked to 'flavonoid biosynthesis', 'biosynthesis of secondary metabolites', and 'metabolic pathways' were seen in the root. Furthermore, we identified 17 differentially expressed TFs, which mainly are important in the waterlogging tolerance of citrus. Therefore, we proposed the working model of waterlogging tolerance based on those evidences. Our findings will deepen our understandings for the mechanism of waterlogging tolerance in citrus.

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

Conceptualization and supervision, WH and XW. Methodology, WH and LL. Investigation, LL, ZW, SY, ML, YL, YZ, YXL, YTZ, and HW. Bioinformatic analyses, WH and QC. Data curation, WH, LL, RX, and JC. Manuscript preparation, WH and YW. Writing-review and editing, WH, LL, and XW. All authors contributed to the article and approved the submitted version.

Funding

This research was founded by Sichuan Provincial Postdoctoral Science Foundation, and Shuangzhi Project Innovation Team of Sichuan Agricultural University (Grant No. P202107).

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1198930/full#supplementary-material

References

Bailey-Serres, J., Fukao, T., Gibbs, D. J., Holdsworth, M. J., Lee, S. C., Licausi, F., et al. (2012). Making sense of low oxygen sensing. *Trends Plant Sci.* 17, 129–138. doi: 10.1016/j.tplants.2011.12.004

Blokhina, O., and Fagerstedt, K. V. (2010). Oxidative metabolism, ROS and no under oxygen deprivation. *Plant Physiol. Biochem.* 48, 359–373. doi: 10.1016/j.plaphy.2010.01.007

Brocard, I. M., Lynch, T. J., and Finkelstein, R. R. (2002). Regulation and role of the *Arabidopsis* abscisic acid-insensitive 5 gene in abscisic acid, sugar, and stress response. *Plant Physiol.* 129, 1533–1543. doi: 10.1104/pp.005793

Cheng, X. X., Yu, M., Zhang, N., Zhou, Z. Q., Xu, Q. T., Mei, F. Z., et al. (2016). Reactive oxygen species regulate programmed cell death progress of endosperm in winter wheat (*Triticum aestivum* l.) under waterlogging. *Protoplasma* 253, 311–327. doi: 10.1007/s00709-015-0811-8

Christianson, J. A., Llewellyn, D. J., Dennis, E. S., and Wilson, I. W. (2010). Global gene expression responses to waterlogging in roots and leaves of cotton (*Gossypium hirsutum* l.). *Plant Cell Physiol.* 51, 21–37. doi: 10.1093/pcp/pcp163

Emenecker, R. J., and Strader, L. C. (2020). Auxin-abscisic acid interactions in plant growth and development. *Biomolecules* 10, 281. doi: 10.3390/biom10020281

Fu, X. K., Huang, X. J., Chen, T., Zhang, J., Wang, Y., Chen, Q., et al. (2017). A new citrus rootstock 'Pujiang xiangcheng' (*Citrus junos*). *J. Fruit Sci.* 34, 917–920. doi: 10.13925/j.cnki.gsxb.20170069(In Chinese)

García-Sánchez, F., Syvertsen, J. P., Gimeno, V., Botía, P., and Perez-Perez, J. G. (2007). Responses to flooding and drought stress by two citrus rootstock seedlings with different water-use efficiency. *Physiol. Plant* 130, 532–542. doi: 10.1111/j.1399-3054.2007.00925.x

Ghaffari, H., Tadayon, M. R., Bahador, M., and Razmjoo, J. (2021). Investigation of the proline role in controlling traits related to sugar and root yield of sugar beet under water deficit conditions. *Agric. Water Manage.* 243, 106448. doi: 10.1016/j.agwat.2020.106448

Gong, Z. Z., Xiong, L. M., Shi, H. Z., Yang, S. H., Herrera-Estrella, L. R., Xu, G. H., et al. (2020). Plant abiotic stress response and nutrient use efficiency. *Sci. China Life Sci.* 63, 635–674. doi: 10.1007/s11427-020-1683-x

Han, D. G., Han, J. X., Xu, T. L., Li, T. M., Yao, C. Y., Wang, Y. J., et al. (2021). Isolation and preliminary functional characterization of *MxWRKY64*, a new WRKY transcription factor gene from *Malus xiaojinensis* Cheng et jiang. *In Vitro Cell Dev-Pl* 57, 202–213. doi: 10.1007/s11627-021-10171-7

Hattori, Y., Nagai, K., Furukawa, S., Song, X. J., Kawano, R., Sakakibara, H., et al. (2009). The ethylene response factors *SNORKE1* and *SNORKE2* allow rice to adapt to deep water. *Nature* 460, 1026–1030. doi: 10.1038/nature08258

He, W., Wang, Y., Chen, Q., Sun, B., Tang, H. R., Pan, D. M., et al. (2018). Dissection of the mechanism for compatible and incompatible graft combinations of citrus grandis (l.) osbeck ('Hongmian miyou'). *Int. J. Mol. Sci.* 19, 505. doi: 10.3390/ijms19020505

He, W., Xie, R., Li, H., Wang, Y., Chen, Q., Lin, Y. X., et al. (2021). Evaluation of suitable qRT-PCR normalization genes for various citrus rootstocks. *Plant Biotechnol. Rep.* 16, 101–111. doi: 10.1007/s11816-021-00725-x

He, W., Xie, R., Wang, Y., Chen, Q., Wang, H., Yang, S. F., et al. (2022). Comparative transcriptomic analysis on compatible/incompatible grafts in *Citrus*. *Hortic. Res.* 9, b72. doi: 10.1093/hr/uhab072

Khan, M., Hu, J. B., Dahro, B., Ming, R. H., Zhang, Y., Wang, Y., et al. (2021). ERF108 from *Poncirus trifoliata* (L.) raf. functions in cold tolerance by modulating raffinose synthesis through transcriptional regulation of *PtrRafs*. *Plant J.* 108, 705–724. doi: 10.1111/tpj.15465

Kim, Y. H., Hwang, S. J., Waqas, M., Khan, A. L., Lee, J. H., Lee, J. D., et al. (2015). Comparative analysis of endogenous hormones level in two soybean (*Glycine max* l.) lines differing in waterlogging tolerance. *Front. Plant Sci.* 6. doi: 10.3389/fpls.2015.00714

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., Salzberg, S. L., et al. (2013). Tophat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14, R36. doi: 10.1186/gb-2013-14-4-r36

Langfelder, P., and Horvath, S. (2008). WGCNA: an r package for weighted correlation network analysis. *BMC Bioinform.* 9, 559. doi: 10.1186/1471-2105-9-559

Li, B., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. *BMC Bioinform.* 12, 323. doi: 10.1186/1471-2105-12-323

Li, S. J., Liu, S. C., Lin, X. H., Grierson, D., Yin, X. R., and Chen, K. S. (2022). Citrus heat shock transcription factor *CitHSFa7*-mediated citric acid degradation in response to heat stress. *Plant Cell Environ.* 45, 95–104. doi: 10.1111/pce.14207

Li, C. W., Su, J. S., Zhao, N., Lou, L., Ou, X. L., Yan, Y. J., et al. (2023). *CmERF5-CmRAP2.3* transcriptional cascade positively regulates waterlogging tolerance in *Chrysanthemum morifolium*. *Plant Biotechnol. J.* 21, 270–282. doi: 10.1111/pbi.13940

Liang, X. Q., Li, Y. M., Yao, A. Q., Liu, W. D., Yang, T. Y., Zhao, M. F., et al. (2022). Overexpression of *MxbHLH18* increased iron and high salinity stress tolerance in *Arabidopsis thaliana*. *Int. J. Mol. Sci.* 23, 8007. doi: 10.3390/ijms23148007

Liu, B., Jiang, Y. Z., Tang, H., Tong, S. F., Lou, S. L., Shao, C., et al. (2021). The ubiquitin E3 ligase SR1 modulates the submergence response by degrading phosphorylated *WRKY33* in *Arabidopsis*. *Plant Cell* 33, 1771–1789. doi: 10.1093/plcell/koab062

Merewitz, E. B., Gianfagna, T., and Huang, B. (2010). Photosynthesis, water use, and root viability under water stress as affected by expression of SAG12-IPT controlling cytokinin synthesis in *Agrostis stolonifera*. *J. Exp. Bot.* 62, 383–395. doi: 10.1093/jxb/erq285

Müllner, D. (2013). Fastcluster: fast hierarchical, agglomerative clustering routines for r and Python. *J. Stat. Software* 53(9), 1–18. doi: 10.18637/jss.v053.i09

Muñoz, N., González, C., Molina, A., Zirulnik, F., and Luna, C. M. (2008). Cadmium-induced early changes in O^{2}–, H_2O_2 and antioxidative enzymes in soybean (*Glycine max* l.) leaves. *Plant Growth Regul.* 56, 159–166. doi: 10.1007/s10725-008-9297-0

Pan, J., Sharif, R., Xu, X., and Chen, X. (2021). Mechanisms of waterlogging tolerance in plants: research progress and prospects. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.627331

Phukan, U. J., Jeena, G. S., Tripathi, V., and Shukla, R. K. (2018). *MaRAP2-4*, a waterlogging-responsive erf from *Mentha*, regulates bidirectional sugar transporter *AtSweet10* to modulate stress response in *Arabidopsis*. *Plant Biotechnol. J.* 16, 221–233. doi: 10.1111/pbi.12762

Qi, X. H., Li, Q. Q., Ma, X. T., Qian, C. L., Wang, H. H., Ren, N. N., et al. (2019). Waterlogging-induced adventitious root formation in cucumber is regulated by ethylene and auxin through reactive oxygen species signaling. *Plant Cell Environ.* 42, 1458–1470. doi: 10.1111/pce.13504

Ren, B., Yu, W., Liu, P., Zhao, B., and Zhang, J. (2023). Responses of photosynthetic characteristics and leaf senescence in summer maize to simultaneous stresses of waterlogging and shading. *Crop J.* 11, 269–277. doi: 10.1016/j.cj.2022.06.003

Sampaio, A. H. R., Silva, R. O., Brito, R. B. F., Filho, W. D. S. S., Gesteira, A. D. S., Souza, L. D., et al. (2021). Sweet orange acclimatisation to water stress: a rootstock dependency. *Sci. Hortic.* 276, 109727. doi: 10.1016/j.scienta.2020.109727

Shahid, M. A., Balal, R. M., Khan, N., Simón-Grao, S., Alfosea-Simón, M., Cámara-Zapata, J. M., et al. (2019). Rootstocks influence the salt tolerance of kinnow mandarin trees by altering the antioxidant defense system, osmolyte concentration, and toxic ion accumulation. *Sci. Hortic.* 250, 1–11. doi: 10.1016/j.scienta.2019.02.028

Shen, C. W., Yuan, J. P., Qiao, H., Wang, Z. J., Liu, Y. H., Ren, X. J., et al. (2020). Transcriptomic and anatomic profiling reveal the germination process of different wheat varieties in response to waterlogging stress. *BMC Genet.* 21, 93. doi: 10.1186/s12863-020-00901-y

Tahjib-Ul-Arif, M., Hasan, M. T., Rahman, M. A., Nuruzzaman, M., Rahman, A. M. S., Hasanuzzaman, M., et al. (2023). Plant response to combined salinity and waterlogging stress: current research progress and future prospects. *Plant Stress* 7, 100137. doi: 10.1016/j.stress.2023.100137

Teoh, E. Y., Teo, C. H., Baharum, N. A., Pua, T., and Tan, B. C. (2022). Waterlogging stress induces antioxidant defense responses, aerenchyma formation and alters metabolisms of banana plants. *Plants* 11, 2052. doi: 10.3390/plants11152052

Voesenek, L. A. C. J., and Bailey Serres, J. (2015). Flood adaptive traits and processes: an overview. *New Phytol.* 206, 57–73. doi: 10.1111/nph.13209

Wang, Y. X., Xu, Y., Xu, J. M., Sun, W. X., Lv, Z. X., Manzoor, M. A., et al. (2023). Oxygenation alleviates waterlogging-caused damages to cherry rootstocks. *Mol. Hortic.* 3, 8. doi: 10.1186/s43897-023-00056-1

Wei, X. N., Xu, H. J., Rong, W., Ye, X. G., and Zhang, Z. Y. (2019). Constitutive expression of a stabilized transcription factor group vii ethylene response factor enhances waterlogging tolerance in wheat without penalizing grain yield. *Plant Cell Environ.* 42, 1471–1485. doi: 10.1111/pce.13505

Xie, R. J., Zheng, L., Jiao, Y., and Huang, X. (2021). Understanding physiological and molecular mechanisms of citrus rootstock seedlings in response to root zone hypoxia by RNA-seq. *Environ. Exp. Bot.* 192, 104647. doi: 10.1016/j.envexpbot.2021.104647

Yao, C. Y., Li, X. G., Li, Y. M., Yang, G. H., Liu, W. D., Shao, B. T., et al. (2022b). Overexpression of a *Malus baccata* MYB transcription factor gene *MbMYB4* increases cold and drought tolerance in *Arabidopsis thaliana*. *Int. J. Mol. Sci.* 23, 1794. doi: 10.3390/ijms23031794

Yao, C. Y., Li, W. H., Liang, X. Q., Ren, C. K., Liu, W. D., Yang, G. H., et al. (2022a). Molecular cloning and characterization of *MbMYB108*, a *Malus baccata* MYB transcription factor gene, with functions in tolerance to cold and drought stress in transgenic *Arabidopsis thaliana*. *Int. J. Mol. Sci.* 23, 4846. doi: 10.3390/ijms23094846

Yin, D. M., Sun, D. Y., Han, Z. Q., Ni, D., Norris, A., and Jiang, C. Z. (2019). *PhERF2*, an ethylene-responsive element binding factor, plays an essential role in waterlogging tolerance of *Petunia*. *Hortic. Res.* 6, 83. doi: 10.1038/s41438-019-0165-z

Yu, Z. P., Zhang, F., Friml, J., and Ding, Z. J. (2022). Auxin signaling: research advances over the past 30 years. *J. Integr. Plant Biol.* 64 (2), 371–392. doi: 10.1111/jipb.13225

Zaher-Ara, T., Boroomand, N., and Sadat-Hosseini, M. (2016). Physiological and morphological response to drought stress in seedlings of ten citrus. *Trees* 30, 985–993. doi: 10.1007/s00468-016-1372-y

Zhang, J. Y., Huang, S. N., Mo, Z. H., Xuan, J. P., Jia, X. D., Wang, G., et al. (2015). *De novo* transcriptome sequencing and comparative analysis of differentially expressed genes in kiwifruit under waterlogging stress. *Mol. Breed.* 35, 208. doi: 10.1007/s11032-015-0408-0

Zhang, Y., Ming, R. H., Khan, M., Wang, Y., Dahro, B., Xiao, W., et al. (2022). ERF9 of *Poncirus trifoliata* (L.) raf. undergoes feedback regulation by ethylene and modulates cold tolerance *via* regulating aglutathione s-transferase u17 gene. *Plant Biotechnol. J.* 20, 183–200. doi: 10.1111/pbi.13705

Zhang, C. L., Wang, Y. X., Hu, X., Zhang, Y. L., Wang, G. L., You, C. X., et al. (2020). An apple AP2/EREBP-type transcription factor, *MdWRI4*, enhances plant resistance to abiotic stress by increasing cuticular wax load. *Environ. Exp. Bot.* 180, 104206. doi: 10.1016/j.envexpbot.2020.104206

Zhang, Q., Yuan, W., Wang, Q. W., Cao, Y. Y., Xu, F. Y., Dodd, L. C., et al. (2022). ABA regulation of root growth during soil drying and recovery can involve auxin response. *Plant Cell Environ.* 45, 871–883. doi: 10.1111/pce.14137

Zhang, H., Zhao, Y., and Zhu, J. K. (2020). Thriving under stress: how plants balance growth and the stress response. *Dev. Cell* 55, 529–543. doi: 10.1016/j.devcel.2020.10.012

Zhang, H. M., Zhu, J. H., Gong, Z. Z., and Zhu, J. K. (2022). Abiotic stress responses in plants. *Nat. Rev. Genet.* 23, 104–119. doi: 10.1038/s41576-021-00413-0

Zhao, Q., Fan, Z., Qiu, L., Che, Q. Q., Wang, T., Li, Y. Y., et al. (2020). *MdbHLH130*, an apple bHLH transcription factor, confers water stress resistance by regulating stomatal closure and ROS homeostasis in transgenic tobacco. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.543696

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 07 June 2023
DOI 10.3389/fpls.2023.1182977

OPEN ACCESS

EDITED BY
Chenxia Cheng,
Qingdao Agricultural University, China

REVIEWED BY
Yaping Kou,
Chinese Academy of Agricultural Sciences (CAAS), China
Ming Sun,
Beijing Forestry University, China

*CORRESPONDENCE
Jiahui Liang
jiahuiliang1230@163.com

RECEIVED 09 March 2023
ACCEPTED 10 May 2023
PUBLISHED 07 June 2023

CITATION
Cao X, Sui J, Li H, Yue W, Liu T, Hou D, Liang J and Wu Z (2023) Enhancing heat stress tolerance in Lanzhou lily (*Lilium davidii* var. *unicolor*) with Trichokonins isolated from *Trichoderma longibrachiatum* SMF2.
Front. Plant Sci. 14:1182977.
doi: 10.3389/fpls.2023.1182977

Enhancing heat stress tolerance in Lanzhou lily (*Lilium davidii* var. *unicolor*) with Trichokonins isolated from *Trichoderma longibrachiatum* SMF2

Xing Cao[1], Juanjuan Sui[2], Haiyan Li[1], Wenxiu Yue[1], Tao Liu[1], Dong Hou[3], Jiahui Liang[4]* and Ze Wu[5]

[1]Department of Environmental Art Design, College of Architecture, Yantai University, Yantai, China, [2]Engineering Technology Research Center of Anti-aging Chinese Herbal Medicine, Biology and Food Engineering College, Fuyang Normal University, Fuyang, China, [3]Vegetable Research Institute, Gansu Academy of Agricultural Sciences, Lanzhou, China, [4]Institute of Grassland, Flowers and Ecology, Key Laboratory of Urban Agriculture (North), Ministry of Agriculture, Beijing Academy of Agriculture and Forestry Sciences, Beijing, China, [5]Key Laboratory of Landscaping Agriculture, Ministry of Agriculture and Rural Affairs, College of Horticulture, Nanjing Agricultural University, Nanjing, China

Lanzhou lily (*Lilium davidii* var. *unicolor*) is a renowned edible crop produced in China and relatively sensitive to high temperature (HT). Trichokonins (TKs) are antimicrobial peptaibols secreted from *Trichoderma longibrachiatum* strain SMF2. Here, we report that TKs application improves the thermotolerance of Lanzhou lily. The activity of the antioxidant enzyme system (SOD, CAT, and POD), the level of heat-resistance-associated phytohormones (ABA, SA, and JA), the relative water content (RWC), the content of chlorophyll (Chl), and the net photosynthetic rate (P_n) were promoted by TKs treatment in Lanzhou lily plants subjected to heat stress (HS). TKs treatment also mitigated cell injury as shown by a lower accumulation of malondialdehyde (MDA) and relative electrolyte leakage (REL) under HS conditions. RNA-seq data analysis showed that more than 4.5 times differentially expressed genes (DEGs) responded to TKs treatment under HS compared to non-HS, and TKs treatment reduced protein folding and enhanced cellular repair function under HS conditions. The analyses of DEGs involved in hormone (ABA, SA and JA) synthesis and signaling pathways suggested that TKs might improve Lanzhou lily heat tolerance by promoting ABA synthesis and signal transduction. TKs highly induced DEGs of the HSF-HSP pathway under HS, in which *HSFA2* accounted for most of the HSF family. Furthermore, TKs treatment resulted in the upregulation of heat-protective genes *LzDREB2B*, *LzHsfA2a*, *LzMBF1c*, *LzHsp90*, and *LzHsp70* involved in HSF-HSP signal pathway after long-term HS. LzHsfA2a-1 likely plays a key role in acquisition of TKs-induced thermotolerance of Lanzhou lily as evidenced by the sustained response to HS, the enhanced response to TKs treatment under long-term HS, and the high sequence similarity to LlHsfA2a which is a key regulator for the improvement of heat tolerance in *Lilium longiflorum*. Our results reveal the underlying mechanisms of TKs-mediated thermotolerance in Lanzhou lily and highlight an attractive approach to protecting crop plants from damage caused by HS in a global warming future.

KEYWORDS

Lanzhou lily, Trichokonins, thermotolerance, physiological response, RNA-Seq, Hsf

Introduction

High temperature (HT) is a major environmental stress that limits plant growth and development, particularly crop yields. As sessile organisms, plants have evolved various physiological and molecular defense mechanisms in response to heat stress (HS). HS causes the overproduction of reactive oxygen species (ROS), lipid peroxidation, photoinhibition, protein denaturation and degradation, and RNA damage, which are responsible for an imbalance in cellular homeostasis (Bokszczanin and Fragkostefanakis, 2013). Meanwhile, factors involved in signaling cascades and transcriptional control, such as ion transporters, proteins, osmoprotectants and antioxidants, are activated by HS to maintain cellular homeostasis (Hasanuzzaman et al., 2013). Understanding the mechanisms underlying heat stress response (HSR) are essential to improving thermotolerance *via* conventional breeding, genetic engineering or protectants applications. The unfolded protein response in endoplasmic reticulum and the heat shock response in cytoplasmic are two evolutionarily conserved systems that protect plants from HS (Li et al., 2020). A complex signal transduction network may integrate signals from all these different sensors involving calcium fluxes, ROS, calmodulin, CDPKs, MAPKs, phosphatases, and transcriptional regulators such as heat stress transcription factor (HSF), MBF1c, WRKY, NAC, DREB, bZIP, MYB, and bHLH (Koini et al., 2009; Zhao et al., 2017; Li et al., 2020; Cai et al., 2021; Haider et al., 2021; Saini et al., 2022; Tian et al., 2022). Among the complex HS signal transduction networks, heat stress transcription factor- heat shock proteins (HSF-HSP) pathway plays a central role in the acquisition of heat tolerance in plants. HS leads to the synthesis and accumulation of HSPs functioning as molecular chaperones to control protein homeostasis. *HSP* genes are regulated by HSFs that can recognize and bind to heat shock elements (HSEs) in the *HSP* promoter regions. It is reported that HSFA2 is a master transcriptional regulator of thermomemory in *Arabidopsis* (Liu et al., 2019; Friedrich et al., 2021).

Lanzhou lily (*Lilium davidii* var. *unicolor*), an endemic species in China, is a famous edible crop for its large, white, sweet bulb scales (Cao et al., 2020). Because of its high edible, medicinal, and ornamental value, Lanzhou lily plays important roles in landscape agriculture and rural revitalization. Lanzhou lily well adapts to cool and humid conditions, and HT may lead to growth stagnation and bulb degradation. Due to its poor heat resistance, HT is a key limiting factor to the annual production of Lanzhou lily bulbs. Measures such as breeding or plant growth regulator (PGR) application must be taken to overcome this limitation. In some studies, OT lily cultivar 'Jinmen' selected from the cross-combination 'T11'(♀)and 'D74'(♂) shows strong heat resistance (Cui et al., 2014). CaM3, HSFA1, HSFA2, HSFA3, and DREB2 are important components of HS signal transduction in lilies (Xin et al., 2010; Cao et al., 2013; Gong et al., 2014; Xin et al., 2017; Wu et al., 2018a; Wu et al., 2018b). Pretreatment of salicylic acid (SA) increases the heat tolerance of lily by enhancing the activities of antioxidant systems (Chen et al., 2008). To date, protectants for improving the thermotolerance of Lanzhou lily have not been reported.

In recent times, exogenous applications of protectants in the form of osmoprotectants (e.g., proline, Pro; glycine betaine, GB; and trehalose, Tre) (Wahid and Shabbir, 2005; Kaushal et al., 2011; Luo et al., 2014), phytohormones (e.g., abscisic acid, ABA; jasmonic acid, JA; epibrassinolide, EBL; and salicylic acid, SA) (Wang and Li, 2006; Scalabrin et al., 2016; Alam et al., 2018; Islam et al., 2018), signaling molecules (e.g., nitric oxide, NO) (Li X, et al., 2014), polyamines (e.g., putrescine, Put; spermidine, Spd; and spermine, Spm) (Fuller and Gerner, 1982; Kumar et al., 2012), trace elements (e.g., selenium, Se; and silicon, Si) (Malerba and Cerana, 2018; Hussain et al., 2019) and nutrients (e.g., nitrogen, N; phosphorus, P; potassium, K, and calcium, Ca) (Waraich et al., 2012; Cao et al., 2013) have been found effective in mitigating HT stress-induced damage in plants. Some of the protectants, also known as plant inducers, can be synthetic compounds, natural products or living microorganisms. When applied exogenously at low doses before upcoming HS or at the early stage of stress occurrence, inducers can rapidly activate the plant's defense systems and then enhance plant resistance to HS. Beneficial microbes, such as plant growth-promoting bacteria and plant growth-promoting fungi, can modulate the plant transcriptome and metabolome to induce thermotolerance (Shekhawat et al., 2022). Compounds produced by *Enterobacter* sp. SA187 trigger ethylene signaling in *Arabidopsis* to regulate HS-tolerance *via* increased expression of *EIN3* and *HSFA2* genes (Zélicourt et al., 2018). 2-Amino-3-methylhexanoic acid (AMHA) isolated from *Alternaria alternata* facilitates HT resistance by alleviating physiological damage in field-grown tea plants due to improved photosynthetic performance, osmotic adjustments and antioxidant enzyme activities (Yang et al., 2023). Trichokonins (TKs) are antimicrobial peptaibols extracted from *Trichoderma longibrachiatum* strain SMF2 (Song et al., 2006). Our previous studies revealed that TKs exhibited broad-spectrum antimicrobial activity against bacteria, fungal phytopathogens and TMV (Luo et al., 2010; Shi et al., 2012; Li H, et al., 2014). However, to our knowledge, no studies have addressed the function of TKs on plant abiotic stresses. Here, we demonstrate the effect and mechanism of TKs on regulating the thermotolerance of Lanzhou lily.

Materials and methods

Plant materials and growth conditions

Lanzhou lily (*Lilium davidii* var. *unicolor*) bulbs were derived from Qilihe district, Lanzhou city, China, and cultivated in Dongchangfu district, Liaocheng city (S 115°97', W 36°45'), China. Sterilized bulbs with uniform size (bulb circumference 5.0 ± 0.15 cm) were potted in plastic pots (height: 9 cm, top square with side length: 10 cm, bottom square with side length: 8 cm) containing a sterile 1:1 mixture of sphagnum peat and sand and grown in a greenhouse under suitable growth conditions (16 h/8 h light/dark and 25°C/20°C light/dark). After cultivation for 21 d, healthy uniform size plants with 9.40 ± 0.20 cm high and 0.34 ± 0.02 cm stem diameter were prepared for thermotolerance test. After cultivation for 28 d, healthy uniform size plants with 10.61 ± 0.35 cm high and 0.43 ± 0.03 cm stem diameter were prepared for physiology and RNA-seq experiments.

Trichokonins preparation and heat treatment

Trichokonins (TKs) stock solution with a 5 mg/mL concentration prepared as previously described (Luo et al., 2010). In different experiments, distilled water was used for further dilution of TKs. Roots of Lanzhou lily plants were treated with 30 mL solution of 0.5, 1, 2, 4, and 8 mg/L TKs, respectively, at 22°C for 12 h, then subjected to a direct HS at 40°C for 72 h. Control plants were treated with distilled water. After HS, the survival rate was recorded with a 30-d-recovery period. Each treatment consisted of three replicates and twelve individual plants served as one replicate.

Physiological measurements

After root-irrigation with 30 mL solution of 2 mg/L TKs or 30 mL distilled water (non-treated control) at 22°C for 12 h, Lanzhou lily plants were exposed to different durations (24, 48, and 72 h) at 40°C (HS conditions) or 22°C (non-stress control). The middle part leaves of plants were collected for physiological indices measurement. Five plants were pooled together as one biological replicate, and each treatment was repeated three times.

The relative water content (RWC) was calculated as previously described (Lafitte, 2002). Relative electrolyte leakage (REL) was evaluated according to the method of Lutts et al. (1996). Chlorophyll (Chl) content was quantified following the method of Frank et al. (2005) with slight modification. Net photosynthetic rate (P_n) was measured as described by Zhang et al. (2013) using the portable photosynthesis measurement system LI-6800 (LI-COR, USA). The malondialdehyde (MDA) content and the activities of superoxide dismutase (SOD, EC 1.15.1.1), catalase (CAT, EC 1.11.1.6), and peroxidase (POD, EC 1.11.1.7) were assayed using MDA Test Kit (A003-1), Total SOD Assay Kit (A001-1), CAT Assay Kit (A007-1), and Plant POD Assay Kit (A084-3), respectively, according to the mentioned manufacturer's protocol (Nanjing Jiancheng, Nanjing, China).

The ABA, SA, and JA were extracted and measured as described previously, with minor modifications (Liu et al., 2014). Briefly, about 500 mg of plant leaf tissue was ground into a fine powder in liquid nitrogen and then phytohormones were extracted twice with acetonitrile and purified with a Poroshell 120 SB-C18 column. The quantification of ABA, SA, and JA was determined by high-performance liquid chromatography-electrospray ionization-tandem mass spectrometry (HPLC-ESI-MS/MS) using an Agilent 1260 HPLC coupled to an 6420A MS (Agilent Technologies Inc., USA).

De Novo transcriptome sequencing and analysis

Two groups of Lanzhou lily plants were prepared for RNA-seq. One group was the Lanzhou lily plants treated with 2 mg/L TKs or distilled water for 12 h at 22°C (non-HS, RT), marked as "TKs" and "W" respectively. The other group of samples was Lanzhou lily plants treated with 2 mg/L TKs or distilled water for 12 h at RT followed by 40°C HS treatment for 12 h, marked as "HS+TKs" and "HS" respectively. The middle part leaves of these Lanzhou lily plants were immediately collected and quickly put into liquid nitrogen for RNA sequencing. The total RNA of the leaf samples was extracted using a CTAB-PBIOZOL reagent (Bioflux, Hangzhou, China). RNA integrity, purity and concentration were evaluated using 1% agarose gel electrophoresis and NanoDrop 2000 (Thermo Fisher Scientific Inc., USA), and Agilent 2100 (Agilent Technologies Inc., USA). The transcriptome analysis was performed in BGI (Shenzhen, China). The mRNA libraries were constructed according to Illumina standard instructions and subsequently sequenced on a BGISEQ-500 platform, which generated raw data of 150-bp paired-end (PE150) reads. Then, *de novo* assembly based on the clean data was performed using the Trinity program. All unigene sequences were aligned to Nt nucleotide databases using Blastn. The NCBI non-redundant (Nr) database, the Swiss-Prot protein database, the eukaryotic orthologous groups/clusters of orthologous groups of proteins (KOG/COG) database, the gene ontology (GO) database, and the Kyoto encyclopedia of genes and genomes (KEGG) database were aligned by Blastx. Each unigene was functionally annotated based on the protein sharing the highest sequence similarity with the given unigene. Fragments per kilobase of the transcript per million mapped reads (FPKM) were used to estimate the quantification of gene expression levels. A corrected p-value of <0.05 and |log2foldchange| ≥1 was set as the threshold for significant differential expression. All sequencing reads were entered into the NCBI, and the sequencing data are available at accession number PRJNA934416. KEGG analysis and GO analysis using the Omic Share platform (https://www.omicshare.com/). TBtools (https://github.com/CJ-Chen/TBtools/releases) and Graphpad prism 9 (https://www.graphpad.com/) were used to make heat maps and pie charts respectively.

Genes expression analysis by qRT-PCR

Lanzhou lily plants were treated with 30 mL solution of 2 mg/L TKs or 30 mL distilled water at 22°C for 12 h, then subjected to a direct HS at 40°C at intervals of 0, 3, and 12 h. The middle parts of the leaves of the plantlets were collected and immediately frozen in liquid N_2 for RNA extraction. Total RNA was extracted from 100 mg of Lanzhou lily leaves with an RNAprep Pure Plant Kit (Tiangen, China). First-strand cDNA was synthesized using HiScript III 1st Strand cDNA Synthesis Kit (+gDNA wiper) (Vazyme, China). The expression levels of *LzDREB2B*, *LzHsfA2a*, *LzMBF1c*, *LzHsp90*, and *LzHsp70* genes in Lanzhou lily were determined by real-time quantitative RT-PCR (qRT-PCR). qRT-PCR was performed using the SYBR Green Supermix (Takara) on a Roche LightCycler 480 II (Roche, Switzerland). Gene expression data was normalized using *18S* rRNA as internal control. Primers designed for the qRT-PCR analysis are listed in Table 1.

Cloning and sequence analysis of LzHsfA2a

Total RNA isolation and first-strand cDNA synthesis were performed as described above. The specific forward primer SPF1

(5′-ATGGCAAGTGAGATGACGAAG-3′) and specific reverse primer SPR1 (5′-TTAAGTCTGGGAATCAAGATAACC-3′) were designed to amplify the coding sequence of *LzHsfA2a*. The PCR procedure was as follows:1 cycle of 5 min at 95°C; 35 cycles of 30 s at 95°C, 30 s at 54°C and 1 min 30 s at 72°C; and finally 1 cycle of 10 min at 72°C. The coding sequence of *LzHsfA2a* was obtained by sequencing and then translated into amino acids using ExPASy (http://web.expasy.org/translate/).The protein amino acid sequence was analyzed using the ExPASy Proteomics Tools online (http://au.expasy.org/tools/). Amino acid multiple alignments were performed using DNAMAN5.0 software, and a phylogenetic tree was constructed by the neighbor-joining method using MEGA7.0 software.

Statistical analysis

Data are the average of three independent experiments and shown as mean ± SD. The statistical analysis was performed by One-way ANOVA and Duncan's multiple range test using SPSS 18.0 software. A p-value < 0.05 was considered statistically significant.

Results

TKs treatment rescued Lanzhou lily seedlings from death under HS

The survival rate of Lanzhou lily plants treated with different concentrations of TKs or distilled water was assayed with a 30-d-recovery period after HS at 40°C for 72 h (Figures 1A, B). Compared to the control, application of TKs improved the survival rate. Treatment with 0.5 or 1 mg/L TKs resulted in increased survival rate, and the effect of 2 or 4 mg/L TKs was more significant, while the survival rate decreased sharply with 8 mg/L TKs. The survival rate of Lanzhou lily plantlets treated with 2 mg/L TKs was 44.4 percentage points higher than that of the control. According to the results, treatment of 2 mg/L TKs was used for the subsequent physiology and molecular biology experiments.

TABLE 1 Primers sequences used for qRT-PCR assays.

Gene	Primer sequence (5′-3′)
LzDREB2B	F: CACTGCTGTGTTGGTATTCCAGACT
	R: CACTAGCAGCATACAAGCCTAATCCT
LzHsfA2a	F: CAGACTGAGGTCGAGTTGGAAG
	R: AACACAGCCCTCTTATCTTCTCG
LzMBF1c	F: AATCAGGCGGTGCTTGCGAAGAT
	R: CGAAGACAATGAACTAAAGTCCATCC
LzHsp90	F: CTGAGGTACTTGGGAAGAAGTTCATC
	R: GTGATGTTGAGTTCAAGGCTCTACTG
LzHsp70	F: CCAGAGGCAGATTATGAAGAGGTC
	R: GGCTTAACAGTCCTCACTAAGGGTAG
18S rRNA	F: AGTTGGTGGAGCGATTTGTCT
	R: CCTGTTATTGCCTCAAACTTCC

Forward (F) primer / Reverse (R).

TKs treatment mitigated membrane injury and photosynthetic damage caused by HS

REL is the efficient indicator of cell membrane stability. TKs treatment had little effect on the changes in REL compared with non-treated control at 22°C ($p>0.05$) (Figure 2A). REL increased significantly with time during HS, indicating heat stress damaged the membrane system and caused the leakage of electrolytes from cells. After a 72-h heat treatment, a 2.93-fold increase in REL was found in the presence of 2 mg/L TKs, significantly lower than the corresponding value of the non-treated control (4.38-fold). These results showed that exogenous TKs significantly alleviated the electrolyte leakage caused by HS.

HS brings about a water deficit in the plant, and plant water status can be measured as leaf RWC. TKs and non-TKs treated plants maintained similar levels of leaf RWC under the non-stress conditions (Figure 2B). Leaf RWC of both non-TKs and TKs treated plants decreased during prolonged periods of HS compared to the non-stress control, while TKs treatment resulted in greater leaf RWC compared to non-TKs treated plants. After 72 h of HS, the leaf RWC of TKs-treated plants was 7.04 percentage points significantly higher than that of the non-TKs treated plants. The findings suggested that TKs treatment gave rise to retardation in leaf dehydration under HS.

The photosynthetic capacity of plants was positively correlated with chlorophyll content. Under the non-stress control conditions, TKs application had no significant effect on chlorophyll content (Figure 2C). HS led to decrease in chlorophyll content, especially in the later stage (72 h). However, the chlorophyll content was significantly higher in TKs-treated seedlings than in the non-treated control. P_n showed the same tendency with chlorophyll content during the HT period (Figure 2D). P_n of the TKs-treated plants decreased by 51.69% at 72 h with HS, significantly lower than the corresponding value (92.68%) of plants without TKs. These results indicated that applying TKs might compensate for the photosynthetic damages resulting from HS in Lanzhou lily.

TKs treatment enhanced ROS-scavenging ability under HS

HS causes overproduction of ROS resulting in oxidative stress. The end products of lipid peroxidation due to heat-induced oxidative stress are reactive aldehydes, one of which is MDA. No significant differences in MDA content were observed between the TKs-treated plantlets and non-treated control under normal temperature (Figure 3A). In comparison with the non-stress control, long-term HS (48-72 h) significantly increased MDA content in Lanzhou lily plants with or without TKs. MDA content of plants with TKs treatment increased by 2.46-fold after

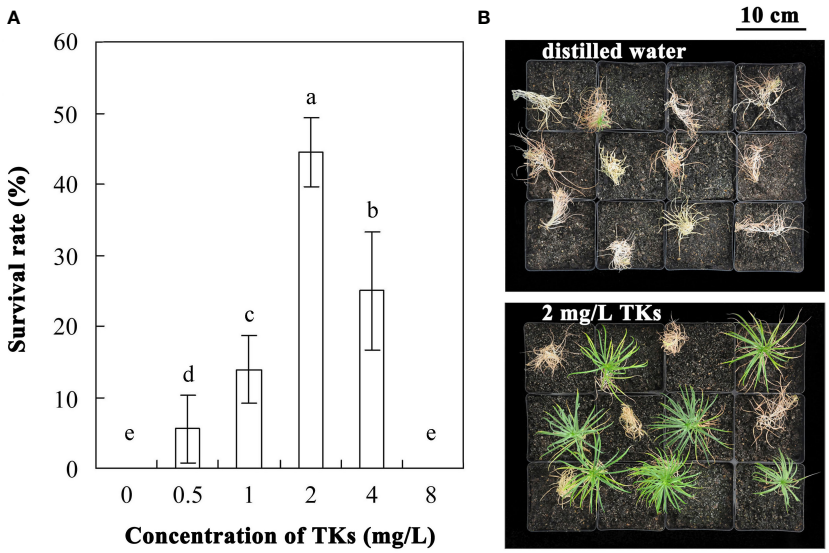


FIGURE 1
Effects of TKs on the thermotolerance of Lanzhou lily plants. **(A)** The survival rate of Lanzhou lily plants treated with different concentrations of TKs (0.5, 1, 2, 4, or 8 mg/L) or distilled water (control) after HS at 40°C for 72 h followed by a 30-d-recovery period. **(B)** Phenotype of Lanzhou lily plants treated with distilled water (upper) or 2 mg/L TKs (lower) with a 30-d-recovery period after HS at 40°C for 72 h. Each treatment included twelve plants. Data are means ± SD of three biological replications. Different letters indicate significant differences at P <0.05 (Duncan test).

HS, significantly lower than that in no-treated control (increased by 4.32-fold).

Correspondingly, plants have evolved a protective enzymatic system which consists of SOD, CAT, POD, etc., to mitigate and repair the damage initiated by ROS. TKs application increased the activities of SOD, CAT, and POD in the early stage (24 h) under non-HS conditions. Under HS conditions, treatment with TKs significantly enhanced SOD and CAT activities within 72 h, and POD activity at 48 h (Figures 3B–D). These results indicated that TKs treatment improved antioxidant enzyme (SOD, CAT, and POD) activities under HT conditions, which might be responsible for the decreased MDA content.

TKs treatment increased heat-resistance-associated hormones level under HS

Phytohormones, such as ABA, JA, SA, Gibberellin (GA), and Brassinosteroids (BRs), are well-known plant growth regulators that mediate adaptations to HT conditions (Waadt et al., 2022). In this study, HPLC-ESI-MS/MS assays were conducted to detect the change of ABA, SA, and JA content in Lanzhou lily plants treated with TKs under non-HS and HS conditions. While TKs application did not significantly changed the ABA content in the absence of HS, the HT-induced ABA level was further promoted by TKs treatment (Figure 4A). Under stress-free conditions, SA and JA content increased by 59.29% and 26.03%, respectively, at 24 h by the addition of TKs (Figures 4B, C). Compared to the non-treated control, Lanzhou lily plants supplement with TKs maintained a higher SA content during 48 h HS and JA content during 72 h HS. These results demonstrated that TKs treatment promoted the level of heat-resistance-associated phytohormones (ABA, SA, and JA) in Lanzhou lily plants exposed to HS.

TKs treatment reduced protein folding and enhanced cellular repair function under HS

Transcriptome sequencing and analysis of Lanzhou lily were performed to investigate the gene activity treated with TKs under non-HS and HS. A total of 12 cDNA libraries were prepared, and 62.9 Gb clean reads were obtained. An average of 95.59% of the sequence bases had quality scores (Q-scores) of Q20% or higher (Supplementary Table 1). To evaluate the completeness of transcriptome libraries and annotations of all the assembled unigenes, the annotated sequences corresponded to the known nucleotide sequences of plant species, with 25.61%, 19.46%, and 6.64% matching with *Elaeis guineensis*, *Phoenix dactylifera*, and *Ananas comosus*, respectively, and the remaining 48.29% matching with others (except for the top three matched species) (Supplementary Figure 1A). Through the BLASTx alignment against the public protein databases including Nt, SwissProt, KOG, KEGG, and Nr, a total of 40,466 unigenes were annotated and matched to all of the public protein databases (Supplementary Figure 1B).

By analyzing W vs. TKs under non-HS, 7266 differentially expressed genes (DEGs) [|Log2 fold-change| (|Log2 FC|) ≥ 1, p-value < 0.05] were identified by comparison, of which 3865 DEGs were down-regulated by TKs treatment, while 3401 DEGs were up-regulated (Supplementary Figure 2). During the TKs treatment and non-TKs treatment under HS, 33,387 DEGs were identified that was more than 4.5 times the number of DEGs under

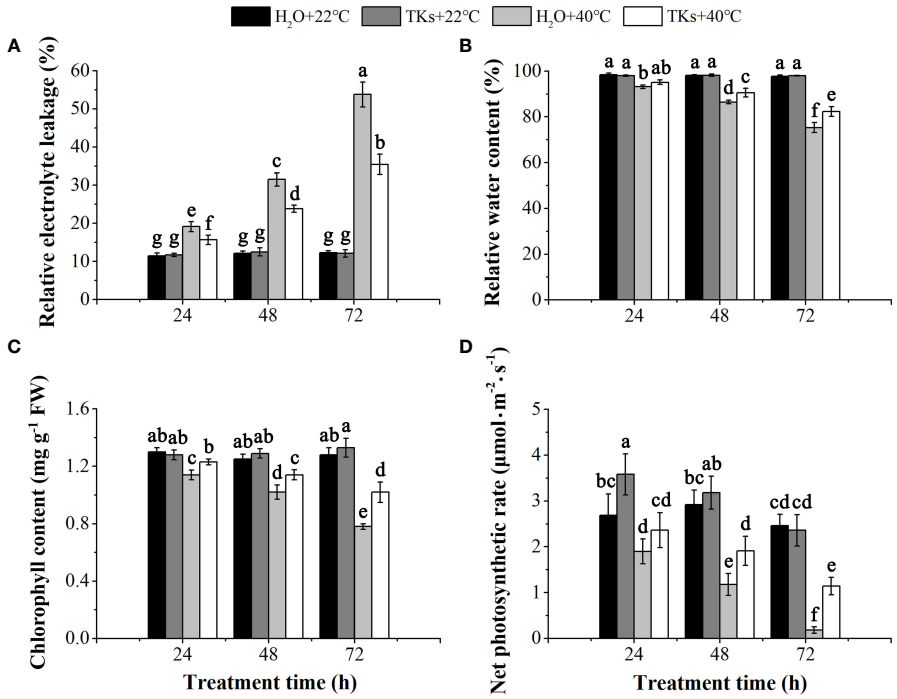


FIGURE 2
Effects of TKs on REL, RWC, *Pn*, and Chl content of Lanzhou lily plants under HS conditions (24, 48, and 72 h at 40°C) and non-stress conditions (24, 48, and 72 h at 22°C). **(A)** Relative electrolyte leakage (REL); **(B)** Relative water content (RWC); **(C)** Chlorophyll (Chl) content; **(D)** Net photosynthetic rate (*Pn*). Different lowercase letters indicate significant differences at P <0.05 (Duncan test).

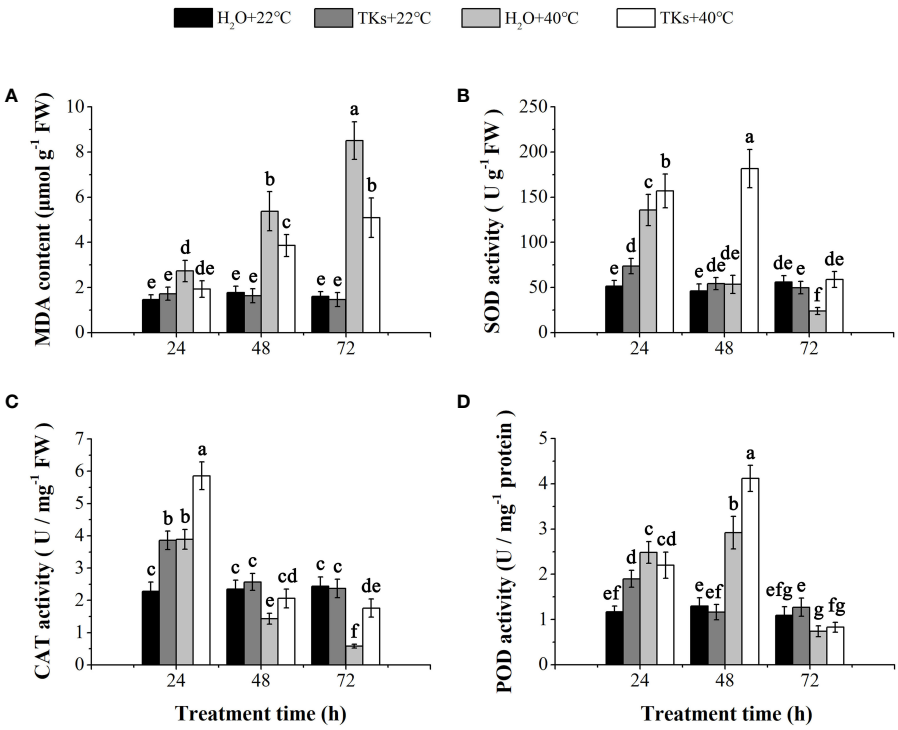


FIGURE 3
Effects of TKs on MDA content, and SOD, CAT, and POD activities of Lanzhou lily plants under HS conditions (24, 48, and 72 h at 40°C) and non-HS conditions (24, 48, and 72 h at 22°C). **(A)** MDA content; **(B)** SOD activity; **(C)** CAT activity; **(D)** POD activity. Different lowercase letters indicate significant differences at P <0.05 (Duncan test).

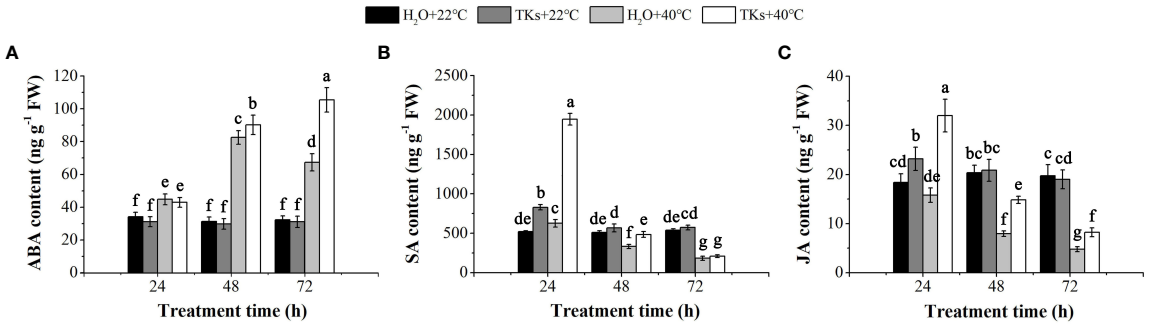


FIGURE 4
Effects of TKs on ABA **(A)**, SA **(B)** and JA **(C)** content of Lanzhou lily plants under HS conditions (24, 48, and 72 h at 40°C) and non-HS conditions (24, 48, and 72 h at 22°C). Different lowercase letters indicate significant differences at P <0.05 (Duncan test).

non-HS. Of these, 13,097 DEGs were revised downwards, while 20,290 DEGs were revised upwards (Supplementary Figure 3).

GO enrichment analysis and KEGG enrichment analysis were performed to explore the functions of DEGs responding to TKs treatment under non-HS and HS. The top 25 GO enrichment terms are shown in Figure 5. Under the non-HS conditions, the functions of DEGs were mainly concentrated in the "carbohydrate metabolic process (GO:0005975)" and "organonitrogen compound catabolic process (GO:1901565)" of Biological Process, "oxidoreductase activity (GO:0016491)", "hydrolase activity, acting on glycosyl bonds (GO:0016798)", "hydrolase activity, hydrolyzing O-glycosyl compounds (GO:0004553)" of Molecular Function, and "integral component of membrane (GO: 0016021)" of Cellular Component (Figure 5A). Compared with the GO enrichment of DEGs at non-HS, the GO enrichment of DEGs of HS vs. HS+TKs had significant changes. First, the proportion of GO_Level of "Cellular Component" had increased in the top 25 GO enrichment, especially in the "Cell (GO:0005623)" and "intracellular (GO:0005622)" of the Cellular Component. Second, in the GO_Level of Biological Process and Molecular Function, "DNA integration (GO:0015074)" and "protein folding (GO:0006457)", "protein binding (GO:0005515)" and "unfolded protein binding (GO:0051082)" showed higher p-value, respectively (Figure 5B).

In KEGG enrichment analysis under non-HS, the following terms were most enriched: "Global and overview maps" and "Carbohydrate metabolism" of Metabolism; "Folding, sorting and degradation" of Genetic Information Processing; "Signal transduction" of Environmental Information Processing; and "Environmental adaptation" of Organismal Systems (Figure 6A). Compared with KEGG enrichment of W vs. TKs under non-HS conditions, KEGG enrichment under HS did not change significantly, except for the relative increase in the proportion of "Energy metabolism" of Metabolism and "Translation" of Genetic Information Processing in the top 20 KEGG enrichment (Figure 6B).

These results suggested that TKs treatment at non-HS conditions may mainly affect the carbohydrate metabolism of Lanzhou lily plants, while TKs treatment under HS may improve heat tolerance by reducing protein folding and enhancing cellular repair function.

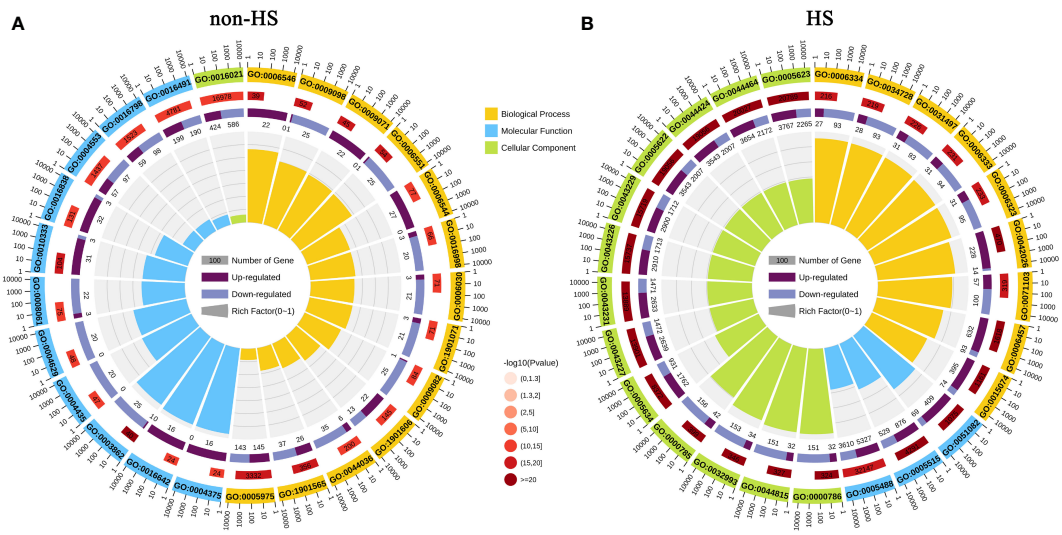


FIGURE 5
GO enrichment analysis of DEGs in response to TKs treatment under non-HS conditions **(A)** and HS conditions **(B)**.

TKs treatment affected the synthesis and signal transduction of ABA, JA and SA under HS

Hormone pathway is considered to be one of the main pathways of plant HS signal transduction. Based on our previous results (Figures 4-6), we further focused on the activity of key genes in the synthesis pathway and signaling pathway of ABA, JA, and SA. Compared to HS treatment, the genes encoding the key enzyme for ABA synthesis—*NCEDs* (*9-cis-epoxycarotenoid dioxygenases*) were significantly up-regulated by HS+TKs treatment, and the key ABA catabolism enzyme encoding genes—*CYP707As* were partially down-regulated (Figure 7A), indicating that TKs treatment could promote ABA synthesis under HS, which was consistent with the previous results (Figure 4A). The expression trends of three key genes, *PYL* (*pyrabactin resistance 1-like*), *PP2C* (*type 2C protein phosphatases*), and *SnRK2* (*SNF1-related protein kinases 2*) in the ABA signaling pathway were inconsistent, but the expression of all *ABFs*, the downstream genes of ABA signaling pathway, were up-regulated, indicating that TKs treatment might activate the signal transduction of ABA under HS.

LOXs (*lipoxygenases*) are genes encoding the key enzyme in the JA synthesis pathway. Among 27 differentially-expressed-*LOXs* under HS+TKs treatment compared to HS treatment, five were up-regulated (Figure 7B). As genes encoding key enzymes in the isochorismate synthase pathway of SA biosynthesis, the expression of *ICSs* (*isochorismate synthase*) were all significantly up-regulated, but the expression of *PALs* (*phenylalanine ammonia-lyase*), genes encoding key enzymes in the phenylalanine ammonia-lyase pathway of SA biosynthesis, were significantly down-regulated by HS+TKs treatment (Figure 7C). The up-regulation of *ICS* and *LOX* genes under HS+TKs treatment compared to HS treatment may be partly responsible for the increase in SA and JA levels (Figures 4B, C). In addition, we found that the expression of key genes in the JA signaling pathway (*JAR*, *JAZ*, and *MYC2*) and SA signaling pathway (*NPR1* and *TGA*) was consistently down-regulated at HS vs. HS+TKs, but the expression of *PR1* (*pathogenesis-related gene 1*), the gene downstream in the SA signaling pathway and involved in plant disease resistance, was significantly up-regulated.

These results suggested that TKs treatment had effects on the synthesis and signal transduction of these above phytohormones under HS, and TKs could improve the heat tolerance of Lanzhou lily mainly by promoting ABA synthesis and signal transduction.

TKs treatment had a significant impact on HSF-HSP pathway under HS

Transcription factors (TFs) are important proteins that regulate gene expression and play critical roles in plant abiotic stress responses. Among the DEGs, 194 annotated TFs responding to the TKs treatment under non-HS were identified, among which the MYB family, bHLH family, HSF family, AP2-EREBP family, NAC family and WRKY family accounted for 15.46%, 14.43%, 12.88%, 10.82%, 7.22%, and 5.67%, respectively (Figure 8A). Of the 92 annotated-*HSFs* in transcriptome data, 25 differentially-expressed-*HSFs* in response to TKs treatment were all down-regulated (Figure 8C). It was shown that TKs might act as a suppressor for *HSFs* under non-HS conditions.

We further explored TFs in response to TKs treatment at HS. Among the 33,387 DEGs, 741 differentially-expressed-TFs were

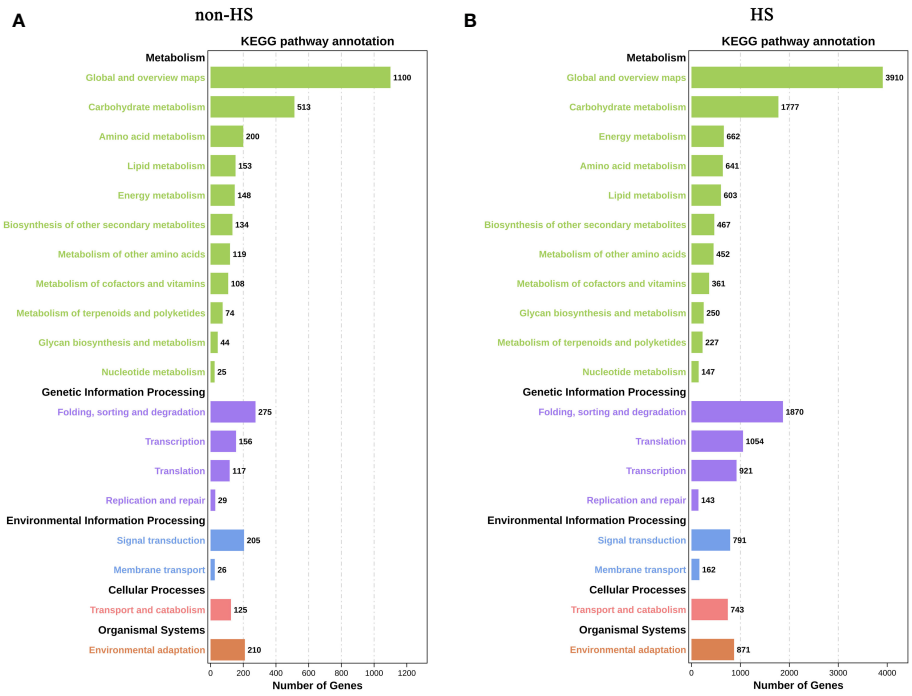


FIGURE 6
KEGG enrichment analysis of DEGs in response to TKs treatment under non-HS conditions **(A)** and HS conditions **(B)**.

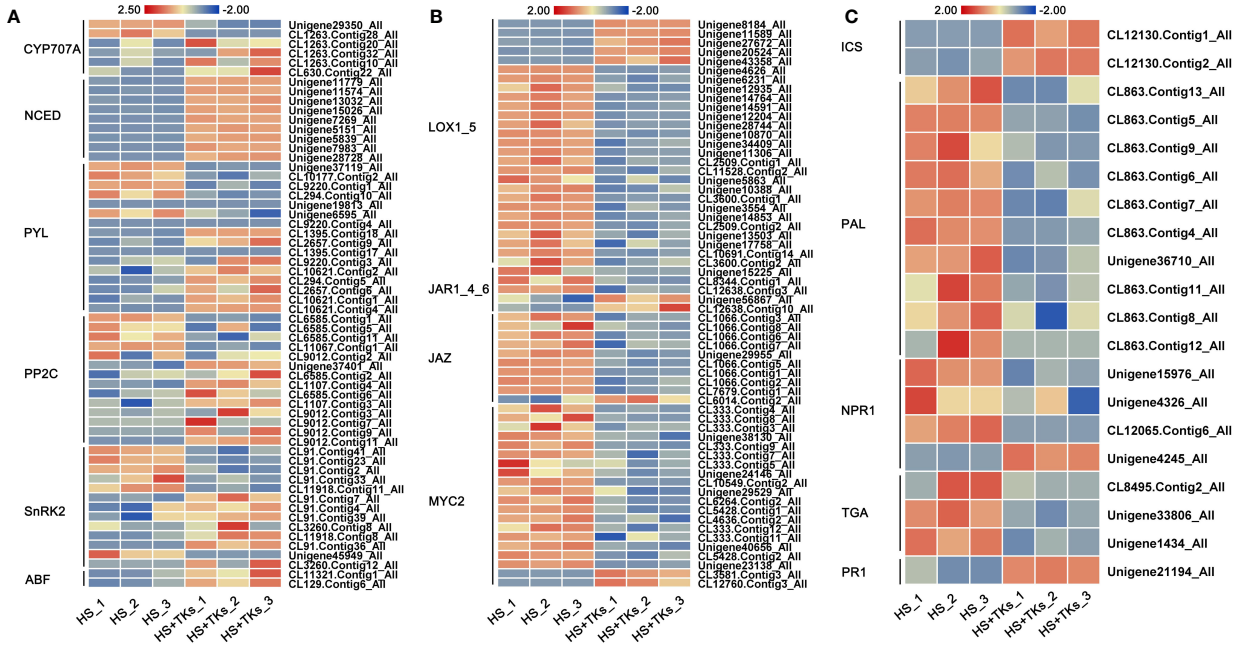


FIGURE 7
The heat map of DEGs involved in hormone synthesis and signaling pathways. Heatmaps showing the relative expression of DEGs involved in the synthesis and signaling pathways of ABA **(A)**, JA **(B)**, and SA **(C)**.

identified, mainly distributed in the families of bHLH, MYB, WRKY, NAC, HSF and C3H. Several WRKYs and HSFs were identified as crucial positive TFs in HSR of ornamental lily based on our earlier investigations (Wu et al., 2018a; Wu et al., 2019; Ding et al., 2021; Wu et al., 2022a), so we focused on the analysis of WRKY and HSF in this study. Notably, compared with the TKs treatment at non-HS, the members of the WRKY family that responded to TKs treatment increased by 50 WRKYs under HS (Figures 8A, B). In 61 differentially-expressed-*WRKY* genes responding to TKs under HS, most *WRKYs* were down-regulated including 90% of *WRKY53* and all *WRKY2* (Figure 9A). Only eight *WRKY* genes were up-regulated under HS+TKs, of which *WRKY33* accounted for 37.5%. In the HSF family, 40 *HSFs* responded to HS+TKs treatment, of which 37 *HSFs* expression were up-regulated by TKs (Figure 8D). Moreover, we noted that the number of *HSFA2* transcripts accounts for 73% of the up-regulated *HSFs*. Interestingly, 14 *HSFs* down-regulated by TKs under non-HS were significantly up-regulated by HS+TKs. These results suggested that *HSFs*, especially *HSFA2*, may play a central role in heat tolerance induced by TKs.

Among the 50 most highly induced genes following HS+TKs treatment compared with HS treatment, 22% are *HSP70/90* genes (Supplementary Table 2). HSPs serve as molecular chaperones in the HSR, which are well-known targets of HSFs. 373 differentially-expressed-*HSP* genes were identified, and about 91% of the *HSP* genes were up-regulated by HS+TKs treatment compared with HS treatment, indicating that TKs treatment also had a strong promoting effect on *HSP* genes under HS conditions (Figure 9B; Supplementary Table 3).

The analysis of DEGs provided candidate genes, such as *HSFA2*, *WRKY33*, *HSP90*, and *HSP70* for further exploration of function in TKs-elicited thermotolerance. The above findings revealed that the mechanism of HS tolerance regulated by TKs in Lanzhou lily might be mainly involved in the HSF-HSP pathway.

TKs treatment up-regulated heat-protective genes transcripts under HS

RNA-seq analysis revealed that the HSF-HSP pathway might mainly regulate the heat response networks in TKs-primed Lanzhou lily, so several heat-protective genes (*DREB*, *HSFA*, *HSP* and *MBF1c*) involved in this signal chain were selected for gene expression assay. The expression profiles of *DREB2B*, *HSFA2a*, *MBF1c*, *HSP90*, and *HSP70* genes in Lanzhou lily in response to HS and TKs treatment were quantified using qRT-PCR (Figures 10A–E). There was no significant difference in the expression level of these genes between TKs-treated and non-treated control plants under non-stressed conditions (22°C) except *LzHsfA2a* (Figure 10B). The expression of these genes in non-treated plants raised significantly after short-term (3 h) HS treatment, and then declined substantially after long-term (12 h) HS treatment with an exception of *LzHsfA2a* (Figures 10A–E). The results implied that *LzHsfA2a* might play an important role in the later stage of HSR in Lanzhou lily. TKs application had no influence on genes expression in short-term HS compared to non-treated control, but caused induction of those genes in long-term HS. The results indicated that TKs treatment could enhance the upregulation of heat-protective genes during long-term HS.

The five gene expression profiles of qRT-PCR in response to HS or HS+TKs treatment at 12 h were strongly consistent with those obtained from RNA-Seq (Figures 10A–E), suggesting that the data from the transcriptome were reliable.

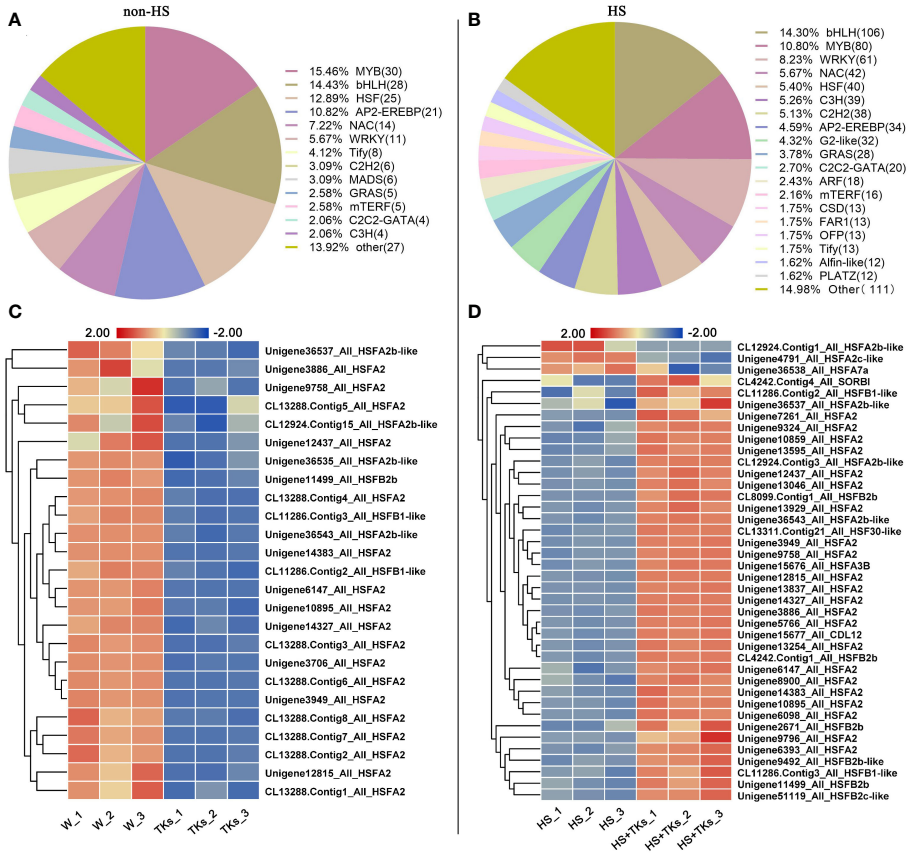


FIGURE 8
Analysis of differentially expressed transcription factors and differentially expressed *HSF* members. **(A)** Transcription factor ratio of W vs. TKs under non-HS conditions. **(B)** Transcription factor ratio of HS vs. HS+TKs. **(C, D)** The heat map of differentially-expressed-*HSFs* responding to TKs treatment at non-HS **(C)** and HS conditions **(D)**.

LzHsfA2a-1 shared high similarity to LlHsfA2 sequence

The gene expression profiles derived from RNA-seq and qRT-PCR revealed that *LzHsfA2a* in response to long-term HS with or without TKs treatment, which might contribute to the improved thermotolerance of Lanzhou lily treated with TKs, so *LzHsfA2a* was chosen as a functional candidate gene for further study. The *LzHsfA2a-1* (*LzHsfA2a*) and *LzHsfA2a-2* genes were isolated from Lanzhou lily, which contains a 1059-bp open reading frame (ORF) encoding a protein of 352 amino acid residues and a 1071-bp ORF encoding a protein of 356 amino acid residues, respectively (Figure 11A). Phylogenetic analysis was performed among LzHsfA2a-1 and LzHsfA2a-2, HSFs in lily and all the members of HSFs in *Arabidopsis* (Figure 11B). LzHsfA2a-1 (UOA68387) and LzHsfA2a-2 respectively showed 95.44% and 89.33% similarity to LlHsfA2a (ADM47610) which had properties of heat resistance in ornamental lily (Xin et al., 2010; Zhou Y, et al., 2022). Multiple alignments and sequence analysis revealed that the deduced amino acid sequence of LzHsfA2a-1 contained typical functional domains, including DNA binding domain (DBD), oligomerization domain (OD), hydrophobic heptad repeat A/B (HR-A/B), nuclear localization signal (NLS), aromatic, hydrophobic and acidic amino acid residues (AHA), and nuclear export signal (NES) (Figure 11A).

Because of the sustained high-level expression in HSR and high similarity to LlHsfA2a sequence, we speculate that LzHsfA2a-1 has the function of heat tolerance in Lanzhou lily. Moreover, the long-term HS-induced *LzHsfA2a-1* expression was up-regulated by TKs treatment, so LzHsfA2a-1 likely plays a crucial role in TKs-induced thermotolerance of Lanzhou lily.

Discussion

TKs from *Trichoderma longibrachiatum* SMF2 enhanced tolerance to HS on Lanzhou lily

HS is becoming one of the major limiting factors for food security due to global warming, so it is necessary to take effective measures to improve plant thermotolerance. Improving heat tolerance in crop plants through conventional breeding is a long and capital-intensive process, while genetic engineering is associated with ethical and social acceptance issues (Backer et al., 2018). Previous studies have suggested that application of plant protectants or inducers, such as

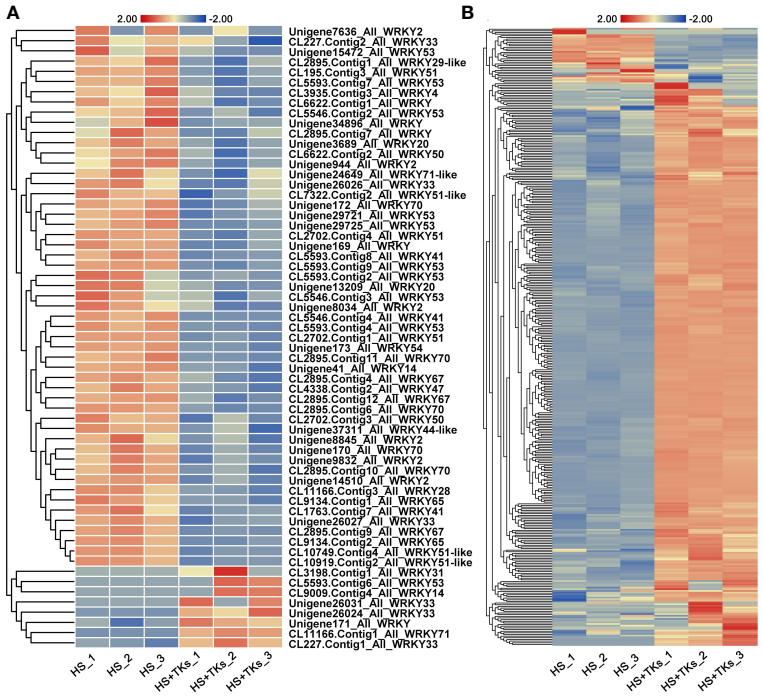


FIGURE 9
The heat map of differentially-expressed-*WRKYs* **(A)** and *HSP* family genes **(B)** responding to TKs treatment under HS conditions. The gene_ID in **(B)** was shown in Supplementary Table 3.

osmoprotectants, PGRs, signaling molecules, mineral elements, amino acids, polypeptides, or plant growth-promoting microorganisms (PGPMs), could be another way to solve this problem (Hasanuzzaman et al., 2013; Backer et al., 2018). Melatonin pretreatment has been shown to improve heat stress tolerance of *Arabidopsis* and tall fescue (Shi et al., 2015; Alam et al., 2018), while the thermotolerance of maize and lily was enhanced by exogenous Ca^{2+} (Gong et al., 1997; Cao et al., 2013). Maize seedlings treated with exogenous trehalose show a higher survival percentage under HS (Li Z, et al., 2014). SA has been implicated in heat resistance in *Arabidopsis* (Larkindale and Knight, 2002; Clarke et al., 2004), grapevine (Wang et al., 2010), rice (Lu et al., 2009), pea (Pan et al., 2006), lily (Chen et al., 2008) and grape berry (Wen et al., 2008). Exogenous application of ABA and methyl jasmonate protects rice and *Arabidopsis* from HS injury, respectively (Clarke et al., 2009; Liu et al., 2022). Pretreatment with EBL in melon alleviates HT-caused growth suppression (Zhang et al., 2013), and the similar findings were observed in tomato and *Camellia sinensis* L. (Mazorra et al., 2002; Li et al., 2016). It has been reported that PGPMs and elicitors secreted from PGPMs have an important function in plant stress tolerance, including heat tolerance (Backer et al., 2018). *Pst* DC3000-induced defense response alleviates the injury caused by subsequent HS in *Arabidopsis* (Tuang et al., 2020). GAs produced by *Gibberella fujikuroi* was reported to promote plant growth attributes of date palms under HT stress (Khan et al., 2020). *Trichoderma* spp. are well known as important biological control agents of plant diseases (Howell, 2003). Trichokonins are peptaibols produced by *Trichoderma longibrachiatum* SMF2, consisting of two major isoforms, 20-aa TKs (TKA) and 11-aa TKs (TKB) (Xie et al., 2014). They were previously found to exhibit broad-spectrum antimicrobial activity against bacteria, fungal phytopathogens and TMV. Furthermore, they induced systemic resistance against *Botrytis cinerea* infection in moth orchid and *Pcc* infection in Chinese cabbage (Li H, et al., 2014; Zhao et al., 2018). However, there have been no studies on the function of TKs on plant abiotic stresses. This study showed that TKs treatment improved the thermotolerance of Lanzhou lily plants, as evidenced by a higher survival rate under HS (Figure 1). This is the first report that exogenous application of TKs can enhance plant resistance to HS. Given its advantages of being highly effective, eco-friendly, easily obtainable, and low-cost, TKs is expected to be a commercial plant inducer in the future.

TKs treatment altered heat-resistance-associated physiology properties of Lanzhou lily to improve thermotolerance

Plants develop various defense strategies to buffer the damages imposed by HS, including morphological, physiological and biochemical responses. Leaf RWC is an important indicator of plant water status. In this study, leaf RWC showed a decreasing trend with increasing exposure time to HT, which was in agreement with results reported in hyacinth bean (Rai et al., 2015) and *Festuca arundinacea* (Xu et al., 2007). Our results revealed that TKs priming maintained a relatively higher leaf RWC under HS conditions (Figure 2B). Similar results were obtained by Li et al. (2014) who found that RWC was recovered by the application of NO during HS. HS-induced impairment of photosynthesis includes direct damage to the photosynthetic apparatus and inhibition of chlorophyll

10.3389/fpls.2023.1182977

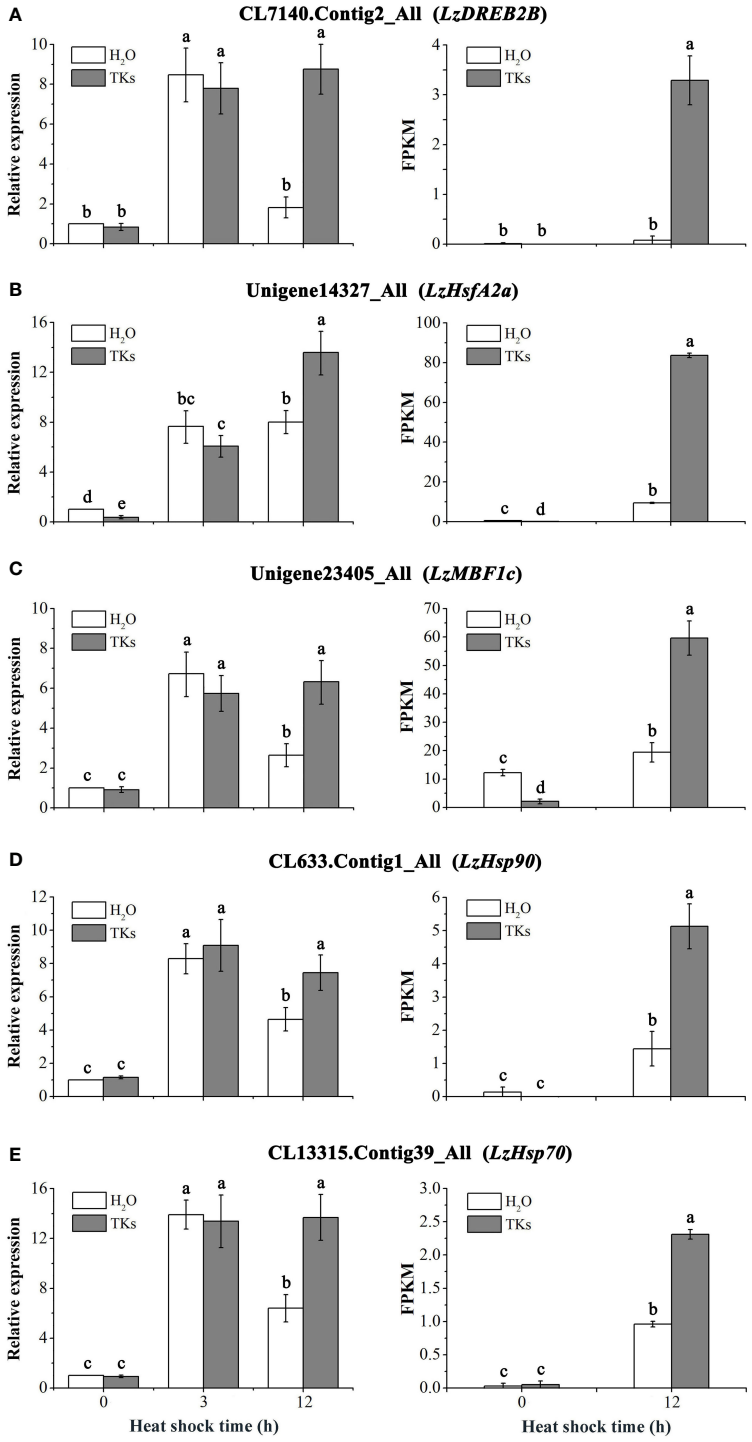


FIGURE 10
qRT-PCR analysis of HS-induced genes in Lanzhou lily leaves and validation of DEGs. Expression of *LzDREB2B* **(A)**, *LzHsfA2a* **(B)**, *LzMBF1c***(C)**, *LzHsp90* **(D)**, and *LzHsp70* **(E)** in response to distilled water or 2 mg/L TKs treatment at 40°C heat treatment for different periods (0, 3, and 12 h). Samples treated with distilled water under normal conditions (22°C) were used for normalization, with their gene expression level set to 1. The relative expression level of each gene was calculated using the $2^{-\Delta\Delta Ct}$ method after normalizing for levels of *18S rRNA* in each sample. Each treatment included three plants. Data are means ± SD of three independent experiments. Different lowercase letters indicate significant differences at P <0.05 (Duncan test).

biosynthesis. A reduction in chlorophyll content and P_n as a result of HT was observed in peas (Haldimann and Feller, 2005), which was in conformity with our findings in this research. We found that TKs treatment prevented the decrease in chlorophyll content and P_n caused by HS (Figures 2C, D), potentially resulting in enhanced thermotolerance in Lanzhou lily. The similar observation had also been made by Alam et al. (2018) in tall fescue by exogenous melatonin. ROS, such as superoxide radical (O_2^-), hydroxyl radical

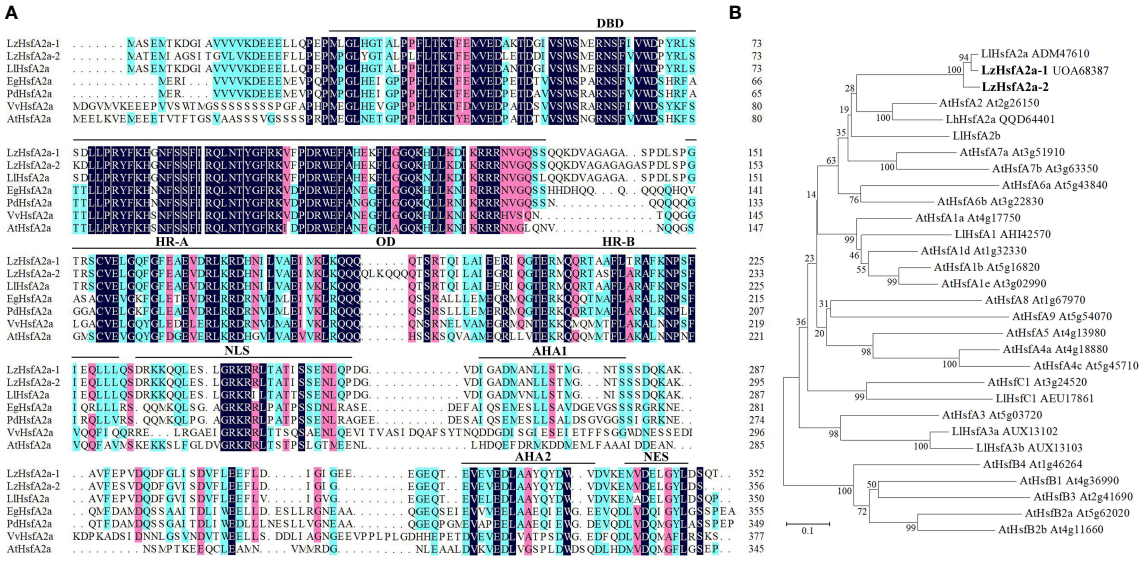


FIGURE 11
Sequences alignment and phylogenetic relationship of LzHsfA2a with its homologs from other species. **(A)** Alignment of the deduced amino acid sequences of LzHsfA2a-1, LzHsfA2a-2 and other plant HsfA2a. LzHsfA2a proteins contained conserved domains: DBD (DNA binding domain), OD (oligomerization domain), HR-A/B (hydrophobic amino acid residues), NLS (nuclear localization signal), AHA (aromatic, hydrophobic and acidic amino acid residues), and NES (nuclear export signal). **(B)** Phylogenetic relationship between LzHsfA2a-1, LzHsfA2a-2, and other plant HSF proteins. The phylogenetic tree was constructed by the neighbor joining method using MEGA 7.0 software. LzHsfA2a-1 and LzHsfA2a-2 are members of class HSFA2. Node values are bootstrap percentages based on 1000 replicates. The branch length scale bar indicates the evolutionary distance of 0.1 amino acid substitutions per site.

(OH), hydrogen peroxide (H_2O_2), and single oxygen ($^1O_2$), serve as signaling messengers required for plant HSR at low levels (Mittler, 2017). However, the injury mechanisms under HT involve the overproduction of ROS leading to lipid peroxidation and increased electrolyte leakage. Consistent with studies on tall fescue and *Lilium longiflorum* (Yin et al., 2008; Alam et al., 2018), increases in MDA content and REL were found in Lanzhou lily leaves after long-term HS (Figures 2A, 3A). To mitigate and repair the damage initiated by ROS, plants have evolved ROS-scavenging mechanism containing enzymatic systems (SOD, CAT, POD, APX, and GR) and non-enzymatic systems (AsA and GSH). As reported in ornamental lily (*Lilium longiflorum*), the enzyme activities of SOD, POD, CAT, APX and GR were stimulated and the content of AsA and GSH maintained high levels during 10 h HS at 37°C and 42°C, which resulted in low O_2^- and H_2O_2 concentrations (Yin et al., 2008). There is increasing evidence that exogenous protectants can alleviate membrane injury induced by HS through improving antioxidant defense capacity. Applying exogenous rosmarinic acid in tomatoes enhances the activities of APX, CAT, GR, and DHAR and modulates the redox status of GSH and AsA to mitigate the lipid peroxidation caused by heat-induced oxidative stress (Zhou Z, et al., 2022). Employing exogenous ABA in rice relieves cell injury, as shown by a lower accumulation of MDA and REL under HS (Liu et al., 2022). In the present investigation, TKs showed a parallel function and mechanism in regulating the tolerance of Lanzhou lily plants to HS. SOD, POD and CAT activity were further promoted by TKs treatment in Lanzhou lily plants subjected to HS (Figures 3B–D), contributing to the reduced MDA and REL accumulation. The results were in coincidence with findings in tea plants treated with AMHA (Yang et al., 2023). The phytohormones ABA, BRs, SA, JA, auxin (IAA), cytokinin (CK), and ethylene (ET) integrate HT stimuli and endogenous signals to regulate plant defensive response to HS (Li et al., 2021). Generally, ABA is recognized as a stress hormone in plants. HS elicites a rapid and transient increase in endogenous ABA levels that confer thermal tolerance by increasing ROS levels to enhance antioxidant capacity (Larkindale et al., 2005; Kaya et al., 2019). In this work, the HT-induced ABA content was further promoted by TKs treatment (Figure 4A), which was accordant to the elevated antioxidase activity. Many researches have also reported the role and mechanism of SA in protecting plants against heat-induced damage. SA application ameliorates HS-induced damage by the improvement on plant growth, antioxidants level and photosynthetic efficiency (Shah Jahan et al, 2019; Wassie et al., 2020). Although the positive role of JA in plant heat tolerance is well documented, the underlying mechanisms are not well understood. Unlike the response pattern of ABA to HS, SA and JA content displayed a sustained reduction during 48-72 HS (Figures 4B, C). However, SA and JA levels were induced by TKs under HS conditions, suggesting SA and JA were also related to TKs-induced thermotolerance acquisition of Lanzhou lily. Accordingly, TKs treatment affected the synthesis and signal transduction of these above phytohormones under HS, and TKs improved the heat tolerance of Lanzhou lily possibly by promoting ABA synthesis and signal transduction (Figure 7). Interestingly, TKs treatment increased SOD, CAT, and POD activities, and SA and JA content at early stage (24 h) under non-HS conditions, which coincides with the discovery that TKs are known as elicitor for plant induced resistance (Luo et al., 2010; Li H, et al., 2014; Zhao et al., 2018).

HSF-HSP pathway was the dominant thermal response pathway in TKs-primed Lanzhou lily and LzHsfA2a-1 probably played a central role in TKs-induced thermotolerance

At the molecular level, genes responsible for the expression of osmoprotectants, detoxifying enzymes, transporters, and regulatory proteins are induced for protection from HS (Hasanuzzaman et al., 2013). HS signal transduction networks in plant cells mainly involve HSF-HSP, Ca^{2+}-CaM, ROS, and hormone pathways, of which HSF-HSP is considered as the fundamental and master signaling pathway (Saini et al., 2022; Zhang et al., 2022). HSPs are major functional proteins induced by HS that are divided into HSP100, HSP90, HSP70, HSP60, and sHSP based on their molecular weight. HSPs function as molecular chaperones to renature a variety of proteins denatured by HS, which are associated with acquired thermotolerance in *Arabidopsis*, rice, tomato, Lanzhou lily, and other plants (Katiyar-Agarwal et al., 2003; Su and Li, 2008; Mu et al., 2013; Zhuang et al., 2020). HSFs are the terminal components of the HS signal transduction chain that can recognize and bind to HSEs (5′-nGAAnnTTCn-3′) in promoters to regulate the expression of *HSP* and other heat-responsive genes. According to the peculiarities of their flexible linkers and oligomerization domain, plant HSFs are classified into three families, HSFA, B, and C (Guo et al., 2016). In *Arabidopsis*, HSFA1s are identified as the master regulators in the early phase of HSR, while HSFA2 serves as core regulators in the late phase of HSR (Schramm et al., 2006; Yoshida et al., 2011). Upon HS, the transcript levels of *HSFA2* and *A6* members become the dominant *HSFs* in wheat (Xue et al., 2014). The natural variations of HSFA2 enhance thermotolerance in grapevine (Liu et al., 2023). HS-responsive transcription factors DREB2A and DREB2C are identified as acting upstream of HSFA3 for the establishment of thermotolerance (Schramm et al., 2008; Chen et al., 2010), while *MBF1c* closely involved in plant HSR is activated by HSFA1b, HSFA2, and DREB2A (Sakuma et al., 2006; Ogawa et al., 2007; Bechtold et al., 2013; Qin et al., 2015; Liu et al., 2023). Plant-specific WRKY TFs also function in plant HSR. AtWRKY25, AtWRKY26, and AtWRKY33 positively regulate the cooperation between the ET-activated and HSPs-related signaling pathways that mediate responses to HS (Li et al., 2011). HSFA1, HSFA2a, HSFA2b, HSFA3a, HSFA3b, HSFA4, DREB2B, MYB305, WRKY22, WRKY39, MBF1c, NAC014, HB16, and ERF110 have been identified from ornamental lily, and all of these TFs are positively correlated with lily thermotolerance (Xin et al., 2010; Gong et al., 2014; Wu et al., 2018a; Wu et al., 2018b; Wu et al., 2019; Ding et al., 2021; Wang C, et al., 2022; Wang Y, et al., 2022; Wu et al., 2021; Wu et al., 2022a; Wu et al., 2022b; Xin et al., 2017; Zhou Y, et al., 2022; Wu et al., 2023). However, only heat-inducible *HSP16.45* has been reported in edible Lanzhou lily (Mu et al., 2013).

To elucidate the molecular mechanism of TKs enhancing the thermotolerance of Lanzhou lily, RNA-seq and data analysis were conducted in this study. GO and KEGG enrichment analyses revealed that TKs treatment at non-HS conditions might mainly affect the carbohydrate metabolism of Lanzhou lily plants, while TKs treatment at HS conditions might improve the heat tolerance by reducing protein folding and enhancing cellular repair function (Figures 5, 6). The analysis of DEGs revealed that the mechanism of HS tolerance regulated by TKs in Lanzhou lily might be mainly involved in the HSF-HSP pathway, in which *HSFA2* might be a central regulator (Figures 8, 9). The expression patterns of *LzDREB2B*, *LzHsfA2a*, *LzMBF1c*, *LzHsp90*, and *LzHsp70* genes involved in HSF-HSP pathway in response to HS and TKs treatment were studied using qRT-PCR (Figure 10). The results indicated that the thermal-induced *LzDREB2B*, *LzHsfA2a*, *LzMBF1c*, *LzHsp90*, and *LzHsp70* transcripts were up-regulated by TKs treatment after long-term (12 h) HS, which validated the accuracy of the RNA-Seq data and was probably responsible for the improved thermotolerance of Lanzhou lily. We also discovered that *LzHsfA2a* activity positively and lastingly responded to HS, so *LzHsfA2a* was chosen as potential functional gene for further study. Subsequently, *LzHsfA2a-1* and *LzHsfA2a-2* genes were isolated from Lanzhou lily (Figure 11), and *LzHsfA2a-1* (*LzHsfA2a*) encodes a protein of 352 amino acid residues that shows 95.44% similarity to LlHsfA2a which had properties of heat resistance in *Lilium longiflorum* (Xin et al., 2010; Zhou Y, et al., 2022). As the result of the lasting response to HS, the elevated response to HS coupled with TKs treatment, and the high similarity to LlHsfA2a sequence, it can be speculated that LzHsfA2a-1 probably plays a vital role in regulation of TKs-induced thermotolerance of Lanzhou lily. Future studies to confirm the molecular function of *LzHsfA2a-1* and other candidate genes, such as *LzWRKY33*, and *LzHSPs*, will provide new data for understanding the TKs-regulated mechanism of HSR in Lanzhou lily.

Conclusion

In summary, TKs application enhanced thermotolerance in Lanzhou lily plants *via* promoting photosynthetic capacity, water retention ability, antioxidant enzyme system activity, HS-related phytohormones level, and HS-protective genes transcription regulation. Therefore, this study not only provides a direct link between TKs and plant thermotolerance, but also supplies a promising strategy that can alleviate plant injury resulted from HS.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here https://www.ncbi.nlm.nih.gov/bioproject/PRJNA934416.

Author contributions

XC and JL contributed to the conceptualization of the study and writing of the original draft. XC, JS, DH, JL and ZW contributed to the methodology. HL, WY and TL performed data curation and graphing. DH contributed to the material preparation. XC, JS, HL, WY, TL, DH, JL and ZW contributed to the writing, reviewing, and editing. JL supervised the work. XC and JL contributed to funding acquisition. All the authors contributed to the manuscript revision and approved the submitted version.

Funding

This work was supported by Shandong Provincial Natural Science Foundation, China (ZR2022MC048), the National Natural Science Foundation of China (31601788) and Scientific and Technological Innovation Capacity Building Project of Beijing Academy of Agriculture and Forestry Sciences (KJCX20230801).

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1182977/full#supplementary-material

References

Alam, M. N., Zhang, L. H., Yang, L., Islam, M. R., Liu, Y., Luo, H., et al. (2018). Transcriptomic profiling of tall fescue in response to heat stress and improved thermotolerance by melatonin and 24-epibrassinolide. *BMC Genomics* 19, 224. doi: 10.1186/s12864-018-4588-y

Backer, R., Rokem, J. S., Ilangumaran, G., Lamont, J., Praslickova, D., Ricci, E., et al. (2018). Plant growth-promoting rhizobacteria: context, mechanisms of action, and roadmapto commercialization of biostimulantsfor sustainable agriculture. *Front. Plant Sci.* 9, 1473. doi: 10.3389/fpls.2018.01473

Bechtold, U., Albihlal, W. S., Lawson, T., Fryer, M. J., Sparrow, P. A. C., Richard, F., et al. (2013). Arabidopsis HEAT SHOCK TRANSCRIPTION FACTORA1b overexpression enhances water productivity, resistance to drought, and infection. *J. Exp. Bot.* 64, 3467–3481. doi: 10.1093/jxb/ert185

Bokszczanin, K. L., and Fragkostefanakis, S. (2013). Perspectives on deciphering mechanisms underlying plant heat stress response and thermotolerance. *Front. Plant Sci.* 4, 315. doi: 10.3389/fpls.2013.00315

Cai, W. W., Yang, S., Wu, R. J., Cao, J. S., Shen, L., Guan, D. Y., et al. (2021). Pepper NAC-type transcription factor NAC2c balances the trade-off between growth and defense responses. *Plant Physiol.* 186, 2169–2189. doi: 10.1093/plphys/kiab190

Cao, X., Yi, J., Wu, Z., Luo, X., Zhong, X. H., Wu, J., et al. (2013). Involvement of Ca^{2+} and CaM3 in regulation of thermotolerance in lily (*Lilium longiflorum*). *Plant Mol. Biol. Rep.* 31, 1293–1304. doi: 10.1007/s11105-013-0587-y

Cao, X., Zhang, M. S., Zhang, X. S., Yu, S. C., Zhao, P. B., Hou, D., et al. (2020). First report of *Fusarium redolens* causing root and bulb rot disease on lanzhou lily (*Lilium davidii* var. *unicolor*) in China. *Plant Dis.* 104, 583. doi: 10.1094/PDIS-03-19-0597-PDN

Chen, H., Hwang, J. E., Chan, J. L., Kim, D. Y., Sang, Y. L., and Lim, C. O. (2010). *Arabidopsis* DREB2C functions as a transcriptional activator of *HsfA3* during the heat stress response. *Biochem. Bioph. Res. Co.* 401, 238–244. doi: 10.1016/j.bbrc.2010.09.038

Chen, Q. M., Yin, H., Li, X. Y., and Yi, M. F. (2008). Effects of salicylic acid on the activities of antioxidant systems in lily plants under high temperature stress. *J. China Agric. Univ.* 13, 44–48.

Clarke, S. M., Cristescu, S. M., Miersch, O., Harren, F. J., Wasternack, C., and Mur, L. A. (2009). Jasmonates act with salicylic acid to confer basal thermotolerance in *Arabidopsis thaliana*. *New Phytol.* 182, 175–187. doi: 10.1111/j.1469-8137.2008.02735.x

Clarke, S. M., Mur, L. A. J., Wood, J. E., and Scott, I. M. (2004). Salicylic acid dependent signaling promotes basal thermotolerance but is not essential for acquired thermotolerance in *Arabidopsis thaliana*. *Plant J.* 38, 432–447. doi: 10.1111/j.1365-313X.2004.02054.x

Cui, G. F., Wang, X. N., Jiang, Y. L., Wu, L. F., Jia, W. J., Duan, Q., et al. (2014). OT Lily cultivar 'Jinmen'. *Acta Hortic. Sin.* 41, 1517–1518.

Ding, L. P., Wu, Z., Teng, R. D., S.J., Xu, Cao, X., Yuan, G. Z., et al. (2021). LlWRKY39 is involved in thermotolerance by activating *LlMBF1c* and interacting with LlCaM3 in lily (*lilium longiflorum*). *Hortic. Res.* 8, 36. doi: 10.1038/s41438-021-00473-7

Frank, W., Ratnadewi, D., and Reski, R. (2005). Physcomitrella patens is highly tolerant against drought, salt and osmotic stress. *Planta.* 220, 384–394. doi: 10.1007/s00425-004-1351-1

Friedrich, T., Oberkofler, V., Trindade, I., Altmann, S., Brzezinka, K., Lämke, J., et al. (2021). Heteromeric HSFA2/HSFA3 complexes drive transcriptional memory after heat stress in *Arabidopsis*. nat. *Commun.* 12, 3426. doi: 10.1038/s41467-021-23786-6

Fuller, D. J., and Gerner, E. W. (1982). Polyamines: a dual role in the modulation of cellular sensitivity to heat. *Radiat. Res.* 92, 439–444. doi: 10.2307/3575915

Gong, M., Li, Y. J., Dai, X., Tian, M., and Li, Z. G. (1997). Involvement of calcium and calmodulin in the acquisition of heat-shock induced thermotolerance in maize seedlings. *J. Plant Physiol.* 150, 615–621. doi: 10.1016/S0176-1617(97)80328-8

Gong, B. H., Yi, J., Wu, J., Sui, J. J., Khan, M. A., Wu, Z., et al. (2014). LlHSFA1, a novel heat stress transcription factor in lily (*Lilium longiflorum*), can interact with LlHSFA2 and enhance the thermotolerance of transgenic *Arabidopsis thaliana*. *Plant Cell Rep.* 33, 1519–1533. doi: 10.1007/s00299-014-1635-2

Guo, M., Liu, J. H., Ma, X., Luo, D. X., Gong, Z. H., and Lu, M. H. (2016). The plant heat stress transcription factors (HSFs): structure, regulation, and function in response to abiotic stresses. *Front. Plant Sci.* 7, 114. doi: 10.3389/fpls.2016.00114

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 Cell Rep.* 40, 2247–2271. doi: 10.1007/s00299-021-02696-3

Haldimann, P., and Feller, U. (2005). Growth at moderately elevated temperature alters the physiological response of the photosynthetic apparatus to heat stress in pea (*Pisum sativum* l.) leaves. *Plant Cell Environ.* 28, 302–317. doi: 10.1111/j.1365-3040.2005.01289.x

Hasanuzzaman, M., Nahar, K., Alam, M., Roychowdhury, R., and Fujita, M. (2013). Physiological, biochemical, and molecular mechanisms of heat stress tolerance in plants. *Int. J. Mol. Sci.* 14, 9643–9684. doi: 10.3390/ijms14059643

Howell, C. R. (2003). Mechanisms employed by *Trichoderma* species in the biological control of plant diseases: the history and evolution of current concepts. *Plant Dis.* 87, 4–10. doi: 10.1094/PDIS.2003.87.1.4

Hussain, I., Parveen, A., Rasheed, R., Ashraf, M. A., Ibrahim, M., Riaz, S., et al. (2019). Exogenous silicon modulates growth, physio-chemicals and antioxidants in barley (*Hordeum vulgare* l.) exposed to different temperature regimes. *Silicon.* 11, 2753–2762. doi: 10.1007/s12633-019-0067-6

Islam, M. R., Feng, B., Chen, T., Fu, W. M., Zhang, C. X., Tao, L. X., et al. (2019). Abscisic acid prevents pollen abortion under high temperature stress by mediating sugar metabolism in rice spikelets. *Physiol. Plantarum.* 165, 644–663. doi: 10.1111/ppl.12759

Katiyar-Agarwal, S., Agarwal, M., and Grover, A. (2003). Heat-tolerant basmati rice engineered by over-expression of hsp101. *Plant Mol. Biol.* 51, 677–686. doi: 10.1023/A:1022561926676

Kaushal, N., Gupta, K., Bhandhari, K., Kumar, S., Thakur, P., and Nayyar, H. (2011). Proline induces heat tolerance in chickpea (*Cicer arietinum* l.) plants by protecting vital enzymes of carbon and antioxidative metabolism. *Physiol. Mol. Biol. Pla.* 17, 203–213. doi: 10.1007/s12298-011-0078-2

Kaya, H., Takeda, S., Kobayashi, M. J., Kimura, S., Iizuka, A., and Imai, A. (2019). Comparative analysis of the reactive oxygen species-producing enzymatic activity of *Arabidopsis* NADPH oxidases. *Plant J.* 98, 291–300. doi: 10.1111/tpj.14212

Khan, A., Bilal, S., Khan, A. L., Imran, M., Shahzad, R., Al-Harrasi, A., et al. (2020). Silicon and gibberellins: synergistic function in harnessing aba signaling and heat stress tolerance in date palm (*Phoenix dactylifera* l.). *Plants*. 9, 620. doi: 10.3390/plants9050620

Koini, M. A., Alvey, L., Allen, T., Tilley, C. A., Harberd, N. P., Whitelam, G. C., et al. (2009). High temperature-mediated adaptations in plant architecture require the bhlh transcription factor PIF4. *Curr. Biol.* 19, 408–413. doi: 10.1016/j.cub.2009.01.046

Kumar, R. R., Singh, G. P., Sharma, S. K., Singh, K., Goswami, S., and Rai, R. D. (2012). Molecular cloning of *HSP17* gene (*sHSP*) and their differential expression under exogenous putrescine and heat shock in wheat (*triticum aestivum*). *Afr. J. Biotechnol.* 11, 16800–16808. doi: 10.5897/AJB12.2855

Lafitte, R. (2002). Relationship between leaf relative water content during reproductive stage water deficit and grain formation in rice. *Field Crop Res.* 76, 165–174. doi: 10.1016/S0378-4290(02)00037-0

Larkindale, J., Hall, J. D., Knight, M. R., and Vierling, E. (2005). Heat stress phenotypes of *Arabidopsis* mutants implicate multiple signaling pathways in the acquisition of thermotolerance. *Plant Physiol.* 138, 882–897. doi: 10.1104/pp.105.062257

Larkindale, J., and Knight, M. R. (2002). Protection against heat stress-induced oxidative damage in *Arabidopsis* involves calcium, abscisic acid, ethylene, and salicylic acid. *Plant Physiol.* 128, 682–695. doi: 10.1104/pp.010320

Li, N., Euring, D., Cha, J. Y., Lin, Z., Lu, M. Z., Huang, L. J., et al. (2021). Plant hormone-mediated regulation of heat tolerance in response to global climate change. *Front. Plant Sci.* 11, 627969. doi: 10.3389/fpls.2020.627969

Li, S. J., Fu, Q. T., Chen, L. G., Huang, W. D., and Yu, D. Q. (2011). *Arabidopsis thaliana* WRKY25, WRKY26, and WRKY33 coordinate induction of plant thermotolerance. *Planta.* 233, 1237–1252. doi: 10.1007/s00425-011-1375-2

Li, X., Gong, B., Wang, Y., and Xu, K. (2014). Heat stress mitigation by exogenous nitric oxide application involves polyamine metabolism and PSII physiological strategies in ginger leaves. *Sci. Agric. Sin.* 47, 1171–1179.

Li, Z. X., Li, X., and Han, W. Y. (2016). Physiology mechanism of exogenous 24-epibrassinolide-induced heat resistance in tea plants (*Camellia sinensis* l.). *Acta Agric. Zhejiangensis.* 28, 959–965.

Li, H. Y., Luo, Y., Zhang, X. S., Shi, W. L., Gong, Z. T., Shi, M., et al. (2014). Trichokonins from *Trichoderma pseudokoningii* SMF2 induce resistance against gram-negative *Pectobacterium carotovorum* subsp. *carotovorum* in Chinese cabbage. *FEMS Microbiol. Lett.* 354, 75–82. doi: 10.1111/1574-6968.12427

Li, Z. G., Luo, L. J., and Zhu, L. P. (2014). Involvement of trehalose in hydrogen sulfide donor sodium hydrosulfide-induced the acquisition of heat tolerance in maize (*Zea mays* l.) seedlings. *Bot. Stud.* 55, 20. doi: 10.1186/1999-3110-55-20

Li, Z. X., Tang, J., Srivastava, R., Bassham, D. C., and Howell, S. H. (2020). The transcription factor bzip60 links the unfolded protein response to the heat stress response in maize. *Plant Cell.* 32, 3559–3575. doi: 10.1105/tpc.20.00260

Liu, X. N., Chen, H. Y., Li, S. C., Lecourieux, D., Duan, W., Fan, P. G., et al. (2023). Natural variations of HSFA2 enhance thermotolerance in grapevine. *Hortic. Res.* 10, uhac250. doi: 10.1093/hr/uhac250

Liu, J. F., Ding, J., Yuan, B. F., and Feng, Y. Q. (2014). Magnetic solid phase extraction coupled with *in situ* derivatization for the highly sensitive determination of acidic phytohormones in rice leaves by UPLC-MS/MS. *Analyst.* 139, 5605–5613. doi: 10.1039/C4AN01186D

Liu, J. Z., Feng, L. L., Gu, X. T., Deng, X., Qiu, Q., Li, Q., et al. (2019). An H3K27me3 demethylase-HSFA2 regulatory loop orchestrates transgenerational thermomemory in *Arabidopsis*. *Cell Res.* 29, 379–390. doi: 10.1038/s41422-019-0145-8

Liu, X. L., Ji, P., Yang, H. T., Jiang, C. J., Liang, Z. W., Chen, Q. Z., et al. (2022). Priming effect of exogenous ABA on heat stress tolerance in rice seedlings is associated with the upregulation of antioxidative defense capability and heat shock-related genes. *Plant Growth Regul.* 98, 23–28. doi: 10.1007/s10725-022-00828-7

Lu, J., Zhang, R., Zong, X. F., Wang, S. G., and He, G. H. (2009). Effect of salicylic acid on heat resistance of rice seedling under heat stress. *Chin. J. Eco-Agric.* 17, 1168–1171. doi: 10.3724/SP.J.1011.2009.01168

Luo, Y., Gao, Y. M., Wang, W., and Zou, C. J. (2014). Application of trehalose ameliorates heat stress and promotes recovery of winter wheat seedlings. *Biol. Plantarum.* 58, 395–398. doi: 10.1007/s10535-014-0397-6

Luo, Y., Zhang, D. D., Dong, X. W., Zhao, P. B., Chen, L. L., Song, X. Y., et al. (2010). Antimicrobial peptaibols induce defense responses and systemic resistance in tobacco against tobacco mosaic virus. *FEMS Microbiol. Lett.* 313, 120–126. doi: 10.1111/j.1574-6968.2010.02135.x

Lutts, S., Kinet, J. M., and Bouharmont, J. (1996). NaCl-Induced senescence in leaves of rice (*Oryza sativa* l.) cultivars differing in salinity resistance. *Ann. Bot-London.* 78, 389–398. doi: 10.1006/anbo.1996.0134

Malerba, M., and Cerana, R. (2018). Effect of selenium on the responses induced by heat stress in plant cell cultures. *Plants.* 7, 64. doi: 10.3390/plants7030064

Mazorra, L. M., Núnez, M., Hechavarria, M., Coll, F., and Sánchez-Blanco, M. J. (2002). Influence of brassinosteroids on antioxidant enzymes activity in tomato under different temperatures. *Biol. Plantarum.* 45, 593–596. doi: 10.1023/A:1022390917656

Mittler, R. (2017). ROS are good. *Trends Plant Sci.* 22, 11–19. doi: 10.1016/j.tplants.2016.08.002

Mu, C. J., Zhang, S. J., Yu, G. Z., Chen, N., Li, X. F., and Liu, H. (2013). Overexpression of small heat shock protein LimHSP16.45 in *Arabidopsis* enhances tolerance to abiotic stresses. *PloS One* 8, e82264. doi: 10.1371/journal.pone.0082264

Ogawa, D., Yamaguchi, K., and Nishiuchi, T. (2007). High-level overexpression of the *Arabidopsis HsfA2* gene confers not only increased thermotolerance but also salt/osmotic stress tolerance and enhanced callus growth. *J. Exp. Bot.* 58, 3373–3383. doi: 10.1093/jxb/erm184

Pan, Q. H., Zhan, J. C., Liu, H. T., Zhang, J. H., Chen, J. Y., and Wen, P. F. (2006). Salicylic acid synthesized by benzoic acid 2-hydroxylase participates in the development of thermotolerance in pea plants. *Plant Sci.* 171, 226–233. doi: 10.1016/j.plantsci.2006.03.012

Qin, D. D., Wang, F., Geng, X. L., Zhang, L. Y., Yao, Y. Y., Ni, Z. F., et al. (2015). Overexpression of heat stress-responsive *TaMBF1c*, a wheat (*Triticum aestivum* l.) multiprotein bridging factor, confers heat tolerance in both yeast and rice. *Plant Mol. Biol.* 87, 31–45. doi: 10.1007/s11103-014-0259-9

Rai, N., Rai, K. K., Tiwari, G., and Singh, P. K. (2015). Changes in free radical generation, metabolites and antioxidant defense machinery in hyacinth bean (*Lablab purpureus*. l) in response to high temperature stress. *Acta Physiol. Plant* 37, 46. doi: 10.1007/s11738-015-1791-1

Saini, N., Nikalje, G. C., Zargar, S. M., and Suprasanna, P. (2022). Molecular insights into sensing, regulation and improving of heat tolerance in plants. *Plant Cell Rep.* 41, 799–813. doi: 10.1007/s00299-021-02793-3

Sakuma, Y., Maruyama, K., Osakabe, Y., Qin, F., Seki, M., Shinozaki, K., et al. (2006). Functional analysis of an *Arabidopsis* transcription factor, DREB2A, involved in drought-responsive gene expression. *Plant Cell.* 18, 1292–1309. doi: 10.1105/tpc.105.035881

Scalabrin, E., Radaelli, M., and Capodaglio, G. (2016). Simultaneous determination of shikimic acid, salicylic acid and jasmonic acid in wild and transgenic *Nicotiana langsdorffii* plants exposed to abiotic stresses. *Plant Physiol. Bioch.* 103, 53–60. doi: 10.1016/j.plaphy.2016.02.040

Schramm, F., Ganguli, A., Kiehlmann, E., Englich, G., Walch, D., and Koskull-Döring, P. V. (2006). The heat stress transcription factor HsfA2 serves as a regulatory amplifier of a subset of genes in the heat stress response in arabidopsis. *Plant Mol. Biol.* 60, 759–772. doi: 10.1007/s11103-005-5750-x

Schramm, F., Larkindale, J., Kiehlmann, E., Ganguli, A., Englich, G., Vierling, E., et al. (2008). A cascade of transcription factor DREB2A and heat stress transcription factor HsfA3 regulates the heat stress response of arabidopsis. *Plant J.* 53, 264–274. doi: 10.1111/j.1365-313X.2007.03334.x

Shah Jahan, M., Wang, Y., Shu, S., Zhong, M., Chen, Z., Wu, J. Q., et al. (2019). Exogenous salicylic acid increases the heat tolerance in tomato (*Solanum lycopersicum* l) by enhancing photosynthesis efficiency and improving antioxidant defense system through scavenging of reactive oxygen species. *Sci. Hortic.* 247, 421–429. doi: 10.1016/j.scienta.2018.12.047

Shekhawat, K., Almeida-Trapp, M., García-Ramírez, G. X., and Hirt, H. (2022). Beat the heat: plant- and microbe-mediated strategies for crop thermotolerance. *Trends Plant Sci.* 27, 802–813. doi: 10.1016/j.tplants.2022.02.008

Shi, M., Chen, L., Wang, X. W., Zhang, T., Zhao, P. B., Song, X. Y., et al. (2012). Antimicrobial peptaibols from *Trichoderma pseudokoningii* induce programmed cell death in plant fungal pathogens. *Microbiology.* 158, 166–175. doi: 10.1099/mic.0.052670-0

Shi, H., Reiter, R. J., Tan, D. X., and Chan, Z. L. (2015). Indole-3-acetic acid inducible 17 positively modulates natural leaf senescence through melatonin-mediated pathway in *Arabidopsis*. *J. Pineal Res.* 58, 26–33. doi: 10.1111/jpi.12188

Song, X. Y., Shen, Q. T., Xie, S. T., Chen, X. L., Sun, C. Y., and Zhang, Y. Z. (2006). Broad-spectrum antimicrobial activity and high stability of trichokonins from *Trichoderma koningii* SMF2 against plant pathogens. *FEMS Microbiol. Lett.* 260, 119–125. doi: 10.1111/j.1574-6968.2006.00316.x

Su, P. H., and Li, H. M. (2008). Arabidopsis stromal 70-kD heat shock proteins are essential for plant development and important for thermotolerance of germinating seeds. *Plant Physiol.* 146, 1231–1241. doi: 10.1104/pp.107.114496

Tian, X. J., Qin, Z., Zhao, Y., Wen, J. J., Lan, T. Y., Zhang, L. Y., et al. (2022). Stress granule-associated TaMBF1c confers thermotolerance through regulating specific mRNA translation in wheat (*Triticum aestivum*). *New Phytol.* 233, 1719–1731. doi: 10.1111/nph.17865

Tuang, Z. K., Wu, Z. J., Jin, Y., Wang, Y. Z., Oo, P. P. Z., Zuo, G. X., et al. (2020). Pst DC3000 infection alleviates subsequent freezing and heat injury to host plants *via* a salicylic acid-dependent pathway in *Arabidopsis*. *Plant Cell Environ.* 43, 801–817. doi: 10.1111/pce.13705

Waadt, R., Seller, C. A., Hsu, P. K., Takahashi, Y., Munemasa, S., and Schroeder, J. I. (2022). Plant hormone regulation of abiotic stress responses. *Nat. Rev. Mol. Cell Biol.* 23, 680–694. doi: 10.1038/s41580-022-00479-6

Wahid, A., and Shabbir, A. (2005). Induction of heat stress tolerance in barley seedlings by pre-sowing seed treatment with glycinebetaine. *Plant Growth Regul.* 46, 133–141. doi: 10.1007/s10725-005-8379-5

Wang, L. J., Fan, L., Loescher, W., Duan, W., Liu, G. J., Cheng, J. S., et al. (2010). Salicylic acid alleviates decreases in photosynthesis under heat stress and accelerates recovery in grapevine leaves. *BMC Plant Biol.* 10, 34. doi: 10.1186/1471-2229-10-34

Wang, L. J., and Li, S. H. (2006). Salicylic acid-induced heat or cold tolerance in relation to Ca^{2+} homeostasis and antioxidant systems in young grape plants. *Plant Sci.* 170, 685–694. doi: 10.1016/j.plantsci.2005.09.005

Wang, Y., Zhou, Y. Z., Wang, R., Xu, F. X., Tong, S., Song, C. X., et al. (2022). Ethylene response factor LlERF110 mediates heat stress response *via* regulation of *LlHsfA3A* expression and interaction with LlHsfA2 in lilies (*Lilium longiflorum*). *Int. J. Mol. Sci.* 23, 16135. doi: 10.3390/ijms232416135

Wang, C. P., Zhou, Y. Z., Yang, X., Zhang, B., Xu, F. X., Wang, Y., et al. (2022). The heat stress transcription factor LlHsfA4 enhanced basic thermotolerance through regulating ROS metabolism in lilies (*Lilium longiflorum*). *Int. J. Mol. Sci.* 23, 572. doi: 10.3390/ijms23010572

Waraich, E. A., Ahmad, R., Halim, A., and Aziz, T. (2012). Alleviation of temperature stress by nutrient management in crop plants: a review. *J. Soil Sci. Plant Nutr.* 12, 221–244. doi: 10.4067/S0718-95162012000200003

Wassie, M., Zhang, W. H., Zhang, Q., Ji, K., Cao, L. W., and Chen, L. (2020). Exogenous salicylic acid ameliorates heat stress-induced damages and improves growth and photosynthetic efficiency in alfalfa (*Medicago sativa* l.). *Ecotoxicol. Environ. Saf.* 191, 110206. doi: 10.1016/j.ecoenv.2020.110206

Wen, P. F., Chen, J. Y., Wan, S. B., Kong, W. F., Zhang, P., Wang, W., et al. (2008). Salicylic acid activates phenylalanine ammonia-lyasein grape berry in response to high temperature stress. *Plant Growth Regul.* 55, 1–10. doi: 10.1007/s10725-007-9250-7

Wu, Z., Li, T., Cao, X., Zhang, D. H., and Teng, N. J. (2022a). Lily WRKY factor LlWRKY22 promotes thermotolerance through autoactivation and activation of LlDREB2B. *Hortic. Res.* 9, uhac186. doi: 10.1093/hr/uhac186

Wu, Z., Li, T., Liu, X. Y., Yuan, G. Z., Hou, H. Z., and Teng, N. J. (2021). A novel R2R3-MYB transcription factor LlMYB305 from *Lilium longiflorum* plays a positive role in thermotolerance *via* activating heat-protective genes. *Environ. Exp. Bot.* 184, 104399. doi: 10.1016/j.envexpbot.2021.104399

Wu, Z., Li, T., Xiang, J., Teng, R. D., Zhang, D. H., and Teng, N. J. (2023). A lily membrane-associated NAC transcription factor LlNAC014 is involved in thermotolerance *via* activation of the DREB2-HSFA3 module. *J. Exp. Bot.* 74, 945–963. doi: 10.1093/jxb/erac436

Wu, Z., Li, T., Zhang, D. H., and Teng, N. J. (2022b). Lily HD-zip I transcription factor LlHB16 promotes thermotolerance by activating *LlHSFA2* and *LlMBF1c*. *Plant Cell Physiol.* 63, 1729–1744. doi: 10.1093/pcp/pcac131

Wu, Z., Liang, J. H., Wang, C. P., Ding, L. P., Zhao, X., Cao, X., et al. (2019). Alternative splicing provides a mechanism to regulate LlHSFA3 function in response to heat stress in lily. *Plant Physiol.* 181, 1651–1667. doi: 10.1104/pp.19.00839

Wu, Z., Liang, J. H., Wang, C. P., Zhao, X., Zhong, X. H., Cao, X., et al. (2018a). Overexpression of two novel *HsfA3s* from lily in *Arabidopsis* confer increased thermotolerance and salt sensitivity *via* alterations in proline catabolism. *J. Exp. Bot.* 69, 2005–2021. doi: 10.1093/jxb/ery035

Wu, Z., Liang, J. H., Zhang, S., Zhang, B., Zhao, Q. C., Li, G. Q., et al. (2018b). A canonical DREB2-type transcription factor in lily is posttranslationally regulated and mediates heat stress response. *Front. Plant Sci.* 9, 243. doi: 10.3389/fpls.2018.00243

Xie, B. B., Qin, Q. L., Shi, M., Chen, L. L., Shu, Y. L., Luo, Y., et al. (2014). Comparative genomics provide insights into evolution of trichoderma nutrition style. *Genome Biol. Evol.* 6, 379–390. doi: 10.1093/gbe/evu018

Xin, H. B., Zhang, H., Chen, L., Li, X. X., Lian, Q. L., Yuan, X., et al. (2010). Cloning and characterization of HsfA2 from lily (*Lilium longiflorum*). *Plant Cell Rep.* 29, 875–885. doi: 10.1007/s00299-010-0873-1

Xin, H. B., Zhang, H., Zhong, X. H., Lian, Q. L., Dong, A., Cao, L., et al. (2017). Over-expression of *LlhsfA2b*, a lily heat shock transcription factor lacking trans-activation activity in yeast, can enhance tolerance to heat and oxidative stress in transgenic *Arabidopsis* seedlings. *Plant Cell Tiss. Org.* 130, 617–629. doi: 10.1007/s11240-017-1251-2

Xu, S., He, X. Y., Chen, W., and Li, J. L. (2007). Ecophysiological responses of *Festuca arundinacea* to high temperature stress. *Chin. J. Appl. Ecol.* 18, 2219–2226.

Xue, G. P., Sadat, S., Drenth, J., and McIntyre, C. L. (2014). The heat shock factor family from *Triticum aestivum* in response to heat and other major abiotic stresses and their role in regulation of heat shock protein genes. *J. Exp. Bot.* 65, 539–557. doi: 10.1093/jxb/ert399

Yang, Q., Guo, Y. J., Li, J. J., Wang, L., Wang, H., Liu, G. D., et al. (2023). Natural plant inducer 2-Amino-3-Methylhexanoic acid protects physiological activity against high-temperature damage to tea (*Camellia sinensis*). *Sci. Hortic.* 312, 111836. doi: 10.1016/j.scienta.2023.111836

Yin, H., Chen, Q. M., and Yi, M. F. (2008). Effects of short-term heat stress on oxidative damage and responses of antioxidant system in *Lilium longiflorum*. *Plant Growth Regul.* 54, 45–54. doi: 10.1007/s10725-007-9227-6

Yoshida, T., Ohama, N., Nakajima, J., Kidokoro, S., Mizoi, J., Nakashima, K., et al. (2011). *Arabidopsis* HsfA1 transcription factors function as the main positive regulators in heat shock-responsive gene expression. *Mol. Genet. Genomics* 286, 321–332. doi: 10.1007/s00438-011-0647-7

Zélicourt, A. D., Synek, L., Saad, M. M., Alzubaidy, H., Jalal, R., Xie, Y., et al. (2018). Ethylene induced plant stress tolerance by enterobacter sp. SA187 is mediated by 2-keto-4-methylthiobutyric acid production. *PloS Genet.* 14, e1007273. doi: 10.1371/journal.pgen.1007273

Zhang, Y. P., Zhu, X. H., Ding, H. D., Yang, S. J., and Chen, Y. Y. (2013). Foliar application of 24-epibrassinolide alleviates high-temperature-induced inhibition of photosynthesis in seedlings of two melon cultivars. *Photosynthetica.* 51, 341–349. doi: 10.1007/s11099-013-0031-4

Zhang, H. M., Zhu, J. H., Gong, Z. Z., and Zhu, J. K. (2022). Abiotic stress responses in plants. *Nat. Rev. Genet.* 23, 104–119. doi: 10.1038/s41576-021-00413-0

Zhao, P. B., Ren, A. Z., Dong, P., Sheng, Y. S., Chang, X., and Zhang, X. S. (2018). The antimicrobial peptaibol trichokonin IV promotes plant growth and induces systemic resistance against *Botrytis cinerea* infection in moth orchid. *J. Phytopathol.* 166, 346–354. doi: 10.1111/jph.12692

Zhao, Y., Tian, X. J., Wang, F., Zhang, L. Y., Xin, M. M., Hu, Z. R., et al. (2017). Characterization of wheat MYB genes responsive to high temperatures. *BMC Plant Biol.* 17, 208. doi: 10.1186/s12870-017-1158-4

Zhou, Z. W., Li, J. J., Zhu, C. A., Jing, B. Y., Shi, K., Yu, J. Q., et al. (2022). Exogenous rosmarinic acid application enhances thermotolerance in tomatoes. *Plants.* 11, 1172. doi: 10.3390/plants11091172

Zhou, Y. Z., Wang, Y., Xu, F. X., Song, C. X., Yang, X., Zhang, Z., et al. (2022). Small HSPs play an important role in crosstalk between HSF-HSP and ROS pathways in heat stress response through transcriptomic analysis in lilies (*Lilium longiflorum*). *BMC Plant Biol.* 22, 202. doi: 10.1186/s12870-022-03587-9

Zhuang, K. Y., Gao, Y. Y., Liu, Z. B., Diao, P. F., Sui, N., Meng, Q. W., et al. (2020). WHIRLY1 regulates *HSP21.5A* expression to promote thermotolerance in tomato. *Plant Cell Physiol.* 61, 169–177. doi: 10.1093/pcp/pcz189

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 09 June 2023
DOI 10.3389/fpls.2023.1200071

OPEN ACCESS

EDITED BY
Chenxia Cheng,
Qingdao Agricultural University, China

REVIEWED BY
Zhi Li,
Northwest A&F University, China
Bilal Ahmad,
Northwest University, China

*CORRESPONDENCE
Fuchun Zhang
zhangfc@xaas.ac.cn
Xinyu Wu
wuxy@xaas.ac.cn

†These authors have contributed equally to this work

RECEIVED 04 April 2023
ACCEPTED 16 May 2023
PUBLISHED 09 June 2023

CITATION
Zhong H, Yadav V, Wen Z, Zhou X, Wang M, Han S, Pan M, Zhang C, Zhang F and Wu X (2023) Comprehensive metabolomics-based analysis of sugar composition and content in berries of 18 grape varieties.
Front. Plant Sci. 14:1200071.
doi: 10.3389/fpls.2023.1200071

Comprehensive metabolomics-based analysis of sugar composition and content in berries of 18 grape varieties

Haixia Zhong†, Vivek Yadav†, Zhang Wen, Xiaoming Zhou, Min Wang, Shouan Han, Mingqi Pan, Chuan Zhang, Fuchun Zhang* and Xinyu Wu*

The State Key Laboratory of Genetic Improvement and Germplasm Innovation of Crop Resistance in Arid Desert Regions (Preparation), Key Laboratory of Genome Research and Genetic Improvement of Xinjiang Characteristic Fruits and Vegetables, Institute of Horticultural Crops, Xinjiang Academy of Agricultural Sciences, Urumqi, China

Xinjiang is the largest grape-producing region in China and the main grape cultivation area in the world. The Eurasian grape resources grown in Xinjiang are very rich in diversity. The sugar composition and content are the main factors that determine the quality of berries. However, there are currently no systematic reports on the types and contents of sugars in grapes grown in Xinjiang region. In this research, we evaluated the appearance and fruit maturity indicators of 18 grape varieties during fruit ripening and determined their sugar content using GC-MS. All cultivars primarily contained glucose, D-fructose, and sucrose. The glucose content in varieties varied from 42.13% to 46.80% of the total sugar, whereas the fructose and sucrose contents varied from 42.68% to 50.95% and 6.17% to 12.69%, respectively. The content of trace sugar identified in grape varieties varied from 0.6 to 2.3 mg/g. The comprehensive assessment by principal component analysis revealed strong positive correlations between some sugar components. A comprehensive study on the content and types of sugar will provide the foundation to determine the quality of grape cultivars and effective ways to utilize resources to improve sugar content through breeding.

KEYWORDS

sugar content, grape varieties, berry morphology, GC-MS, physicochemical characteristics

1 Introduction

Sugar composition and content are the main factors used to measure the quality of fruit, and they are essential for superior enological characteristics (Torregrosa et al., 2015; Kanayama, 2017). Sugar content in grape berries is another major factor that determines fruit quality (Zhang et al., 2022a). Sugar accumulation plays an important role in the synthesis of flavor substances and secondary metabolites and is a key indicator to judge

whether the fruit is fully mature (Brumos, 2021). The growth and development of berries can be divided into three stages. In the first stage, the accumulation of sugary substances is less, the content of acidic substances is higher, and the fruit grows faster (Bigard et al., 2019; Durán-Soria et al., 2020). Compared to the first and third stages of development, the growth and development of berries are much faster in the second stage. In the third stage, the fruit growth is slower, and the accumulation of various compounds, including sugar substances and anthocyanins, takes place (Manning et al., 2001; Tadeo et al., 2020). The content of acidic substances begins to decline, and the fruit gradually becomes soft (Agasse et al., 2009; Lecourieux et al., 2014). The accumulation of sugar substances in grape berries is dominated by the accumulation of glucose and fructose, supplemented by the accumulation of sucrose (Fontes et al., 2011). These three sugar components and contents play a key role in the formation of fruit quality (Varandas et al., 2004; Liu et al., 2017). Glucose and fructose in grapes account for 97.88%–99.86%, and the proportion of both in the ripening stage is 0.74–1.05 (Li et al., 2020). Sucrose is the main form of sugar transportation and is decomposed into glucose and fructose at maturity. The glucose content in the green fruit stage is higher than that in the overripe stage, and the accumulation of sucrose accounts for less than 4% of the total sugar (Liu et al., 2006). The sugar accumulation in grapes is in direct proportion to their growth and development (Manning et al., 2001). Therefore, the accumulation of fructose, glucose, and sucrose is an important carbohydrate substance affecting fruit quality. The movement and transformation of products of photosynthesis are linked to the sugar accumulation in fruit, and the activity of enzymes involved in sugar metabolism is key to controlling the source-sink relationship (Ruan, 2014; Zhu et al., 2018; Parker et al., 2020).

The accumulation of carbohydrates in fruits is closely related to the activity of metabolism-related enzymes (Liu et al., 2006; Li et al., 2017). High acid invertase (AI) activity in grape berries was observed during growth, development, and maturation, and the activities of sucrose synthase (SS) and sucrose phosphate synthase (SPS) are directly related to the content of sugars (Bilska-Kos et al., 2020; Liao et al., 2022). SS is mainly located in the cytoplasm and can control the functionality of the fruit sink organs. SS catalyzes the following reversible reactions: fructose+UDPG ⟷ sucrose+UDP (guanosine diphosphate) (Durán-Soria et al., 2020). SS-c showed high activity in the early stage of grapefruit development and decreased activity at maturity (Jiang et al., 2014; Liu et al., 2022b). The study found that sucrose synthase exhibits strong activity in the decomposition direction during the early stage of peach fruit development, while sucrose synthase shows strong activity in the synthesis direction during the late stage (Vimolmangkang et al., 2016; Zhang et al., 2019). With the exception of weak SS-c activity during fruit setting and harvest, SS-s activity in orange fruit is strong during other periods (Shi et al., 2016). In watermelon, the activity of SS-c is consistently stronger than that of SS-s, indicating that sucrose synthase is mainly involved in catalyzing the reversible cleavage of sucrose into fructose (Liu et al., 2006). Neutral invertase (NI) and aldose reductase (AI) are negatively correlated with sucrose content and positively correlated with hexose content (Wang et al., 2021). Previous findings have shown that sucrose accumulation in citrus fruits during the young fruit and expansion stages increases gradually with decreasing invertase (INV) activity (Gutiérrez-Miceli et al., 2002). Additionally, it has been observed that acid invertase activity is strong during the young fruit stage of *Myrica rubra* and consistently high in the 'Feizixiao' fruit of litchi (Huicong et al., 2002). In tomato, invertase activity is low during the early stages of development but gradually increases as the development process accelerates, reaching its highest point during maturity. However, sucrose content is low during the later stages (Quinet et al., 2019; Durán-Soria et al., 2020). Therefore, there is a growing interest in exploring various sugar types in fruit and developing an understanding of sugar accumulation based on variety differences.

The grapevine (*Vitis vinifera* L.) is cultivated worldwide due to its lucrative nature as a fruit crop that thrives in various climates (Grassi and De Lorenzis, 2021). Grapes have a high economic value because, in addition to being eaten fresh, they are used to produce juice and wine. Additionally, grapes contain numerous healthy nutrients beneficial to human health. In recent decades, grape cultivation has significantly expanded in scale throughout China (Zhong et al., 2022). Xinjiang, owing to its unique location, long hours of sunshine, and wide range of natural resources, has become the most important region in China for producing high-quality berries (Zhang et al., 2022a). The total cultivated land area for grapevines in Xinjiang has reached 26,000 hectares (Zhang et al., 2022a; Zhang et al., 2022b). Recent studies showed that this grape growing region is an important center from a domestication and evolution point of view. Many elite cultivars have been reported with specific berry characteristics. Comparative studies are essential in determining fruit quality and exploring the potential of crops in the region. Furthermore, different types of sugar metabolism-related enzymes contribute differently to sugar accumulation, but there is a correlation between them. Fruit sugar accumulation is achieved through the cooperation of sugar metabolism-related enzymes. Although various scientific studies and reports have revealed the composition and content of sugar units in popular grape varieties, there have been no systematic comparative studies on the types and contents of sugar composition between popular and local elite grape varieties. Therefore, it is of great significance to study the variation in individual and total soluble sugar contents among different grape varieties. In this paper, 18 grape varieties from Xinjiang were used as research materials. Their appearance and internal basic indicators during fruit ripening were analyzed, and their sugar content was detected and analyzed by GC-MS. By clarifying the content and types of sugar in Xinjiang grape germplasm resources, conducting a comprehensive evaluation, and identifying grape varieties rich in sugar, it will be possible to regulate sugar, which is an important quality, and find more effective ways to improve sugar content, ultimately achieving the breeding goal.

2 Materials and methods

2.1 Grape verities and field cultivation

The experiment was conducted at the grape research base of the Horticulture Institute (87.28'E, 45.56'N), Xinjiang Academy of

Agriculture Sciences, under the fruit quality control post of the national grape industry technology system, from July to December 2018. The experimental base covers an area of four hectares and is equipped for irrigation, fertilization, weeding, and spraying. The base is well constructed and developed. In the current study, 18 varieties of fresh grapes were used as test materials over the course of 6 years. These varieties covered the main types with varying maturities, shapes, and colors, as outlined in Table 1. All varieties were planted on the same location in 2012 and began bearing fruit in 2014. Grapes plants were spaced in 1×3.5 m, and standard cultivation practices were followed. For each variety, 15 healthy trees with uniform growth and the same flowering period were selected. Each group consisted of five trees with three biological repeats and berries were harvested at their optimum technological maturity as per the OIV resolution VITI 1/2008 (OIV, 2008).

2.2 Fruit sample collection, handling, and extraction

After the fruit sample was cleaned well, the exocarp and mesocarp were quickly separated with a scalpel, cut into small pieces, and put in liquid nitrogen. Some extra samples were transferred to the laboratory and stored in a freezer at -80°C. The grape berry samples were freeze-dried in a vacuum, and the freeze-dried berry was ground (30 Hz, 1.5 minutes) to powder by using a grinder with zirconia beads. Weigh 20 mg grape powder and add 500 µL methanol: isopropanol: water (3:3:2 V/V/V) extract, vortex for 3 min, and ultrasonic in ice water for 30 min. Centrifuge, add an internal standard, and then freeze dry. The chemical derivation method is as follows: a small molecule carbohydrate sample and 100 µL methoxy ammonium salt pyridine (15 mg/mL) solution were mixed, and the mixture was incubated at 37°C for 2 h. Then, 100 mL of BSTFA solution was added to the mixture and incubated at 37°C for 30 min to obtain the derivatized solution. The solution was diluted with n-hexane to an appropriate concentration and stored in a brown sampling bottle for analysis.

2.3 Determination of basic indicators in the field

The vertical and horizontal diameters of fruit were measured with an electronic digital vernier caliper, and the fruit shape index was calculated (Barbagallo et al., 2020). The weight of a single fruit was measured with a 1/10000 electronic analytical balance. The content of total soluble solids concentration (°Brix) was determined by a PAL-1 digital refractometer, and the total acid was determined by a PAL-BX/ACID-2 Brix acid-meter.

2.4 GC-MS analysis

Sugar determination was carried out using the GC-MS method (Medeiros and Simoneit, 2007; Ruiz-Matute et al., 2011; Milkovska-Stamenova et al., 2015; Wang et al., 2022). Briefly, sugar substances

TABLE 1 Basic profile of 18 varieties used in the current study.

Name of Cultivar	Abbreviation	Species	Maturity	Colour	Seeds
Summer Black	SB	*V.vinifera* × *V.Labrusca*	Early-maturing	Black	Seedless
Bronx seedless	BS	*V.vinifera* × *V.Labrusca*	Early-maturing	Light red	Seedless
Crimson seedless	CS	*V.vinifera*	Extremely late maturing	Red	Seedless
Flame seedless	FS	*V.vinifera*	Early-maturing	Red	Seedless
Rizamat	RZ	*V.vinifera*	Medium	Red	Seedy
XinYu	XY	*V.vinifera*	Medium	Red	Seedless
Wuhecuibao	WHCB	*V.vinifera*	Early-maturing	Green	Seedless
Shine Muscat	SM	*V.vinifera* × *V.Labrusca*	Medium	Green	Seedy
Victoria	VT	*V.vinifera*	Early-maturing	Green	Seedy
Black Monukka	MK	*V.vinifera*	Mid early maturity	Purple	Seedless
Red Globe	RG	*V.vinifera*	Late maturing	Red	Seedy
Thompson seedless	TS	*V.vinifera*	Mid early maturity	Green	Seedless
Centennial seedless	CT	*V.vinifera*	Medium	Green	Seedless
Munake	MN	*V.vinifera*	Late maturing	Green	Seedy
Yatomi Rosa	FHYDM	*V.vinifera*	Mid early maturity	Red	Seedy
Huozhouheiyu	HZ	*V.vinifera*	Mid early maturity	Black	Seedless
Huozhouhongyu	HZHY	*V.vinifera*	Mid early maturity	Red	Seedless
Melissa	ML	*V.vinifera*	Late maturing	Green	Seedless

in grape berries were analyzed by gas chromatography (Agilent 7890B), mass spectrometry (7000 d), and a DB-5MS column. With helium as the carrier gas, the flow rate is 1 mL/min. The injector and source temperatures were maintained as per standard procedures. The oven temperature ramp progress was maintained at 170°C, 250°C, 280°Agilent 7890B), mass spectrometry (7000 d), and a DB-5MS column. With helium as the carrier gas, the flow rate is 1 mL/min. The injector and source temperatures were maintained as per standard procedures. The oven temperature ramp progress was maintained at 170°C, 250°C, 280°C, and 310°C. The details of the GC-MS analysis for specific parameters are presented in Table 2.

Based on the GC-MS platform, MetWare software (Wuhan, China, http://www.metware.cn/) was used to do a qualitative and quantitative analysis of the sugar components (Wang et al., 2022). For each group of samples, three biological replications were maintained. Sugar standards were procured from Olchemim, Aladdin (Shanghai), and Sigma (America). The detected sugar components include 9 monosaccharides and 4 disaccharides (Table 3).

2.5 GC-MS data evaluation

The Agilent MassHunter qualitative and quantitative software was used for data processing. The total ion current diagram is shown in Figure S1. It can be seen from this study that the repeatability and reliability of QC sample data are well, as shown in Figure S2. Attached Table S1 is the linear equation for different carbohydrate species.

Calculation:

$$\text{Sugar content}(mg/g) = BCE/D/F/1000000$$

B: Sugar concentration value (μg/m)

C: Volume of solution for constant volume (μL)

D: Volume of supernatant (μL)

E: Volume of extract (μL)

F: Weighed sample mass (g)

2.6 Statistical analysis and plotting

The mean of three replicates was used to express all berry quality parameters and different sugar data. IBM SPSS v25.0 software (IBM SPSS Inc., Chicago, IL, USA) was employed to analyze Duncan's test at a different level of significance. Pearson's correlation coefficient was used to determine the correlation between variables. R software was utilized to conduct principal component analysis (PCA) and plot the grapes.

3 Results

3.1 Basic berry quality indexes of different grape varieties at maturity

The phenotypes, such as appearance and morphology, of the 18 grape varieties are known for their own characteristics at fruit maturity (Figure 1). The basic characteristics of each grape variety at fruit maturity are measured, including weight, size, firmness, and total acid, and presented in Table 4. Based on the data in the table, it is evident that all of the varieties used in this study differ significantly. For instance, results showed that in terms of individual bunch weight, 'XinYu' has the largest bunch weight of 954.97g, followed by 'Red Globe', and so on. The 'Wuhecuibao' variety had the smallest panicle weight of 201.46g, while the panicle

TABLE 2 Mass spectrometer condition and specific parameters during GC-MS analysis.

Mass spectrometry conditions	Parameter
Sample quantity	3 μL
Front Inlet Mode	3:1
Carrier Gas	Helium
Column	DB-5MS (30 m x 0.25 mm x 0.25 μm)
Column Flow	1mL min^{-1}
Oven Temperature Ramp	170°C (1min), raised to 250°C at a rate of 10°C/min, raised to 280°C at a rate of 4°C /min, raised to 310°C at a rate of 25°C/ min, 310°C (3.72min)
Front Injection Temperature	250°C
Transfer Line Temperature	240°C
Ion Source Temperature	230°C
Quad Temperature	150°C
Electron Energy	70eV
Scan mode	SIM
Solvent Delay	3.5min

TABLE 3 Ion pair information of different sugar components.

Index	Class	Mol. Weight (Da)	Q1 (Da)	Rt (min)	Compounds	KEGG ID
Ara	Monosaccharide	150	307	4.35	D-Arabinose	C00216
Xylitol	Monosaccharide	152	307	4.723	Xylitol	C00379
Rha	Monosaccharide	164	321	4.834	L-Rhamnose	C00507
Fuc	Monosaccharide	164	117	5.051	L-Fucose	C01019
Fru	Monosaccharide	180	364	6.168	D-Fructose	C10906
Gal	Monosaccharide	180	319	6.354	D-Galactose	C00124
Glu	Monosaccharide	180	364	6.407	Glucose	C00031
Sorbitol	Monosaccharide	182	319	6.78	D-Sorbitol	C00794
Inositol	Monosaccharide	180	265	8.186	Inositol	C00137
Suc	Disaccharide	342	437	13.582	Sucrose	C00089
Lac	Disaccharide	342	361	14.208	Lactose	C00243
Mal	Disaccharide	342	361	14.822	Maltose	C00208
Tre	Disaccharide	342	361	14.9	Trehalose	C01083

weight of 'XinYu' was nearly five times higher. A large variation was observed in the longitudinal diameter of individual grape bunches among different varieties. The results showed that bunch longitudinal diameter ranged from 17.58 to 30.78 cm. The highest bunch longitudinal diameter was observed in 'Centennial seedless' and the smallest was observed in 'Wuhecuibao'. The bunch longitudinal diameter of 'Centennial seedless' was 75.08% larger than 'Wuhecuibao'. In terms of bunch diameter, 'Yatomi Rosa' had the largest diameter of 22.21cm, while 'Wuhecuibao' had the smallest bunch diameter of 8.49 cm. The bunch diameter of 'Yatomi Rosa' was 161.60% larger than that of 'Wuhecuibao'.

Grape berries of different varieties with different colors and shapes are shown in Figure 1. The parameters related to the berry index were also measured to understand the structural variation among the varieties. The longitudinal diameter of barriers ranged from 14.28 to 35.43mm. The longitudinal diameter of 'Rizamat' was the largest, and 'Thompson seedless' had the smallest longitudinal diameter. The fruit longitudinal diameter of 'Rizamat' was 148.11% larger than that of 'Thompson seedless'. In terms of berry transverse diameter, 'XinYu' was in top rank with 27.74 mm, and 'Thompson seedless' is the smallest, reaching 10.9 mm. The berry transverse diameter of 'XinYu' was 53.34% larger than that of the smallest variety. As for single berry weight, 'XinYu' was the largest, and 'Thompson seedless' was the smallest.

Total soluble solid (TSS) is among the most essential attributes that determine the quality of fruits. TSS is dominated by total sugar content, with a minor contribution from soluble proteins, amino acids, and other organic materials (Xu et al., 2022). In this study, the

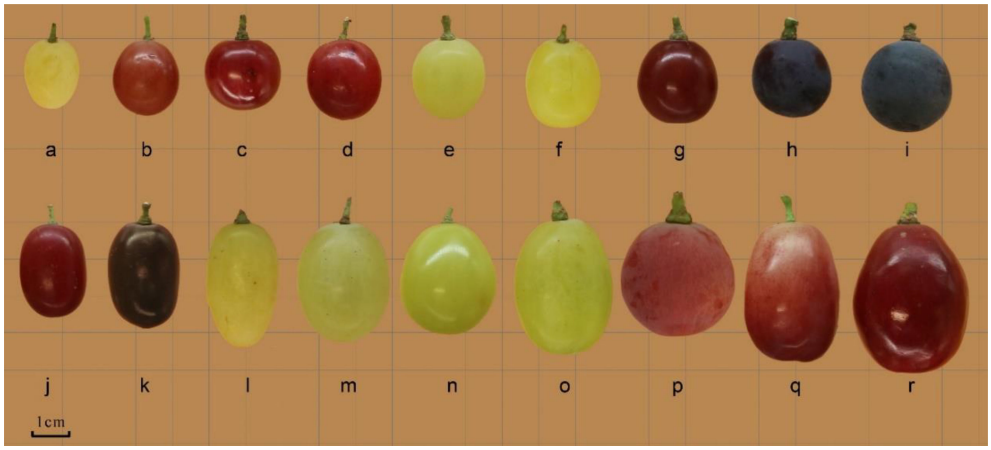


FIGURE 1
Fruit morphology of different grape cultivars. (a. Thompson Seedless; b. Bronx Seedless; c. Huozhouhongyu; d. Flame Seedless; e. Wuhecuibao; f. Melissa; g. Yatomi Rosa; h. Huozhouheiyu; i. Summer Black; j Crimson Seedless; k. Black Monukka; l. Centennial seedless; m. Munake; n. Shine-Muscat; o. Victoria; p. Red Globe; q. Rizamat; r. XinYu).

TABLE 4 Basic berry quality indexes of 18 varieties used in the present study.

Varieties	Panicle weight (g)	Bunch longitudinal diameter (cm)	Bunch diameter (cm)	Bunch shape index	Berry longitudinal diameter (mm)	Transverse diameter (mm)	Berry shape index	Single berry weight (g)	Soluble solids (° Brix)	Total acid (%)	TTS/TA	Harvest date
Summer Black	241.74 ±30.34i^{jIJ}	22.27±1.58^{eCD}	10.92±0.55^{ijFG}	2.06 ±0.21^{bcBCD}	19.46±0.54^{efF}	18.3±0.56^{fgEF}	1.06 ±0.01^{ghG}	4.40±0.28^{eE}	25.12±0.78^{aA}	0.48 ±0.02^{cdCD}	52.43 ±0.21^{deCD}	8.11
Bronx Seedless	372.64 ±9.31^{fghFGH}	27.54±1.60^{abcdABC}	11.90±0.38^{ghijFG}	2.31±0.07^{abAB}	19.24±0.20^{efF}	16.48±0.20^{gFG}	1.17 ±0.01^{fgEFG}	3.04±0.07^{hGH}	17.25 ±0.22^{gEF}	0.53 ±0.04^{bcC}	33.00 ±1.93^{ijkGH}	8.25
Crimson Seedless	600.16 ±37.10^{cBCD}	25.10±0.24^{cdeABC}	17.38 ±0.50^{bcdeABCDE}	1.45 ±0.03e^{fCDEF}	20.42±0.40^{eF}	13.86±0.75^{hGH}	1.48±0.06^{cC}	2.30±0.14^{iH}	22.41±0.13^{bB}	0.36 ±0.03^{eEF}	63.23 ±5.35^{cB}	9.28
Flame Seedless	468.23 ±41.20d^{eEF}	26.23±2.32^{abcdeABC}	15.44 ±2.22^{cdefghBCDEF}	1.73 ±0.09^{cdefBCDEF}	17.08±0.49^{fFG}	16.60±0.70^{gFG}	1.03 ±0.04^{ghG}	2.75±0.38^{ghGH}	24.71±0.53^{aA}	0.48 ±0.01^{bcdCD}	51.24 ±1.59^{deCD}	8.18
Rizamat	611.36 ±13.16^{cBC}	30.54±1.21^{abA}	20.84±1.43^{abAB}	1.48 ±0.08^{efCDEF}	35.43±0.55^{aA}	20.81±0.60^{deCDE}	1.71±0.05^{bB}	7.84±0.52^{cC}	21.89 ±0.49^{bcBC}	0.49 ±0.01^{bcCD}	44.68 ±1.64^{efgDEF}	8.19
XinYu	954.97 ±12.30^{aA}	29.58±1.37^{abcAB}	19.44 ±0.67^{abcA}B^{C}	1.52 ±0.02^{defCDEF}	29.07±0.93^{cCD}	27.74±0.78^{aA}	1.05 ±0.02^{ghG}	11.59±0.28^{aA}	21.93 ±0.18^{bcBC}	0.29 ±0.02^{fFG}	76.92 ±5.31^{bA}	9.13
Wuhecuibao	201.46 ±13.39^{jJ}	17.58±0.65^{fD}	8.49±0.56^{jG}	2.10 ±0.22^{bcABC}	18.20±0.93^{efF}	17.85±0.82^{fgEF}	1.03 ±0.09^{ghG}	3.37±0.15^{fgFG}	19.96 ±0.23^{deCD}	0.23 ±0.01^{fG}	86.24 ±2.23^{aA}	8.16
Shine Muscat	607.58 ±39.75^{cBC}	29.13±1.33^{abcAB}	16.05 ±2.79^{cdefBCDEF}	1.96 ±0.43^{bcdeBCDE}	25.40±0.61^{dE}	25.24±0.27^{bAB}	1.01 ±0.03^{hG}	7.55±0.13^{cC}	21.98 ±0.84^{bcBC}	0.46 ±0.01^{cdCD}	48.23 ±1.56^{defCDE}	9.12
Victoria	509.01 ±34.24^{dCDE}	23.43±1.09^{deBCD}	15.86 ±0.63^{cdefgBCDEF}	1.48 ±0.05^{efCDEF}	33.06±0.20^{bAB}	22.43±0.31^{cdBCD}	1.47±0.03^{cC}	10.12±0.21^{bB}	14.93±0.21^{hG}	0.27 ±0.02^{fG}	56.19 ±3.47^{cdBC}	8.17
Black Monukka	419.13 ±20.10^{efgEFG}	27.88±1.24^{abcdABC}	13.63 ±0.64^{efghiDEFG}	2.06 ±0.17^{bcBCD}	25.33±0.73^{dE}	17.97±0.50^{fgEF}	1.41 ±0.02^{cdCD}	5.52±0.32^{dD}	18.73 ±0.11^{efDE}	0.67 ±0.01^{aA}	28.06 ±0.29^{kH}	8.28
Red Globe	709.47 ±14.85^{bB}	28.25±2.34^{abcABC}	17.77 ±1.46^{bcdABCDE}	1.60 ±0.12^{cdefCDEF}	26.60±0.53^{dDE}	25.05±1.40^{bAB}	1.07 ±0.09^{ghG}	7.97±0.25^{cC}	20.02 ±0.12^{deCD}	0.55 ±0.04^{bBC}	36.77 ±2.43^{ghijFGH}	9.16
Thompson Seedless	332.76 ±16.53^{ghGHI}	27.75±0.84^{abcdABC}	11.65±0.31^{hijFG}	2.38±0.05^{abAB}	14.28±0.37^{gG}	10.90±0.11^{iI}	1.31 ±0.02^{deCDE}	2.29±0.08^{hH}	24.65 ±0.22^{bcBC}	0.64 ±0.01^{aA}	33.97 ±0.65^{hijkFGH}	8.21
Centennial Seedless	277.48 ±14.10h^{ijHIJ}	30.78±0.99^{aA}	11.26±0.46^{ijFG}	2.74±0.08^{aA}	26.55±0.49^{dDE}	12.81±0.28^{hiHI}	2.07±0.05^{aA}	4.09±0.28^{efEF}	19.03 ±0.56^{efDE}	0.67 ±0.02^{aA}	28.49 ±1.54^{jkH}	9.06
Munake	336.51 ±23.28^{ghGHI}	24.74±1.21^{cdeABC}	12.56 ±1.26^{fghiEFG}	2.02 ±0.26^{bcdBCDE}	26.43±1.40^{dDE}	24.94±0.96^{bAB}	1.06 ±0.03^{ghG}	8.29±0.08^{cC}	17.71 ±0.56^{fgEF}	0.49 ±0.01^{bcCD}	36.08 ±1.24^{hijkFGH}	9.27
Yatomi Rosa	702.41 ±25.31^{bB}	29.14±2.37^{abcAB}	22.21±1.43^{aA}	1.33±0.14^{fEF}	31.20±0.82^{bcBC}	23.15±0.88^{bcBC}	1.35 ±0.03^{cdeCDE}	8.21±0.21^{cC}	16.41 ±0.15^{gFG}	0.42 ±0.01^{deDE}	39.4167 ±1.02^{ghiEFGH}	8.22
Huozhouheiyu	431.11 ±33.24^{defEFG}	22.58±0.66^{eCD}	16.39 ±0.08^{cdefBCDEF}	1.38±0.04^{fDEF}	17.98±0.70^{efF}	18.09±0.85^{fgEF}	1.00 ±0.03^{hG}	5.68±0.13^{dD}	20.59 ±0.61^{cdBCD}	0.49 ±0.01^{bcCD}	42.08 ±2.32^{fghDEFG}	8.23

(Continued)

10.3389/fpls.2023.1200071

10.3389/fpls.2023.1200071

TABLE 4 Continued

Varieties	Panicle weight (g)	Bunch longitudinal diameter (cm)	Bunch diameter (cm)	Bunch shape index	Berry longitudinal diameter (mm)	Transverse diameter (mm)	Berry shape index	Single berry weight (g)	Soluble solids (° Brix)	Total acid (%)	TTS/TA	Harvest date
Huozhouhongyu	296.85 ±23.41^{hiHIJ}	17.71±0.46^{fD}	14.24 ±0.61^{defghiCDEF}	1.25±0.02^{fF}	20.05±0.99^{eF}	18.39±0.50^{fgEF}	1.09 ±0.05^{ghFG}	2.74±0.07^{ghGH}	20.72 ±0.44^{cdBCD}	0.61 ±0.05^{aAB}	34.33 ±3.29^{hijkFGH}	8.26
Melissa	495.35 ±42.44^{deDE}	25.81±1.82^{bcdeABC}	18.22 ±1.99^{bcdABCD}	1.43 ±0.10^{fCDEF}	25.14±1.89^{dE}	19.87±0.84^{efDE}	1.26 ±0.06^{efDEF}	4.67±0.38^{eDE}	21.28 ±0.80^{bcdBC}	0.48 ±0.02^{cdCD}	44.91 ±3.49^{efgDEF}	9.12

The different lowercase letters in the column indicate significant differences between cultivars ($p < 0.05$), different capital letters indicate significant differences between cultivars ($p < 0.01$).

TSS of 'Summer Black' and 'Flame Seedless' was found to be 25.12 and 24.71° Brix, respectively, which was higher than other grape varieties, while 'Victoria' had the lowest TSS of 14.93° Brix. 'Summer Black' and 'Flame Seedless' had 68.25% and 65.50% higher TSS than 'Victoria,' respectively. Furthermore, 'Black Monukka' and 'Centennial Seedless' had the highest total acid content of 0.67%, while 'Victoria' had the lowest total acid content of 0.27% among all grape varieties studied.

The study conducted by Huang et al. (2021) used the soluble solids to acid ratio as an indicator to evaluate grape maturity and flavor. The TSS/TA ratio was calculated for all grape varieties, and the results showed that the ratio ranged from 86.24 (Wuhecuibao) to 28.06 (Black Monukka), indicating significant differences among the varieties. Specifically, 'Wuhecuibao' had a TSS/TA ratio that was 207.34% higher than that of 'Black Monukka'. Berry maturity is a crucial stage for harvesting, and fruit quality often depends on maturity indices (Niimi et al., 2017; Niimi et al., 2018). The grape varieties used in this study were classified as early and late based on their harvest dates. 'Summer Black', 'Victoria', and 'Flame seedless' matured early, while 'Crimson seedless' and 'Munake' matured relatively late.

3.2 Variation in monosaccharide and disaccharide substances amongst grape varieties

The GC-MS analysis revealed significant differences in the amounts of sugar components present in the berries of the 18 grape varieties tested (Figure 2). The heatmap showed that, with the exception of maltose, all identified sugar components were detected in all tested varieties, whereas maltose was only detected in five varieties. Notably, the concentration of the other nine sugar components was found to be low, while D-fructose and glucose were present in high concentrations, followed by sucrose. The heatmap depicts the average concentration of different sugar components among the varieties.

3.2.1 Variation in content of D-fructose among grape varieties

Fructose is a crucial carbohydrate in nutrition, and though it is a monosaccharide like glucose and galactose, it has a distinct flavor. Upon analyzing various sugar components, it was found that D-fructose content was highest in all varieties, constituting between 42.68% to 50.95% of the total sugar in different types. D-fructose levels in varieties ranged from 185.66 to 265 mg/g, with the top three varieties containing the highest D-fructose contents being 'Yatomi Rosa', 'Melissa,' and 'Wuhecuibao.' Of these, 'Yatomi Rosa' had the highest D-fructose content at 265 mg/g, followed by 'Wuhecuibao,' Red Globe, and 'Flame Seedless' (Figure 3). The variety with the lowest D-fructose content is 'Victoria,' containing only 185 mg/g. The difference in D-fructose content between 'Yatomi Rosa' and 'Victoria' is the largest, with 'Yatomi Rosa' having 43.24% more D-fructose content than 'Victoria.' Additionally, there is a significant difference in D-fructose content between 'Flame Seedless' and 'Victoria' grapes, with 'Flame

10.3389/fpls.2023.1200071

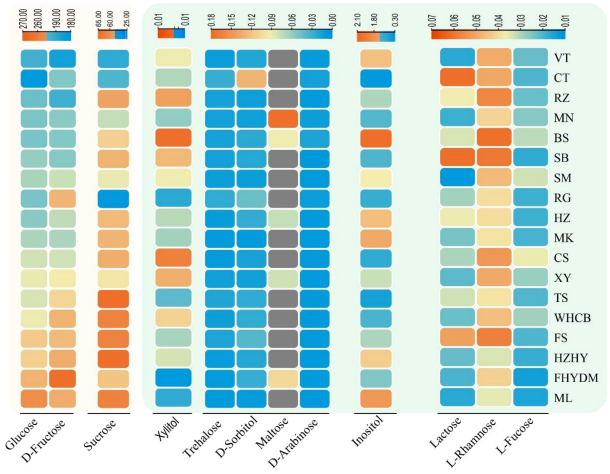


FIGURE 2
Variation and distribution of sugar contents in 18 grape varieties are presented in form of heatmap. The sugar content was measured in mg/g. Thompson Seedless (TS); Bronx Seedless (BS); Huozhouhongyu (HZHY); Flame Seedless (FS); Wuhecuibao (WHCB); Melissa (ML); Yatomi Rosa (FHYDM); Huozhouheiyu (HZ); Summer Black (SB); Crimson Seedless (CS); Black Monukka (MK); Centennial Seedless (CT); Munake (MN); Shine-Muscat (SM); Victoria (VT); Red Globe (RG); Rizamat (RZ); XinYu (XY).

Seedless' having 29.72% higher D-fructose content than 'Victoria' grapes (Figure 3).

3.2.2 Uneven distribution of glucose content in grape varieties

The glucose range in total sugar across the 18 grape varietals varied from 42.13% to 46.80%. The top three varieties with the highest glucose levels are 'Melissa', 'Yatomi Rosa', and 'Flame Seedless', with glucose contents of 254 mg/g, 242 mg/g, and 235 mg/g, respectively (Figure 4). 'Centennial Seedless' and 'Victoria' have the lowest glucose content among all varieties, with 'Centennial Seedless' having the lowest glucose level of only 182 mg/g, followed by 'Victoria' with 193 mg/g. 'Flame Seedless' has a glucose level that is 21.76% higher than 'Victoria', and the glucose content of 'Melissa' is 40% higher than that of 'Centennial Seedless'. The highest difference in glucose content is between 'Melissa' and 'Centennial seedless'. The glucose content in the majority of varieties ranged from 193.6 to 254.6 mg/g (Figure 4). Moreover, results showed that glucose content in varieties did not relate to maturity duration. Some varieties with late and early maturity had similar glucose levels. For instance, there was no significant difference in glucose levels between 'Summer Black' and 'Munake', which are early and late maturity varieties.

3.2.3 Variation in sucrose content in different varieties

Sucrose is the third most common type of sugar found in grape berries, following D-fructose and glucose. The amount of sucrose present in the total sugar content of the 18 grape varieties tested ranged from 6.17% to 12.69%. The three grape varieties with the highest sucrose content were 'Thompson Seedless',

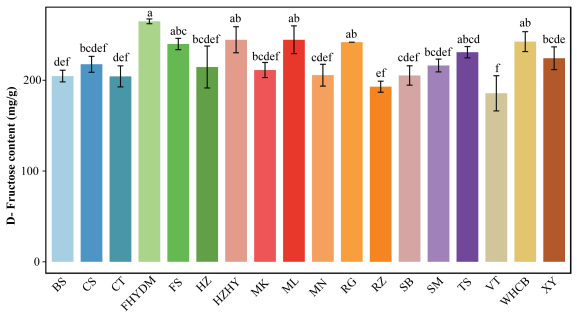


FIGURE 3
Analysis of D-fructose content in 18 varieties Bars of different colors represent the grape varieties used in the current study. The D-fructose content was measured in mg/g. The alphabets on the error bars show statistically significant values between varieties. Values are the means of a minimum of three replicates and expressed as means ± standard deviation (SD). Different superscripts in the same row indicate statistical differences using the Duncan test.

'Huozhouhongyu', and 'Melissa', with 'Thompson Seedless' having the highest sucrose content at 64 mg/g (Figure 5). The three varieties with the lowest sucrose content were 'Red Globe', 'Victoria', and 'Centennial Seedless', with 'Red Globe' having the lowest sucrose content at 29 mg/g. The difference in sucrose content between the lowest (Melissa) and highest (Thompson Seedless) was 121%.

3.2.4 Minor and trace sugar components

The estimated sugar content in grape berries, as determined by GC-MS analysis, shows that there are some sugar units present in small quantities. D-fructose, glucose, and sucrose are present in higher concentrations, while some monosaccharides and disaccharides, including D-arabinose, inositol, lactose, maltose, trehalose, and xylitol, are detected at very low concentrations. Among these, the content of inositol was higher than that of other trace sugar components, with concentrations in grape varieties varying from 0.31 to 2.04 mg/g. Notably, the content of

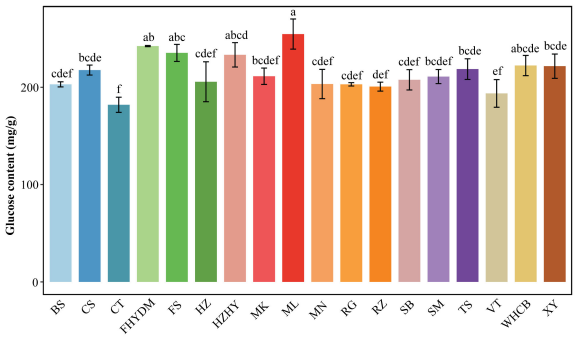


FIGURE 4
Analysis of glucose content in 18 varieties. Bars of different colors represent the grape varieties used in the current study. The D-fructose content was measured in mg/g. The alphabets on the error bars show statistically significant values between varieties. Values are means of minimum three replicates and expressed as means ± standard deviation (SD). Different superscripts with the same row indicate statistical differences using the Duncan test at different level of significance.

10.3389/fpls.2023.1200071

maltose was detected in only a few varieties, including 'Munake', 'Bronx seedless', 'Huozhouhongyu' and 'Yatomi Rosa', with concentrations ranging from 0.01 to 0.07 mg/g. The details about the concentration of trace sugar components are displayed in Figure 6A, where significant differences can be clearly seen in the concentration of inositol in different varieties. The results revealed that the concentration of inositol was exceptionally higher in 'Bronx seedless' (Figure 6B).

3.3 Differences and comparative analysis of total sugar in fruit of different genotype grape varieties

Based on the analysis of the components and contents of various sugars in different grape varieties using GC-MS, principal component analysis (PCA) was conducted to investigate the differences both within and among replicates. The results, depicted in Figure 7, reveal that the three replicates of each of the 18 grape varieties were closely clustered together, suggesting good repeatability and ensuring the accuracy of the data. Moreover, the samples of different grape varieties were clearly distinguishable, indicating that the sugar composition and content among different grape varieties differed significantly, reaching a significant or extremely significant level. The findings highlighted the substantial variation in sugar composition and content among grape varieties.

According to Figure 8A, the total sugar content in these grape varieties is determined by the sum of monosaccharides and disaccharides detected by GC-MS, including their absolute values. 'Melissa' has the highest total sugar content, followed by 'Yatomi Rosa' and 'Huozhouhongyu', and then 'Flame seedless', 'Wuhecuibao', 'Thompson seedless', and 'Xin Yu', among others. Conversely, 'Victoria', 'Centennial Seedless', and 'Rizamat' were found to have lower sugar content. In 18 varieties, D-fructose, glucose, and sucrose contribute to over 99% of the total sugar content, with D-fructose slightly higher than glucose in 12 of them (Figure 8B). Additionally, inositol was identified as a higher contributor to the minor sugar components in all grape varieties (Figure 8C).

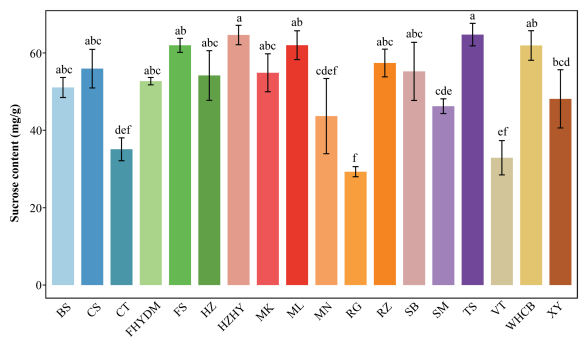


FIGURE 5
Variation in sucrose content across 18 varieties. Values are means of minimum three replicates and expressed as means ± standard deviation (SD). Different superscripts with the same row indicate statistical differences using the Duncan test.

3.4 Comparison and analysis of sugar quality of different grape varieties

The correlations between various sugar components are presented in Figure 9A using Spearman correlation coefficients. The correlation matrix revealed that some sugar components showed a positive correlation with others. For instance, Xylitol content exhibited a positive correlation with L-fucose ($p<0.001$) and L-rhamnose ($p<0.001$). Similarly, lactose content was positively correlated with D-sorbitol ($p<0.01$), L-rhamnose ($p<0.05$), and trehalose ($p<0.05$) content. A strong correlation between sucrose content and glucose content was observed. Furthermore, glucose content in varieties was found to have a strong correlation with D-fructose content ($p<0.001$). Additional positive and negative correlations were identified, and further details are presented in Figure 9A.

Results shown in Figure 9B revealed that there is no significant positive correlation between total sugar and soluble solid content. For example, although 'Yatomi Rosa' has the second-highest total sugar content, its proportion of soluble solids is lower. Similarly, 'Summer Black' has the highest soluble solids, but its total sugar content is not particularly high. In contrast, 'Flame Seedless' has relatively high levels of both total sugar and soluble solids, while 'Victoria' has lower levels of both. There are significant differences in sugar and total soluble content among some varieties. For instance, the total soluble solid content in 'Summer Black', 'Flame Seedless', and 'Thompson Seedless' is similar, but the total sugar content varies greatly. Therefore, the plot showing the relationship between total sugar and total soluble solids indicates that there is no positive correlation between these parameters.

4 Discussion

4.1 Variation in basic berry characteristics

The morphology and characteristics of berries vary among the varieties and different species of grapes (Table 4). Quantifying the phenotypic parameters of grape berries and bunches is important for precision agriculture (Liu et al., 2022a). The cultivated grapes are known to have high variation compared with wild resources, which largely resemble round berries in shape (Rodríguez et al., 2011; Zhang et al., 2021). In many fruit crop markets, demand is closely associated with the shape and quality of fruit crops. Varietal development programs are also associated with fruit shape in many fruit crops (van der Knaap et al., 2014; Zhang et al., 2021). The morphological characteristics observed in our study showed that the popular grape varieties grown in the Xinjiang region are round to oval in shape. Many studies have shown that berry appearance is highly associated with berry index and appears to influence the acceptability and preferences of consumers when it comes to fresh fruit consumption (Lecourieux et al., 2014; Zhang

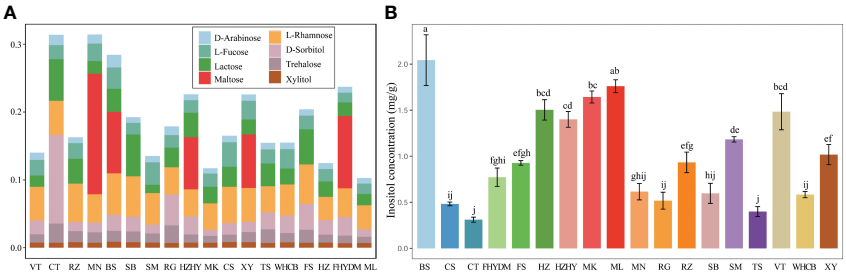


FIGURE 6
Analysis of nine trace sugar content in18 grape varieties. **(A)** The stacked bar plot shows share of eight different trace sugar in different varieties. **(B)** The content of inositol in grape varieties. Colour bars shows different varieties. The significant differences are shown with the letter on top of error bar.

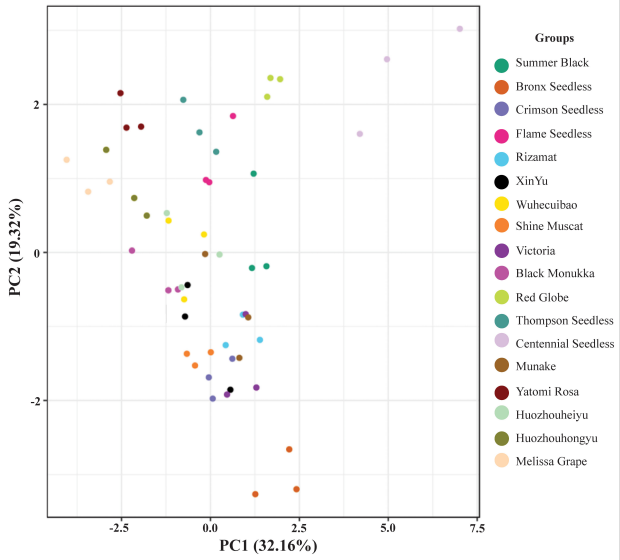


FIGURE 7
Principal component analysis of sugar content in 18 cultivars. Different color dots represents the various varieties of grapes. Multiple dots of same color shows replication of samples.

et al., 2022b). The studies by Zhang et al. (2022a) and Zhang et al. (2022b) revealed that in some regions, spine grapes are eaten because of their large shape and flavor. The berries are characterized by high variability in grapes and these traits are also used to describe different grape varieties and domestication processes (Barbagallo et al., 2020). We measured various berry traits to highlight the phenotypic diversity of berries in different varieties grown in Xinjiang. The varieties used in the current investigation showed higher variability in various berry characteristics. The significant differences were reported in bunch parameters, berry parameters, berry index, TSS, total acid and harvesting time. The technological maturity parameters of Italian table grapes were measured in some popular varieties. The shorting of varieties on a commercial level showed significant variation in different chemical parameters, including TSS, TA, TSS/TA, pH, and different major sugar components (Segade et al., 2013). Barbagallo et al. (2020) used a number of grape varieties from the Sicilian genetic pool and precisely observed the variability in seed and berry traits. Similar research was carried out by Bouby et al. (2013) to identify the difference in elongation of the pip body between primitive cultivars and highly domesticated cultivars. The results revealed that traits of domestication are related to the strength of selection pressure (Bouby et al., 2013).

4.2 Distribution of sugar contents in grape varieties

The quality of grape berries is mainly determined by the type and amount of sugar, soluble solids, and organic acids content. Among these factors, sugar content is the most important. Xinjiang is home to a wide variety of grape germplasm, but different grape varieties have varying levels of sugar. In a study of 18 grape varieties, we found that fruit size, shape, and color characteristics differed at maturity, as did sugar content. Sugar accumulation during fruit development has been extensively studied in various species, and the amount of total soluble sugars typically increases with growth, reaching a peak at maturity or ripening (Basson et al., 2010; Li et al., 2012; Cao et al., 2013). However, the patterns and concentrations of sugar accumulation can differ between species. Glucose and fructose are typically the major proportion of soluble sugars in most fruits, while sucrose is predominant in some species like peaches, citrus, and litchi (Yakushiji et al., 1996; Desnoues et al., 2014; Li et al., 2020).

In our studies, we observed a similar pattern of sugar distribution, where glucose and fructose contributed the majority of sugars. The sugar-to-acid ratio reflects fruit taste, and the flavor of grape berries is closely related to sugar-acid content (Guo et al., 2017; Zhang et al., 2022b). Fructose and glucose are the most common sugars in most fruits, and they also contribute to the flavor of grape berries. Other metabolic changes occur during grape berry ripening, such as the accumulation of sugars in the form of glucose and fructose in the berry vacuoles (flesh and skin) following sucrose translocation from the leaves (Durán-Soria et al., 2020). Recent research by Li et al. (2020) compared sugar profiles in major fruit crops and found that glucose and fructose were the most abundant sugars, which is consistent with our findings. However,

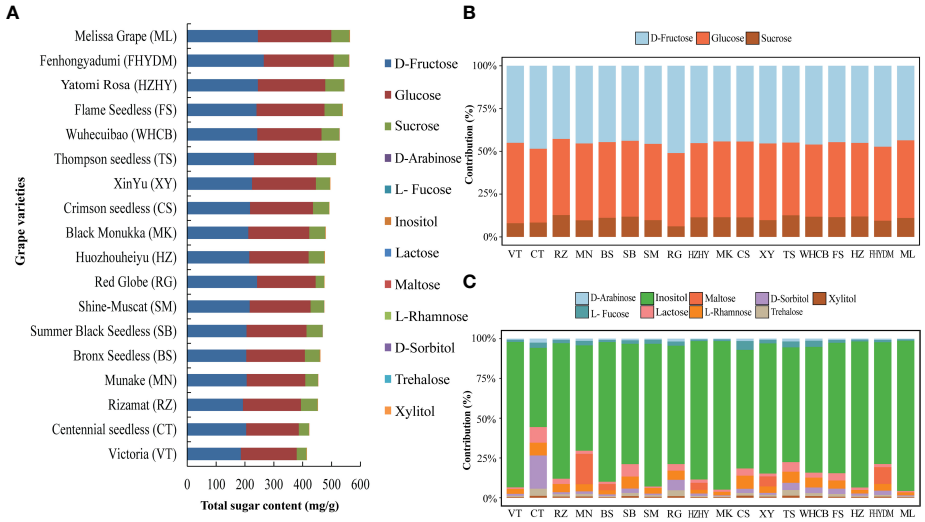


FIGURE 8
Analysis of total sugar content in 18 cultivars and percentage share of major sugar and trace sugar in each variety. **(A)** Total sugar content in all grape varieties **(B)** The percentage contribution of major sugar components to all grape varieties. **(C)** The percentage share of trace sugar components in different varieties is shown with different color bars.

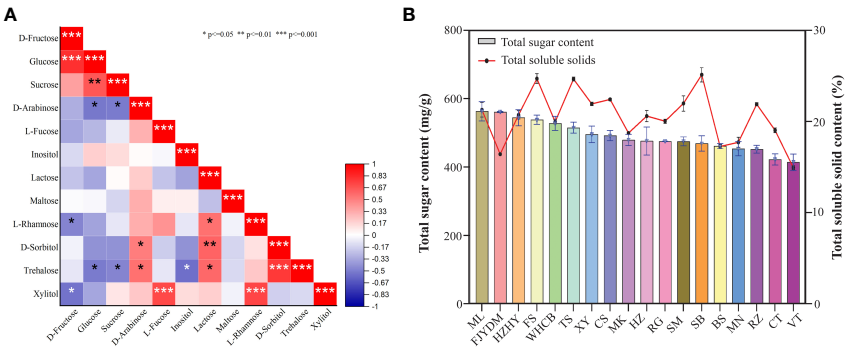


FIGURE 9
Correlation analysis of total sugar content and soluble solids. **(A)** Spearman correlation coefficients among different sugar contents. **(B)**The total sugar content is shown on left y-axis and right y-axis shows percentage of total soluble solids. * p<=0.05 ** p<=0.01 *** p<=0.001.

sucrose was found to be the predominant sugar in some fruits, including peaches.

The recent studies conducted on grapes have shown that the sugar content of wild grapes is primarily divided into fructose and glucose, as reported by Liu et al. (2006); Jiang et al. (2017); Zhang et al. (2022b) and Segade et al. (2013). Our current study supports these findings and indicates that the sugar composition and level vary significantly across grape varieties, as observed in previous studies on grapes (Zhang et al., 2022b) and citrus (Zhou et al., 2018). Specifically, our analysis of individual sugar components revealed that fructose, glucose, and sucrose are the three primary sugar components in grapes, accounting for over 99% of the total sugar content. Among the 18 grape varieties studied, fructose accounted for 42.68%-50.95% of the total sugar, while glucose accounted for 42.13%-46.80%, both accumulating at significantly higher levels than sucrose and other sugar substances, consistent with the findings of Li et al. (2020) and Liu et al. (2006). Similar results were observed in the sugar concentration of other fruits, where D-fructose, sucrose, and glucose were the major sugar contributors (Ma et al., 2017). The proportion of sucrose content in total sugar ranged from 6.17% to 12.69%. Orak (2009) reported a variation in glucose content of 5.98% to 12.21% and fructose content of 5.93% to 12.66% in 24 significant grape varieties, which is similar to our findings. The carbohydrate composition on grape varieties identified by thin layer chromatography, and spectrophotometric Dubois method also revealed that glucose and

fructose as major constituents in black grape varieties (Iosub et al., 2013). However, the variation in fructose and glucose ranges may be due to different analytical methods used for sugar content estimation. Interestingly, our current research found that the majority of grape varieties had trace amounts of monosaccharides and disaccharides, unlike many other studies. The concentration of lactose and maltose was measured in some Korean fruits and vegetables, but they were not present in most of them. Shanmugavelan et al. (2013) found that the concentration of trace sugar was not enough to be detected by HPLC in Campbell early and green varieties of grapes. In contrast, GC-MS-based sugar quantification in our study showed that many trace sugars were present in most of the grape varieties, with only five varieties showing the presence of maltose. The differences in the identification of various sugars can be attributed to analytical methods or variety differences. In our study, the sugar composition determined by GC-MS was combined with the data of soluble solids, and it was found that the sugar content, an important quality factor, differed significantly among the 18 grape varieties. GC-MS has been found to be effective in previous studies for the quantitation of carbohydrate intermediates (Milkovska-Stamenova et al., 2015). In short, current study allowed to characterize economically important grape varieties grown in Xinjiang grapes according to physiochemical and sugar composition.

5 Conclusion

Thirteen carbohydrate components were correctly detected in the berries of 18 different grape varieties using GC-MS technology. Fructose and glucose were the predominant sugar types in grape berries, followed by sucrose. However, the average content of D-arabinose, lactose, maltose, trehalose, and the other nine sugars was very low, ranging from 0.01 to 1.04 mg/g. The grape varieties with higher fructose content were 'Yatomi Rosa', 'Huozhouhongyu', and 'Melissa', while those with higher glucose content were 'Melissa', 'Yatomi Rosa', and 'Flame Seedless'. 'Thompson seedless', 'Huozhouhongyu', and 'Melissa' were the grape varieties with a higher sucrose content. The top three varieties with higher total sugar content were 'Melissa', 'Yatomi Rosa', and 'Huozhouhongyu'. Further analysis of total sugar and soluble solids showed no significant correlation between them. Positive and negative correlations were observed between some major and trace sugars. 'Flame Seedless' scored rather well on the indices of total sugar and soluble solids. 'Flame Seedless' and 'Victoria', having the same population source and maturity stage, had significant differences in sugar content and could be chosen as representatives of high- and low-sugar-type varieties for further study. Collectively, the findings suggested that a phenotypic characteristic with sugar content and type analysis can be used as a comprehensive and objective evaluation system for determining the quality of grape varieties.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

Author contributions

Author Contributions: Conceptualization, HZ and FZ; Data curation, HZ, FZ, VY, ZW, MP, and XW. Formal analysis, HZ, FZ, VY, ZW, SH, and MW; Funding acquisition, HZ and FZ, Methodology, HZ and FZ. Writing – review and editing, HZ, FZ, VY, DY, and CZ. All authors contributed to the article and approved the submitted version.

Funding

This research was financed by the Opening Foundation of Xinjiang Key Laboratory (XJTSGS-2023004), China Agriculture Research System of MOF and MARA, National Natural Science Foundation of China (32160682, 31960575), Xinjiang Uygur Autonomous Region Tianchi Talent - Special Expert Project (Whole genome design and breeding of grapes), Project of Fund for Stable Support to Agricultural Sci-Tech Renovation (xjnkywdzc-2022009), and The central government guides local science and technology development special fund projects(2022) -Germplasm Innovation and Breeding Ability Improvement of Characteristic Fruit Trees.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1200071/full#supplementary-material

References

Agasse, A., Vignault, C., Kappel, C., Conde, C., Gerós, H., and Delrot, S. (2009). "Sugar transport & sugar sensing in grape," in *Grapevine molecular physiology & biotechnology*. Ed. K. A. Roubelakis-Angelakis (Dordrecht: Springer Netherlands).

Barbagallo, M. G., Patti, D., and Pisciotta, A. (2020). Phenotypic traits of berries and seeds of Sicilian grape cultivars (Vitis vinifera l.). *Sci. Hortic.* 261, 109006. doi: 10.1016/j.scienta.2019.109006

Basson, C. E., Groenewald, J. H., Kossmann, J., Cronje, C., and Bauer, R. (2010). Sugar and acid-related quality attributes and enzyme activities in strawberry fruits: invertase is the main sucrose hydrolysing enzyme. *Food Chem.* 121 (4), 1156–1162. doi: 10.1016/j.foodchem.2010.01.064

Bigard, A., Romieu, C., Sire, Y., Veyret, M., Ojéda, H., and Torregrosa, L. (2019). The kinetics of grape ripening revisited through berry density sorting. *OENO One* 53 (4), 1–16. doi: 10.20870/oeno-one.2019.53.4.2224

Bilska-Kos, A., Mytych, J., Suski, S., Magoń, J., Ochodzki, P., and Zebrowski, J. (2020). Sucrose phosphate synthase (SPS), sucrose synthase (SUS) and their products in the leaves of miscanthus × giganteus and zea mays at low temperature. *Planta* 252 (2), 23. doi: 10.1007/s00425-020-03421-2

Bouby, L., Figueiral, I., Bouchette, A., Rovira, N., Ivorra, S., Lacombe, T., et al. (2013). Bioarchaeological insights into the process of domestication of grapevine (*Vitis vinifera* l.) during Roman times in southern France. *PloS One* 8 (5), e63195. doi: 10.1371/journal.pone.0063195

Brumos, J. (2021). Gene regulation in climacteric fruit ripening. *Curr. Opin. Plant Biol.* 63, 102042. doi: 10.1016/j.pbi.2021.102042

Cao, S., Yang, Z., and Zheng, Y. (2013). Sugar metabolism in relation to chilling tolerance of loquat fruit. *Food Chem.* 136 (1), 139–143. doi: 10.1016/j.foodchem.2012.07.113

Desnoues, E., Gibon, Y., Baldazzi, V., Signoret, V., Génard, M., and Quilot-Turion, B. (2014). Profiling sugar metabolism during fruit development in a peach progeny with different fructose-to-glucose ratios. *BMC Plant Biol.* 14 (1), 336. doi: 10.1186/s12870-014-0336-x

Durán-Soria, S., Pott, D. M., Osorio, S., and Vallarino, J. G. (2020). Sugar signaling during fruit ripening. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.564917

Fontes, N., Gerós, H., and Delrot, S. (2011). Grape berry vacuole: a complex and heterogeneous membrane system specialized in the accumulation of solutes. *Am. J. Enol. Viticult.* doi: 10.5344/ajev.2011.10125

Grassi, F., and De Lorenzis, G. (2021). Back to the origins: background and perspectives of grapevine domestication. *Int. J. Mol. Sci.* 22 (9), 4518. doi: 10.3390/ijms22094518

Guo, Y. S., Niu, Z. Z., and Su, K. (2017). Composition and content analysis of sugars and organic acids for 45 grape cultivars from northeast region of China. *Pak. J. Bot.* 49 (1), 155–160.

Gutiérrez-Miceli, F. A., Rodríguez-Mendiola, M. A., Ochoa-Alejo, N., Méndez-Salas, R., Dendooven, L., and Arias-Castro, C. (2002). Relationship between sucrose accumulation and activities of sucrose-phosphatase, sucrose synthase, neutral invertase and soluble acid invertase in micropropagated sugarcane plants. *Acta Physiol. Plant.* 24 (4), 441–446. doi: 10.1007/s11738-002-0041-5

Huang, X., Wang, H., Luo, W., Xue, S., Hayat, F., and Gao, Z. (2021). Prediction of loquat soluble solids and titratable acid content using fruit mineral elements by artificial neural network and multiple linear regression. *Sci. Hortic.* 278, 109873. doi: 10.1016/j.scienta.2020.109873

Huicong, W., Xuming, H., and Huibai, H. (2002). A study on the causative factors retarding pigmentation in the fruit of 'Feizixiao'litchi. *Acta Hortic. Sin.* 29 (5), 408.

Iosub, S., Soare, C., Rau, I., and Meghea, A. (2013). Anthocyanin and carbohydrate content in selective extracts obtained from black grape varieties. *Rev. Chimie* 64 (10), 1078–1082.

Jiang, H.-Y., Li, W., He, B.-J., Gao, Y.-H., and Lu, J.-X. (2014). Sucrose metabolism in grape (Vitis vinifera l.) branches under low temperature during overwintering covered with soil. *Plant Growth Regul.* 72 (3), 229–238. doi: 10.1007/s10725-013-9854-z

Jiang, Y., Meng, J. F., Liu, C. H., Jiang, J. F., Fan, X. C., Yan, J., et al. (2017). Quality characteristics, phenolics content and antioxidant activity of Chinese wild grapes. *Shipin Kexue* 38, 142–148.

Kanayama, Y. (2017). Sugar metabolism and fruit development in the tomato. *Horticult. J.* 86 (4), 417–425. doi: 10.2503/hortj.OKD-IR01

Lecourieux, F., Kappel, C., Lecourieux, D., Serrano, A., Torres, E., Arce-Johnson, P., et al. (2014). An update on sugar transport and signalling in grapevine. *J. Exp. Bot.* 65 (3), 821–832. doi: 10.1093/jxb/ert394

Li, M., Feng, F., and Cheng, L. (2012). Expression patterns of genes involved in sugar metabolism and accumulation during apple fruit development. *PloS One* 7 (3), e33055. doi: 10.1371/journal.pone.0033055

Li, Y., Xin, G., Wei, M., Shi, Q., Yang, F., and Wang, X. (2017). Carbohydrate accumulation and sucrose metabolism responses in tomato seedling leaves when subjected to different light qualities. *Sci. Hortic.* 225, 490–497. doi: 10.1016/j.scienta.2017.07.053

Li, J., Zhang, C., Liu, H., Liu, J., and Jiao, Z. (2020). Profiles of sugar and organic acid of fruit juices: a comparative study and implication for authentication. *J. Food Qual.* 2020, 7236534. doi: 10.1155/2020/7236534

Liao, G., Li, Y., Wang, H., Liu, Q., Zhong, M., Jia, D., et al. (2022). Genome-wide identification and expression profiling analysis of sucrose synthase (SUS) and sucrose phosphate synthase (SPS) genes family in actinidia chinensis and a. eriantha. *BMC Plant Biol.* 22 (1), 215. doi: 10.1186/s12870-022-03603-y

Liu, X. J., An, X. H., Liu, X., Hu, D. G., Wang, X. F., You, C. X., et al. (2017). MdSnRK1.1 interacts with MdJAZ18 to regulate sucrose-induced anthocyanin and proanthocyanidin accumulation in apple. *J. Exp. Bot.* 68 (11), 2977–2990. doi: 10.1093/jxb/erx150

Liu, X., Fan, P., Jiang, J., Gao, Y., Liu, C., Li, S., et al. (2022b). Evolution of volatile compounds composition during grape berry development at the germplasm level. *Sci. Hortic.* 293, 110669. doi: 10.1016/j.scienta.2021.110669

Liu, W., Wang, C., Yan, D., Chen, W., and Luo, L. (2022a). Estimation of characteristic parameters of grape clusters based on point cloud data. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.885167

Liu, H. F., Wu, B. H., Fan, P. G., Li, S. H., and Li, L. S. (2006). Sugar and acid concentrations in 98 grape cultivars analyzed by principal component analysis. *J. Sci. Food Agr.* 86 (10), 1526–1536. doi: 10.1002/jsfa.2541

Ma, Q. J., Sun, M. H., Lu, J., Liu, Y. J., Hu, D. G., Hao, Y. J., et al (2017). Transcription factor AREB2 is involved in soluble sugar accumulation by activating sugar transporter and amylase genes1. *Plant Physiol.* 174, 2348–2362. doi: 10.1104/pp.17.00502

Manning, K., Davies, C., Bowen, H. C., and White, P. J. (2001). Functional characterization of two ripening-related sucrose transporters from grape berries. *Ann. Bot.* 87 (1), 125–129. doi: 10.1006/anbo.2000.1316

Medeiros, P. M., and Simoneit, B. R. T. (2007). Analysis of sugars in environmental samples by gas chromatography–mass spectrometry. *J. Chromatogr. A* 1141 (2), 271–278. doi: 10.1016/j.chroma.2006.12.017

Milkovska-Stamenova, S., Schmidt, R., Frolov, A., and Birkemeyer, C. (2015). GC-MS method for the quantitation of carbohydrate intermediates in glycation systems. *J. Agric. Food Chem.* 63 (25), 5911–5919. doi: 10.1021/jf505757m

Niimi, J., Boss, P. K., Jeffery, D., and Bastian, S. E. P. (2017). Linking sensory properties and chemical composition of *Vitis vinifera* cv. Cabernet sauvignon grape berries to wine. *Am. J. Enol. Viticult.* 68 (3), 357–368. doi: 10.5344/ajev.2017.17083

Niimi, J., Boss, P. K., Jeffery, D. W., and Bastian, S. E. P. (2018). Linking the sensory properties of Chardonnay grape *Vitis vinifera* cv. berries to wine characteristics. *Am. J. Enol. Viticult.* 69 (2), 113–124. doi: 10.5344/ajev.2017.17083

OIV (2008). *Resolution VITI 1/2008. OIV standard on minimum maturity requirements for table grapes* (Paris: Organisation Internationale de la Vigne et du Vin).

Orak, H. H. (2009). Determination of glucose and fructose contents of some important red grape varieties by HPLC. *Asian J. Chem.* 21 (4), 3068–3072.

Parker, A. K., García de Cortázar-Atauri, I., Gény, L., Spring, J.-L., Destrac, A., Schultz, H., et al. (2020). Temperature-based grapevine sugar ripeness modelling for a wide range of vitis vinifera l. cultivars. *Agric. For. Meteorol.* 285-286, 107902. doi: 10.1016/j.agrformet.2020.107902

Quinet, M., Angosto, T., Yuste-Lisbona, F. J., Blanchard-Gros, R., Bigot, S., Martinez, J.-P., et al. (2019). Tomato fruit development and metabolism. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.01554

Rodríguez, G. R., Muños, S., Anderson, C., Sim, S.-C., Michel, A., Causse, M., et al. (2011).). distribution of SUN, OVATE, LC, and FAS in the tomato germplasm and the relationship to fruit shape diversity. *Plant Physiol.* 156 (1), 275–285. doi: 10.1104/pp.110.167577

Ruan, Y. L. (2014). Sucrose metabolism: gateway to diverse carbon use and sugar signaling. *Annu Rev Plant Biol.* 65, 33–67. doi: 10.1146/annurev-arplant-050213-040251

Ruiz-Matute, A. I., Hernández-Hernández, O., Rodríguez-Sánchez, S., Sanz, M. L., and Martínez-Castro, I. (2011). Derivatization of carbohydrates for GC and GC–MS analyses. *J. Chromatogr. B* 879 (17), 1226–1240. doi: 10.1016/j.jchromb.2010.11.013

Segade, S. R., Giacosa, S., de Palma, L., Novello, V., Torchio, F., Gerbi, V., et al. (2013). Effect of the cluster heterogeneity on mechanical properties, chromatic indices and chemical composition of italia table grape berries (*Vitis vinifera* l.) sorted by flotation. *Int. J. Food Sci. Technol.* 48 (1), 103–113. doi: 10.1111/j.1365-2621.2012.03164.x

Shanmugavelan, P., Kim, S. Y., Kim, J. B., Kim, H. W., Cho, S. M., Kim, S. N., et al. (2013). Evaluation of sugar content and composition in commonly consumed Korean vegetables, fruits, cereals, seed plants, and leaves by HPLC-ELSD. *Carbohydr. Res.* 380, 112–117. doi: 10.1016/j.carres.2013.06.024

Shi, S., Wang, W., Liu, L., Shu, B., Wei, Y., Jue, D., et al. (2016). Physico-chemical properties of longan fruit during development and ripening. *Sci. Hortic.* 207, 160–167. doi: 10.1016/j.scienta.2016.05.026

Tadeo, F. R., Terol, J., Rodrigo, M. J., Licciardello, C., and Sadka, A. (2020). "Chapter 12 - fruit growth and development," in *The genus citrus*. Eds. M. Talon, M. Caruso and F. G. Gmitter (Woodhead Publishing), 245–269. doi: 10.1016/B978-0-12-812163-4.00012-7

Torregrosa, L., Vialet, S., Adivèze, A., Iocco-Corena, P., and Thomas, M. R. (2015). "Grapevine (Vitis vinifera l.)," in *Agrobacterium protocols*, vol. 2 . Ed. K. Wang (New York, NY: Springer New York), 177–194.

van der Knaap, E., Chakrabarti, M., Chu, Y. H., Clevenger, J. P., Illa-Berenguer, E., Huang, Z., et al. (2014). What lies beyond the eye: the molecular mechanisms regulating tomato fruit weight and shape. *Front. Plant Sci.* 5. doi: 10.3389/fpls.2014.00227

Varandas, S., Teixeira, M. J., Marques, J. C., Aguiar, A., Alves, A., and Bastos, M. (2004). Glucose and fructose levels on grape skin: interference in lobesia botrana behaviour. *Anal. Chimica Acta* 513 (1), 351–355. doi: 10.1016/j.aca.2003.11.086

Vimolmangkang, S., Zheng, H., Peng, Q., Jiang, Q., Wang, H., Fang, T., et al. (2016). Assessment of sugar components and genes involved in the regulation of sucrose accumulation in peach fruit. *J. Agric. Food Chem.* 64 (35), 6723–6729. doi: 10.1021/acs.jafc.6b02159

Wang, W., Li, D., Quan, G., Wang, X., and Xi, Z. (2021). Effects of leaf removal on hexose accumulation and the expression of sugar unloading-related genes in syrah grapes. *Plant Physiol. Biochem.* 167, 1072–1082. doi: 10.1016/j.plaphy.2021.09.022

Wang, L., Zhang, S., Li, J., Zhang, Y., Zhou, D., Li, C., et al. (2022). Identification of key genes controlling soluble sugar and glucosinolate biosynthesis in Chinese cabbage by integrating metabolome and genome-wide transcriptome analysis. *Front. Plant Sci.* 13, 1043489. doi: 10.3389/fpls.2022.1043489

Xu, M., Sun, J., Yao, K., Wu, X., Shen, J., Cao, Y., et al. (2022). Nondestructive detection of total soluble solids in grapes using VMD-RC and hyperspectral imaging. *J. Food Sci.* 87 (1), 326–338. doi: 10.1111/1750-3841.16004

Yakushiji, H., Nonami, H., Fukuyama, T., Ono, S., Takagi, N., and Hashimoto, Y. (1996). Sugar accumulation enhanced by osmoregulation in Satsuma mandarin fruit. *J. Am. Soc. Hortic. Sci.* 121 (3), 466–472. doi: 10.21273/JASHS.121.3.466

Zhang, C., Fan, X., Liu, C., and Fang, J. (2021). Anatomical berry characteristics during the development of grape berries with different shapes. *Hortic. Plant J.* 7 (4), 295–306. doi: 10.1016/j.hpj.2021.04.002

Zhang, X. M., Liu, S. H., Du, L. Q., Yao, Y. L., and Wu, J. Y. (2019). Activities, transcript levels, and subcellular localizations of sucrose phosphate synthase, sucrose synthase, and neutral invertase and change in sucrose content during fruit development in pineapple (*Ananas comosus*). *J. Hortic. Sci. Biotechnol.* 94 (5), 573–579. doi: 10.1080/14620316.2019.1604169

Zhang, X., Liu, Q., Niu, S., Liu, C., Fan, X., Zhang, Y., et al. (2022b). Varietal differences among the fruit quality characteristic of 15 spine grapes (*Vitis davidii* foe euro x). *Hortscience* 57 (10), 1282–1288. doi: 10.21273/HORTSCI16702-22

Zhang, F., Zhong, H., Zhou, X., Pan, M., Xu, J., Liu, M., et al. (2022a). Grafting with rootstocks promotes phenolic compound accumulation in grape berry skin during development based on integrative multi-omics analysis. *Horticult. Res.* 9. doi: 10.1093/hr/uhac055

Zhong, H., Liu, Z., Zhang, F., Zhou, X., Sun, X., Li, Y., et al. (2022). Metabolomic and transcriptomic analyses reveal the effects of self- and hetero-grafting on anthocyanin biosynthesis in grapevine. *Horticult. Res.* 9. doi: 10.1093/hr/uhac103

Zhou, Y., He, W., Zheng, W., Tan, Q., Xie, Z., Zheng, C., et al. (2018). Fruit sugar and organic acid were significantly related to fruit mg of six citrus cultivars. *Food Chem.* 259, 278–285. doi: 10.1016/j.foodchem.2018.03.102

Zhu, J., Qi, J., Fang, Y., Xiao, X., Li, J., Lan, J., et al. (2018). Characterization of sugar contents and sucrose metabolizing enzymes in developing leaves of *Hevea brasiliensis*. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.00058

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 19 June 2023
DOI 10.3389/fpls.2023.1167789

OPEN ACCESS

EDITED BY
Yi Zheng,
Beijing University of Agriculture, China

REVIEWED BY
Xu Xiaozhao,
China Agricultural University, China
Hunseung Kang,
Chonnam National University, Republic of Korea

*CORRESPONDENCE
Nan Wang
nanwang2019@caf.ac.cn

RECEIVED 16 February 2023
ACCEPTED 30 May 2023
PUBLISHED 19 June 2023

CITATION
Zhang Y, Wang J, Ma W, Lu N, Fu P, Yang Y, Zhao L, Hu J, Qu G and Wang N (2023) Transcriptome-wide m6A methylation in natural yellow leaf of *Catalpa fargesii*. *Front. Plant Sci.* 14:1167789. doi: 10.3389/fpls.2023.1167789

Transcriptome-wide m6A methylation in natural yellow leaf of *Catalpa fargesii*

Yu Zhang[1], Junhui Wang[1], Wenjun Ma[1], Nan Lu[1], Pengyue Fu[1], Yingying Yang[1], Linjiao Zhao[2], Jiwen Hu[1], Guanzheng Qu[1] and Nan Wang[1*]

[1]State Key Laboratory of Tree Genetics and Breeding, Research Institute of Forestry, Chinese Academy of Forestry and Northeast Forestry University, Beijing, China, [2]Hekou Yao Autonomous County Forestry and Grassland Bureau, Hekou, China

N6-methyladenosine (m6A) is the most abundant internal modification in eukaryotic messenger RNA, and involved in various biological processes in plants. However, the distribution features and functions of mRNA m6A methylation have been poorly explored in woody perennial plants. In this study, a new natural variety with yellow-green leaves, named Maiyuanjinqiu, was screened from the seedlings of *Catalpa fargesii*. Based on the preliminary experiment, the m6A methylation levels in the leaves of Maiyuanjinqiu were significantly higher than those in *C. fargesii*. Furthermore, a parallel analysis of m6A-seq and RNA-seq was carried out in different leaf color sectors. The result showed that m6A modification were mostly identified around the 3'-untranslated regions (3'-UTR), which was slightly negatively correlated with the mRNA abundance. KEGG and GO analyses showed that m6A methylation genes were associated with photosynthesis, pigments biosynthesis and metabolism, oxidation-reduction and response to stress, etc. The overall increase of m6A methylation levels in yellow-green leaves might be associated with the decreased the expression of RNA demethylase gene *CfALKBH5*. The silencing of *CfALKBH5* caused a chlorotic phenotype and increased m6A methylation level, which further confirmed our hypothesis. Our results suggested that mRNA m6A methylation could be considered as a vital epigenomic mark and contribute to the natural variations in plants.

KEYWORDS

***Catalpa fargesii*, CfALKBH5, Epitranscriptomics, RNA methylation, N6-methyladenosine, yellow-green leaf**

Introduction

Leaf color variation is a widespread phenomenon in nature, and has been widely used as an ornamental trait in plant kingdoms. In recent years, researchers have drawn attention to the study of leaf color mechanism, because they are excellent materials for investigating pigment metabolism, chloroplast development, photosynthetic efficiency, etc (Wu et al.,

2007). Yellow-green leaves generally have the same genetic background but show two different leaf colors. Epigenetic modification might play an important role in the formation of leaf color (Li et al., 2015).

Approximately one hundred types of chemical modifications have been reported in eukaryotic RNAs, including N6-methyladenosine (m6A), N1-methyladenosine (m1A), 5-methylcytosine (m5C), 5-hydroxymethylcytosine (hm5C) and inosine (Frye et al., 2016; Zhao et al., 2017; Yang et al., 2020). It is worth noting that m6A modification accounts for 80% of all RNA methylation modifications in eukaryotic cells, and more than 50% of methylated nucleotides in polyA mRNA (Kierzek and Kierzek, 2003; Boccaletto and Bagiński, 2021). In recent years, m6A modification of mRNA in plants has been reported in virous species. m6A methylation could affect many biological processes by interfering with mRNA metabolism, including messenger RNA stability, pre-mRNA splicing, nuclear-to-cytoplasmic export and RNA translation efficiency (Wang X et al., 2014; Wang et al., 2015; Xiao et al., 2016; Roundtree et al., 2017; Liang et al., 2020; Luo et al., 2020; Tang et al., 2023). Recent studies suggested that m6A RNA methylation probably participate in the regulation of the plastid and thylakoid in the leaves (Li et al., 2014; Zhang et al., 2021). However, little is known about the pattern and functions of m6A methylation in regulating leaf color. In order to better understand our biological event, m6A modification will pave the way for further in-depth molecular mechanism analysis.

The m6A modification is regulated by three components: writers (written by methyltransferase), erasers (erased by demethylase), and readers (read by m6A-binding proteins) in plants (Yang et al., 2018; Liang et al., 2020; Zheng et al., 2020; Arribas-Hernandez and Brodersen, 2020; Arribas-Hernandez, 2023). In strawberry, m6A RNA methylation MTA and MTB were highly functionally conserved with dynamic modification of mRNA, and indispensable for fruit ripening (Zhou et al., 2021). In *Arabidopsis thaliana* and *Oryza sativa*, m6A methyltransferase AtFIP37 and OsFIP could influence the fate of shoot stem cells and early degeneration of microspores, respectively (Shen et al., 2016; Zhang et al., 2019). Other m6A demethylases, such as ALKBH2, ALKBH4B, ALKBH8B, ALKBH9B and ALKBH10B, have been discovered that mediate mRNA demethylation to influence the stability of target transcripts, consequently regulating flowering time, trichome and root development, fruit ripening and biotic and abiotic stress responses (Duan et al., 2017; Martínez-Pérez et al., 2017; Zhou et al., 2019; He et al., 2021; Amara et al., 2022; Huong et al., 2022; Han et al., 2023). In addition to methylases and demethylases, a class of m6A readers could fine-regulate biological function. For example, the YTH-domain proteins named EVOLUTIONARILIY CONSERVED C-TERMINAL REGION (ECT). Among which, ECT2, ECT3, ECT4, and ECT13 have been proven to function as m6A readers, playing critical roles in leaf and trichome formation, organogenesis, and nitrate signaling in *Arabidopsis* (Wei et al., 2018; Arribas-Hernandez and Brodersen, 2020; Scutenaire et al., 2018; Hou et al., 2021; Shao et al., 2021; Song et al., 2021). However, it is not clear whether there were other members of these three components involved in m6A modification, and whether these reported components are functionally conserved in different species also needed to be further explored. *Catalpa fargesii* Bur is widely distributed in the middle and western regions, and it is a famous timber and ornamental tree species in China (Wang et al., 2018). Maiyuanjinqiu with yellow-green leaf is a new variety (Identification code: 20150150) that was cultivated from the seedings of *C. fargesii*. The total chlorophyll and photosynthesis were significantly lower in Maiyuanjinqiu than those in *C. fargesi* (Wang et al., 2019). Based on the preliminary experiment, the m6A methylation level of total RNA in yellow leaves was higher than in green leaves. However, the key regulatory mechanism of m6A modifications and transcriptional regulation the leaf color formation was still poorly understood. In our study, transcriptome-wide m6A sequencing was performed in different leaf sectors of Maiyuanjinqiu and *C. fargesii*. Interestingly, the silencing of methyltransferase *CfALKBH5* caused a chlorotic phenotype and increased m6A methylation level, which further confirmed our hypothesis. Our results together indicate that m6A modification is an important epigenetic mark related to leaf color in *C. fargesii* and provide new insights into natural leaf color variation *via* epitranscriptome manipulation in forest breeding.

Materials and methods

Plant materials

Maiyuanjinqiu is a variety derived from *C. fargesii* seedlings. The plants samples used in this study were grown in the experimental field of Luoyang, Henan Province in China. All samples were collected, instantly frozen in liquid nitrogen, and stored at -80°C refrigerator. The different leaf color sectors of Maiyuanjinqiu and the responding sectors of *C. fargesii* were divided and collected according to the method of Wang et al. (2019), respectively. For each experiment, at least three biological repeats and technical repetitions were performed.

RNA extraction

Fresh leaf samples were grinded in mortars with liquid nitrogen, and extracted using RNAprep Pure Plant Plus Kit (TIANGEN, DP441, China), according to the manufacturer's protocol. Then, the total RNA of the samples was purified with the PrimeScript™ RT Master Mix Kit (TaKaRa, RR036A, China). The concentration and quality of total RNA were tested on a NanoDrop spectrophotometer (Thermo, USA) and gel electrophoresis.

The quantitative detection of m6A modification

Total RNA isolation and two rounds of PolyA$^+$ mRNA selection were performed to measure the global change of the m6A modification level. The change of global m6A levels in mRNA was measured by EpiQuik m6A RNA Methylation Quantification Kit (EpiGentek, P-9005, NY, USA) following the manufacturer's protocol.

Library construction, m6A-seq and RNA-seq

The poly(A) mRNAs were fragmented into 100-nucleotide (nt)-long oligonucleotides by using divalent cations. The cleaved RNA fragments were immunoprecipitated by incubating in IP buffer (50 mM Tris-HCl, 0.5% Igepal CA-630, 750 mM NaCl, and 0.5 g/L BSA) at 4 °C for 24 hours with an m6A-specific antibody (Synaptic Systems, Goettingen, Germany). According to the library preparation protocol, immunoprecipitated fragments (IP fractions) and input RNA (input control) libraries were constructed. The Illumina Novaseq 6000 platform was used for RNA sequencing at LC-BIO Biotech (Hangzhou, China). For each experiment, three biological repeats and technical repeats were performed.

m6A-seq data analyses

m6A-seq and RNA-seq were analyzed following the described previous method (Meyer and Jaffrey, 2014). Firstly, the FastQC tool was used to remove adaptor contamination, low-quality bases, and undetermined bases of raw data (Martin, 2011). Secondly, the clean data were mapped to the reference genome *via* HISAT2 (http://daehwankimlab.github.io/hisat2) (Kim et al., 2015). The mapping reads of input were then used to identify m6A peaks calling *via* the R package (Meng et al., 2014). HOMER and MEME online tools were used to identify conserved sequence motifs using Perl scripts. FPKM (fragments per kilobase of transcript per million mapped reads) was analyzed for the mRNA expression level of genes (Trapnell et al., 2012). The differentially expressed genes were selected with the standard of absolute |Log2FC| ≥ 1 and $P < 0.05$ *via* R package edgeR (Robinson et al., 2010). KEGG enrichment and GO analysis were performed using the LC-BIO online tools (https://www.omicstudio.cn/index).

qRT-PCR analysis

As described above, total RNAs were extracted from *C. fargesii* leaves. Extracted RNAs were synthesized first-strand cDNA using the PrimeScript™ RT Reagent Kit (TaKaRa, RR037A, China). qRT-PCR assays were analyzed on the Roche LightCycler® 480 Real-Time PCR system following the protocol described methods (Wang L et al., 2014). *CfActin* gene was used as internal reference genes (Supplementary Table S1) (Wang et al., 2018). The $2^{-\Delta\Delta CT}$ method was used to analyze the relative mRNA expression levels of genes (Livak and Schmittgen, 2001). For each experiment, three biological repeats and technical repetitions were performed.

Subcellular localization of CfALKBH5 protein

For the generation of the *CfALKBH5* and GFP fusion gene, the coding region of *CfALKBH5* lacking the stop codon was ligated into the linearized pBI121 vector with the green fluorescent protein (GFP) (Supplementary Table S2). The construct *CfALKBH5*-GFP and 35S::GFP (control) were introduced into *Nicotiana benthamiana* lower leaf epidermal cells according to the method of Zheng et al. (2005). After being transformed for 48 h in the dark, the tobacco cells were pelleted, resuspended in infiltrations solution, including 100 μM acetosyringone and 10 mM $MgCl_2$, and visualized using a confocal laser scanning microscopy (LSM 700, Zeiss, Jena, Germany).

VIGS vectors construction and transformation

The full-length coding sequence (CDS) of *CfALKBH5* was cloned using specific primers (Supplementary Table S2). The target gene fragment and pTRV2 were digested by *Xba*I and *Sac*I and then ligated by T4 ligase (Promega, M1801, China). The specific vector primers were used to detect the recombinant plasmid pTRV2-*CfALKBH5* by PCR (Supplementary Table S2). The constructed vectors were transferred into *Agrobacterium tumefaciens* strain EHA105 by freeze-thaw method (Fire et al., 1998). The *A. tumefaciens* strain was grown at 0.6 OD_{600} with shaking at 120 rpm at 28°C in ILuria-Bertani (LB) medium contained kanamycin (50 mg/L) and rifampicin (50 mg/L). The *A. tumefaciens* strain cells were harvested and re-suspended in an inoculation solution with acetosyringone (200 mM), including 10 mM $MgCl_2$, 10 mM MES, and 200 mM acetosyringone. Incubate for 3 hours at room temperature, a mixture (1:1 v/v) of induced EHA105 cultures containing pTRV1 plus pTRV2 vector (control), or pTRV1 plus pTRV2-*CfALKBH5*, was applied using approximately 450 μL of the mixture into the leaves surface of one-month seedlings. At least 3 biological replicates were performed with 15 transgenic plants per replicate. The number of plants for empty vector and untreated control were both same to that of infected plants. Only the most representative individuals were used to photograph. The infected leaves were collected for subsequent experiment analyses.

Pigment profiling

Fresh leaves were harvested and used to determine total chlorophyll and carotenoid contents following the protocol (Lichtenthaler and Wellburn, 1983). Total chlorophyll and carotenoid contents were extracted with 80% acetone at 4°C overnight in the dark, and then the solution was calculated at 645 nm, 663 nm, and 470 nm against the solvent (acetone) blank with a UV-Vis spectrophotometer (TU-1901, PERSEE, China). For each experiment, three biological repeats and technical repeats were performed.

Statistical analyses

SPSS19 software was used for statistical analyses, and the statistically significant (*, **, or ***; $P < 0.05$, $P < 0.01$, or $P <$

0.001) were considered as significant differences. For each experiment, three biological repeats and technical repetitions were performed.

Results

m6A methylation level is increased in the yellow leaf

Maiyuanjinqiu is a natural variation cultivar derived from *Catalpa fargesii*, whose leaves exhibit yellow-green character (Figure 1A). The yellow leaves of Maiyuanjinqiu are named Y1, while the light green leaves in the middle sectores of Maiyuanjinqiu are named Y2. The corresponding sectors of normal green leaves of the *C. fargesii* are named G1 and G2, respectively. To clarify whether m6A methylation is related to leaf color, we examined the overall total RNA m6A methylation levels in different leaf color sectors of Maiyuanjinqiu and the corresponding leaves of *C. fargesii*. As shown in Figure 1B, the yellow leaves of Maiyuanjinqiu exhibit hypermethylation compared with the green leaves of *C. fargesii*. The m6A methylation level of yellow leaves reached 1.59 times higher than that in green leaves in Maiyuanjinqiu, while the light green leaves in Maiyuanjinqiu were 1.47 times higher than that in the corresponding green leaves of *C. fargesii*, respectively. Based on the above results, we speculated that a correlation might exist between m6A modifications and leaf color.

The landscape of m6A methylation in the leaves of Maiyuanjinqiu and *C. fargesii*

To verify our hypothesis, the transcriptome-wide sequencing of m6A methylation was performed with three biological replicates for Maiyuanjinqiu and *C. fargesii* leaves. After eliminating adapter reads, unidentified bases, and low-quality bases, around 67.34%-89.04% of clean reads were uniquely mapped into the *Catalpa bungei* genome (unpublished) (Supplementary Table S3). The identified confident m6A peaks in three replicates with a high Pearson's correlation coefficient were used for further subsequent analyses (Supplementary Figure S1). On average, approximately 30,236 m6A sites of 21,343 genes from G1, around 31,375 m6A sites of 21,429 genes from G2, approximately 32,258 sites of 21,543 genes from Y1, and 32,198 m6A sites of 21,684 genes from Y2 were successfully identified, respectively (Supplementary Table S4). The whole-genome density of the m6A peaks and transcripts have different distribution modes mapped on the reference genome (Figure 2A). In addition, compared with the leaves of *C. fargesii*, the global hypermethylation of m6A was exhibited in Maiyuanjinqiu leaves (Figure 2B).

Furthermore, we discovered that the distribution of m6A peaks in protein-coding mRNA was mostly enriched around the 3'-UTR region in all groups and relatively less in the 5'-UTR (Figures 2C, D). In *C. fargesii* leaves, peak callings were almost equally distributed in transcription start (TSS) and end (TES) sites, whereas much more m6A peak callings were discovered in the TES regions than the TSS regions in Maiyuanjinqiu leaves (Supplementary Figure S2). Additionally, global m6A peaks were identified two conserved sequences: RRACY (R = G or A; Y = C or U) and UGUAH motifs (H = A, C, or U) using the MEME and HOMER suite (Figure 2E), which were very similar to the results identified in rice, tomato, and barley seedlings (Li et al., 2014; Zhang et al., 2019; Miao et al., 2020; Su et al., 2022).

To further understand the relationship between m6A methylation and leaf color, the yellow edge leaves of Maiyuanjinqiu (Y1) and the corresponding green leaves of *C. fargesii* (G1), which have the more representative color difference, were selected for subsequent analysis. We identified 1,533 green-leaf specific m6A peaks and 1,113 yellow-leaf specific peaks, as well as 15,911 common peaks (Figure 2F). KEGG enrichment analysis showed that m6A-modified genes unique to green leaves are mainly involved in flavone and flavonoid biosynthesis, anthocyanin biosynthesis, and starch and sucrose metabolism (Supplementary Figure S3A). However, m6A-harboring genes specific to yellow leaves are principally involved in photosynthesis-antenna proteins and phenylalanine metabolism (Supplementary Figure S3B).

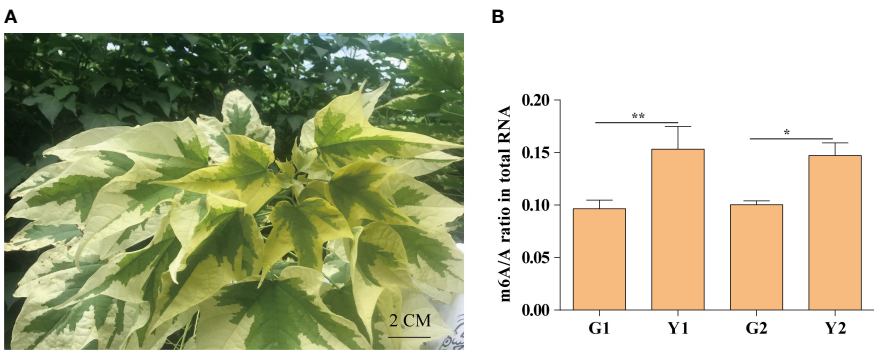


FIGURE 1
The m6A methylation levels in Maiyuanjinqiu and *C. fargesii*. **(A)** The leaf color character *of* Maiyuanjinqiu. **(B)** The levels of m6A methylation in the leaves of G1, G2, Y1 and Y2 sectors. The asterisks (**, ***) indicate significant differences ($P < 0.01$, $P<0.001$) between the two samples. Bar=2 cm.

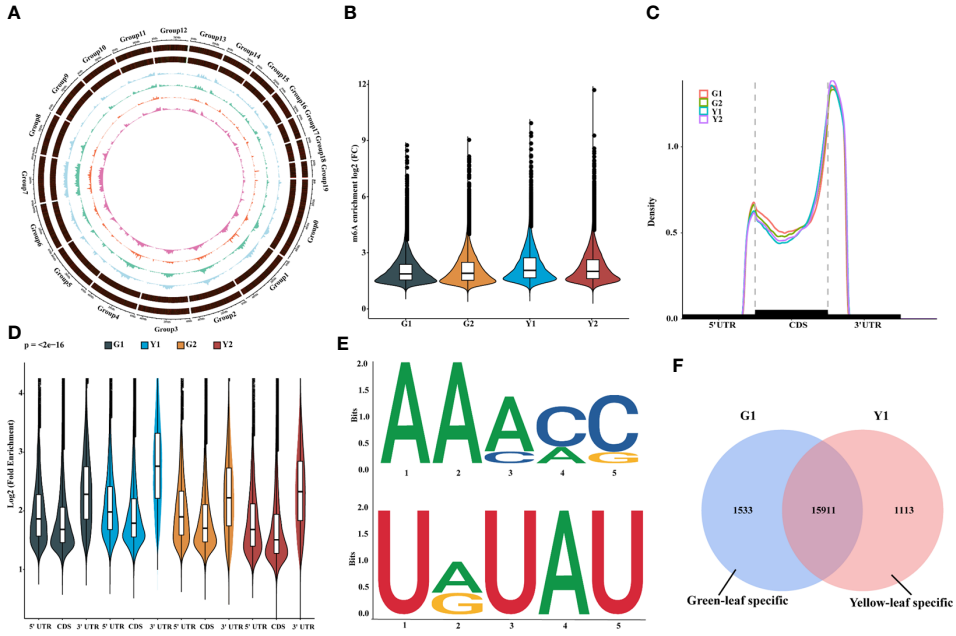


FIGURE 2
Overview of m6A methylome and transcriptome in leaves between Maiyuanjinqiu and *C.fargesii*. **(A)** Circos plot of the distance and density of m6A peak and expression abundance on *Catalpa bungei* chromosomes. **(B)** Violin plot showed the enrichment of m6A peaks in Maiyuanjinqiu and *Catalpa fargesii*. **(C)** Distribution of m6A peaks in transcript segments divided into 5'-UTR (untranslated region), CDS (coding sequence), and 3'-UTR. **(D)** Violin plot representing a comparison of the fold enrichment of m6A peaks in the different gene segments. **(E)** RRACH and URUAY motifs conserved sequence motifs for m6A-containing peak regions. **(F)** Venn diagram showing the overlap of m6A peaks in the yellow leaves and the corresponding sectors of *C. fargesii*.

The m6A levels are slightly negatively correlated with global gene expression

m6A deposition has been reported to influence mRNA abundance (Shen et al., 2016; Wang X et al., 2014; Duan et al., 2017; Zhao et al., 2017). To estimate the potential correlation between m6A mRNA methylation and gene expression during leaf yellowing, m6A-seq and RNA-seq were analyzed for the enhancement in the levels of m6A-containing transcripts and the global gene expression changes between two samples (Figure 3). A total of 9,586 transcripts with differential m6A levels (log2(FC) ≥1, *P* value < 0.05) between Y1 versus G1 were identified. Compared with Y1, 7,326 transcripts had higher m6A methylation levels than G1, while 2,260 transcripts had lower m6A methylation levels than G1 (Figure 3A). Among the 7,326 transcripts, only 669 showed higher expression levels, whereas 905 exhibited lower expression levels in Y1 versus G1 groups (Figure 3B). Accordingly, 719 and 372 displayed higher and lower expression levels among the 4,464 transcripts with lower m6A levels in Y2 compared to that in G2, respectively (Figure 3F, Supplementary Table S5). These data suggested that m6A methylation was slightly negatively correlated with the expression of the global transcripts.

Transcriptional differences in different leaf color sectors

To further explore the relationship between gene expression and leaf color, the yellow leaves (significant phenotypic characteristics) and the corresponding green leaves were selected for transcriptome differential analysis. A total of 3,308 differentially expressed genes (DEGs) were analyzed by the comparison in Y1 versus G1 group, with 772 up-regulated genes and 882 down-regulated genes, respectively (Figure 4A, Supplementary Table S6). Heatmap plot of the differentially expressed genes showed the consistency of three biological replicates and the significant difference between yellow and green leaves, respectively (Figure 4B). KEGG analysis showed that all significant DEGs were enriched in 129 metabolic or biological pathways. Remarkably, DEGs were highly enriched in the pathways of flavonoid biosynthesis, photosynthesis-antenna proteins, carotenoid biosynthesis, photosynthesis, anthocyanin biosynthesis, and RNA transport (Figure 4C). GO analysis revealed these DEGs are highly enriched in response to photosynthesis, light reaction, response to stress, heme binding, chlorophyll metabolism, and metal ion binding (Figure 4D). These results showed that the mRNA expression levels of photosynthesis, pigment, ion-binding and stress-related genes contributed to the formation of leaf color in Maiyuanjinqiu.

Correlation analysis of m6A-modified genes and DEGs expression on in yellow and green leaves

According to the crosstalk analysis of the m6A-seq and RNA-seq, we found 546 DEGs had differentially expressed peaks (DPs) in the comparison between yellow leaf and leaves. Among them, 140

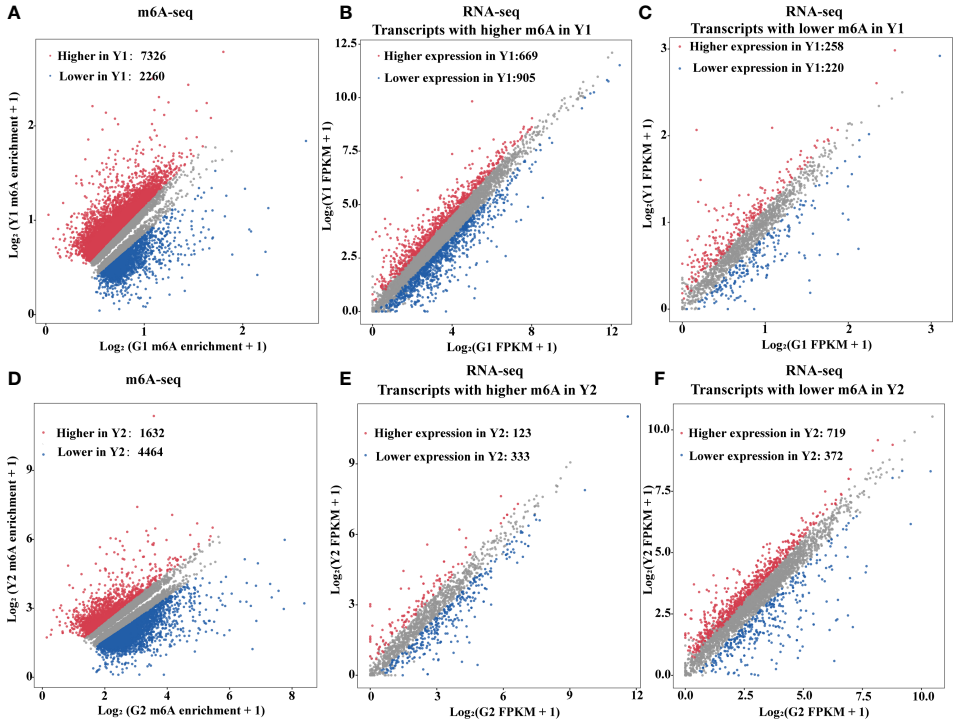


FIGURE 3
m6A RNA methylation is negatively associated with gene abundance. **(A)** Scatter plots exhibited transcripts abundance with differential m6A enrichment between the yellow leaves (Y1) and the corresponding green leaves (G1). The m6A-modified transcripts with substantially higher and lower peak enrichment in Y1 compared to G1 are highlighted in red and blue, respectively ($|Log_2FC| \geq 1$; *P* value < 0.05). **(B)** Expression of m6A-modified transcripts with markedly higher peak enrichment in Y1 than in G1. **(C)** Expression of m6A-modified transcripts with markedly lower peak enrichment in Y1 than in G1. **(D)** Scatter plots exhibited transcripts abundance with differential m6A enrichment between light green leaves (Y2) and the corresponding green leaves (G2). The m6A modified transcripts with substantially higher and lower peak enrichment in Y2 compared to G2 are highlighted in red and blue, respectively ($|Log_2FC| \geq 1$ and P < 0.05). **(E)** Expression of m6A-modified transcripts with markedly higher peak enrichment in Y2 than in G2. **(F)** Expression of m6A-modified transcripts with markedly lower peak enrichment in Y2 than in G2. RNA-seq was used for gene expression analysis. FPKM, fragments per kilobase of exon per million mapped fragments.

transcripts showed m6A hypermethylation and up-regulated expression (Hyper-up), 143 transcripts displayed m6A hypermethylation and down-regulated expression (Hyper-down), 143 transcripts exhibited m6A hypomethylation and up-regulated expression (Hypo-up), and 120 transcripts had m6A hypomethylation and down-regulated expression (Hypo-down) (Figure 5A, Supplementary Table S7). Furthermore, KEGG analysis of those yellow leaf yellow genes in the four-quadrant plot revealed that they were abundant in 69 pathways, including plant hormone signal transduction, peroxisome, porphyrin and chlorophyll metabolism, carotenoid biosynthesis, flavonoid biosynthesis and carbon fixation in photosynthetic organisms (Figure 5B). Meanwhile, GO analysis revealed that those genes were highly enriched in the oxidation-reduction process, response to stress, heme binding and metal ion binding (Figure 5C).

CfALKBH5 is a putative demethylase for m6A RNA that contributes to the yellow leaves

A large number of differential sites of m6A methylation modification were identified (Figure 5A), and the methylation level in yellow leaves was higher than that of green leaves. The differences in m6A levels might result from the regulation of m6A methyltransferases and demethylases. According to the existing literature report, the m6A writers mainly perform biological functions through a complex composed of methyltransferase METTL3 and METTL14 (Shi et al., 2019).The m6A writers mainly perform biological functions through a complex composed of methyltransferase METTL3 and METTL14 (Shi et al., 2019). The homologous genes of *METTL3* and *METTL14* were identified in the genome of *C. fargesii*, named *CfMTA1* and *CfMTA2*, respectively. And both of which contained a conserved MT-A70 domain (Supplementary Figure S4A). Based on the RNA-seq, the mRNA expression levels of *CfMTA1* and *CfMTA2* did not change significantly in the yellow leaves (Supplementary Figure S4B). It is speculated that m6A methyltransferase might not be an important reason for the overall change of m6A methylation modification level in yellow leaves. However, five demethylase *ALKBHs*, from *CfALKBH1* to *CfALKBH5*, were markedly down-regulated in the yellow leaves. The phylogenetic tree showed that CfALKBH3 and CfALKBH5 shared high similarity with AtALKBH10A and AtALKBH10B of *Arabidopsis* (Figure 6A). However, only the expression level of *CfALKBH5* was remarkably decreased in the yellow-green leaves compared with that of *C.*

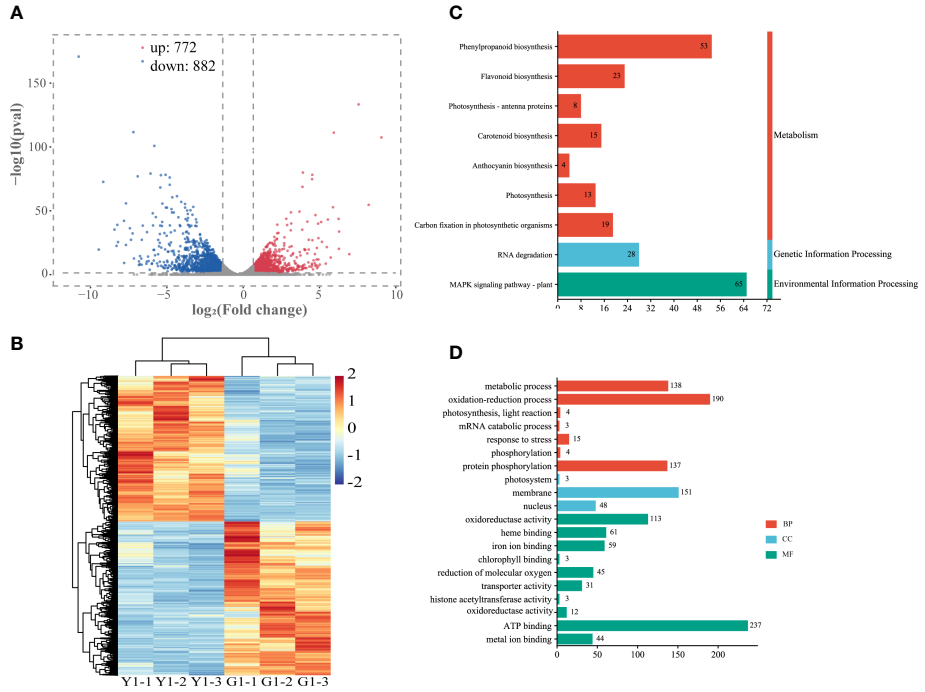


FIGURE 4
Analysis of differentially expressed genes (DEGs) in the leaves of Y1 and G1. **(A)** Volcano plot and Venn graph of regulated genes. **(B)** Heatmap showed the correlation of transcript expression between three biological replicates between Y1 and G1. **(C)** KEGG enrichment pathways of the DEGs between Y1 and G1. **(D)** GO enrichment of DEGs. BP, biological process; CC, cellular component; MF, molecular function.

fargesii among the five *CfALKBH* genes based on RNA-seq analysis results (Figure 6B, Supplementary Table S8). This result was further confirmed by qRT-PCR assay (Figure 6C). Therefore, *CfALKBH5* gene might regulate gene expression by regulating m6A methylation, and thus affect leaf color.

CfALKBH5 is a nuclear protein

To further investigate the nucleus localization of CfALKBH5, the plasmid encoding the 35S::CfALKBH5-GFP and 35S::GFP control were transiently expressed in tobacco epidermal cells. The fluorescence signals of CfALKBH5-GFP were localized in the nucleus of epidermal cells, whereas 35S-GFP was detected to be uniformly distributed throughout the tobacco epidermal cells (Figure 6D). These data indicated that CfALKBH5 was a nuclear-localized protein.

The suppression *CfALKBH5* results in a chlorotic phenotype in *C. fargesii*

The pTRV2-*CfALKBH5* (suppression) and pTRV2 vector (empty vector) were infected using an efficient virus-induced gene silencing (VIGS) system following to previous protocol (Chen et al., 2017; Yang et al., 2020), which has been generally used in functional characterization of genes in plenty of plant species (Chen et al.,

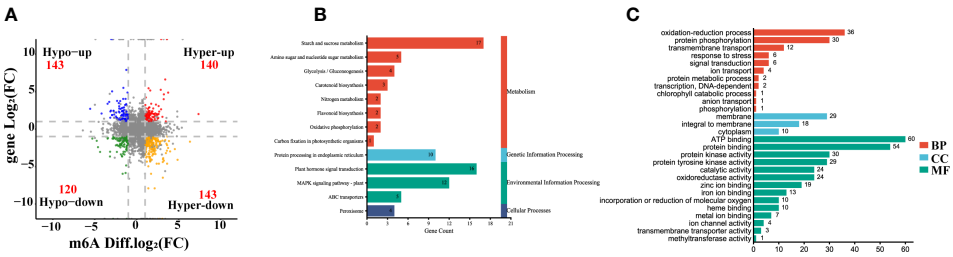


FIGURE 5
Integrating analysis of differentially m6A-modified and expressed transcripts between the yellow leaves (Y1) and the corresponding green leaves (G1). **(A)** Four-quadrant diagram exhibiting a relationship between m6A methylation and transcripts expression. Hyper-up, hypermethylation and up-regulated genes. Hyper-down, hypermethylation and down-regulated genes. Hypo-up, hypomethylation and upregulated genes. Hypo-down, hypomethylation and down-regulated genes. **(B)** KEGG pathways of differentially expressed genes (DEGs) with m6A peaks. **(C)** GO enrichment of DEGs with m6A peaks. MF, Molecular Function; CC, Cellular Component; BP, Biological Process.

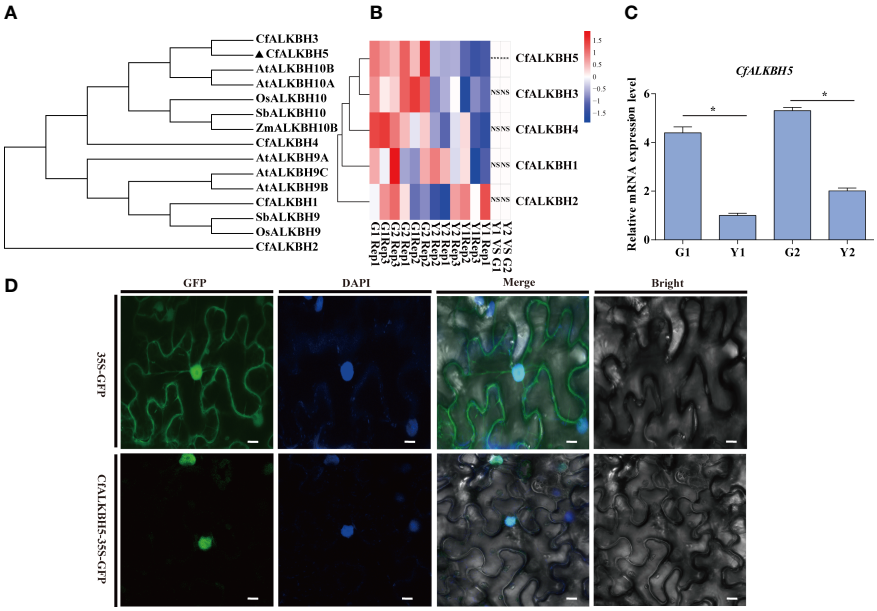


FIGURE 6
Phylogenetic tree construction, gene expression analyses and subcellular localization of m6A RNA demethylases ALKBHs. **(A)** The phylogenetic tree of ALKBHs in *C. fargesii* and other ALKBH proteins by the neighbor-joining method using MEGA 6. The sequences of the ALKBH proteins were achieved from the NCBI website (https://www.ncbi.nlm.nih.gov/protein/), and species names are shown below. Species names are abbreviated as follows: Cf, *C. fargesii*; At, *Arabidopsis thaliana*; OS, *Oryza sativa L*; Sb, *Sorghum bicolor*; Zm, *Zea mays*. **(B)** Heatmap analysis revealing the gene expression of *CfALKBH1-5* in G1, G2, Y1 and Y2 sectors based on RNA-seq. Data are presented as the mean of three independent biological replicates. **(C)** The relative mRNA expression of *CfALKBH5* in yellow leaves and green leaves as determined by qRT-PCR analysis. Asterisks indicate significant differences (*P < 0.05). **(D)** Subcellular localization of CfALKBH5 in tobacco leaf epidermal cells. The 35S::GFP and construct for 35S::CfALKBH5-GFP plasmid was introduced into tobacco epidermal cells by particle bombardment. The nuclei of the tobacco leaf epidermal cells were detected *via* DAPI staining. Bar=20 μm.

2017; Yang et al., 2020). To ensure the accuracy of this experiment, a minimum of three biological replicates were performed, which contain at least 15 transformed *C. fargesii* plants in each replicate. A chlorotic phenotype was observed in pTRV2-*CfALKBH5*-infected plants (Figure 7A, Supplementary Figure S5). The relative mRNA expression level of *CfALKBH5* was significantly lower in pTRV2-*CfALKBH5* infected plants than that in control plants, indicating the suppression was effective in our experiment (Figure 7B). Furthermore, the methylation level was detected by the m6A methylation kit, the result showed that *CfALKBH5* silencing significantly increased the m6A level in the total RNA (Figure 7C). The contents of chlorophyll and carotenoid were analyzed in the leaves of *CfALKBH5*-silenced and control seedlings. Compared with the control plants, the total chlorophyll content of leaves markedly decreased in *CfALKBH5* silencing plants, but the content of carotenoid did not significantly change (Figures 7D, E). The results showed that *CfALKBH5* might play a role in chlorophyll synthesis/metabolism by regulating the methylation levels.

The suppression of *CfALKBH5* could change the expression level of pigment biosynthesis and photosynthesis genes

Based on m6A-seq and RNA-seq data, the m6A levels of some key genes involved in chlorophyll metabolism, carotenoid biosynthesis, and photosynthesis were altered in yellow leaves (Supplementary Table S9). Therefore, qRT-PCR was performed to explore whether these genes also changed in pTRV2-*CfALKBH5*-infected and control plants. Compared with the control leaves, the expression levels of chlorophyll biosynthesis-related genes *CfHEMA*, *CfCAO* and *CfGLK* and photosynthesis-related genes *CfLHCA3* and *CfPsbP* were down-regulated in pTRV2-*CfALKBH5*-infected plants (Figure 8). In contrast, the gene expression of *CfPSY* and *CfVDE* involved in carotenoid synthesis were up-regulated in the *CfALKBH5*-silenced leaves (Figure 8). Taken together, *CfALKBH5* suppression indeed induced the change of gene expression profiles in pigment biosynthesis and photosynthesis.

Discussion

m6A methylation is an important mechanism of epigenetic regulation at the post-transcriptional level and has been studied in eukaryotic, including mice, human, Arabidopsis, rice, barley, tomato, sea buckthorn and apple (Meyer and Jaffrey, 2014; Zhao et al., 2014; Zhang et al., 2016; Zhou et al., 2019; Hu et al., 2021; Zhang et al., 2021; Hou et al., 2022; Su et al., 2022). However, as a valuable horticultural tree, the abundance, distribution, and function of m6A modification of *C. fargesii* remain unclear. In this research, we confirmed the presence of m6A modification of yellow-green leaves in *C. fargesii*. Firstly, we used an m6A RNA methylation quantification kit to detect the presence differences of

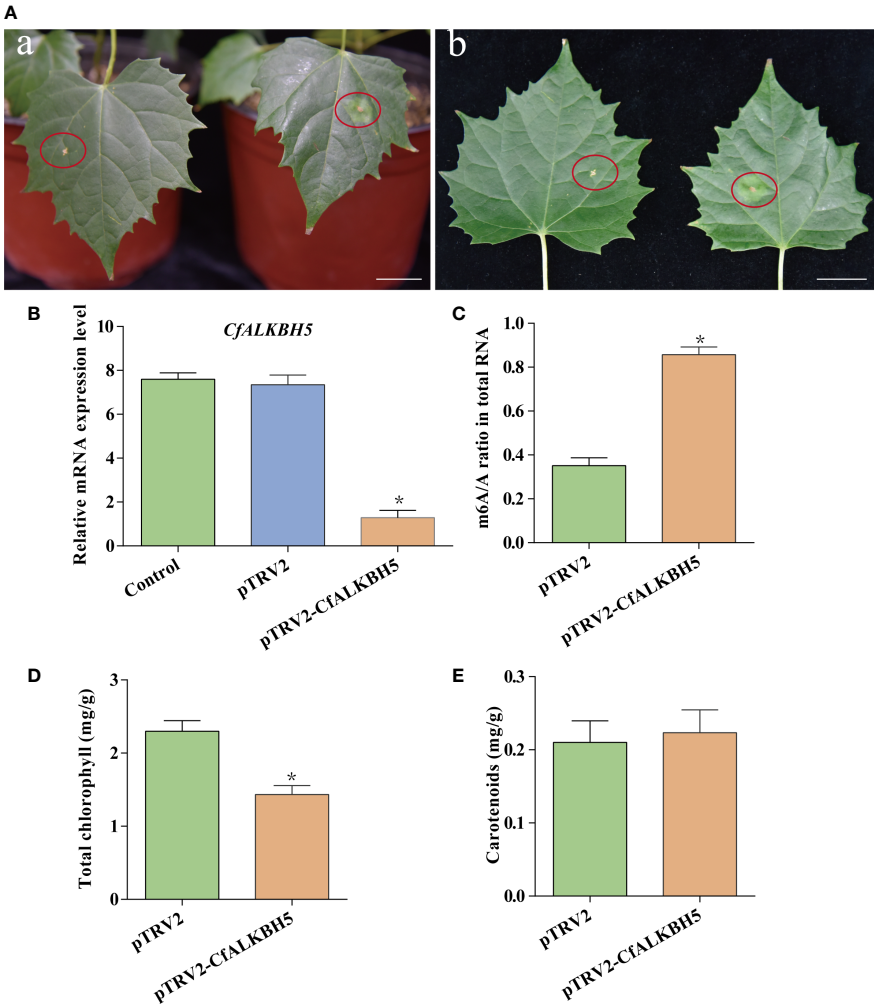


FIGURE 7
The function analysis of a putative methyltransferase CfALKBH5 by silencing *CfALKBH5* gene *in C.fargesii*. **(A)** VIGS-mediated CfALKBH5-silenced plants (Figure 7A) compared to pTRV2 vector-infected plants (control, Figure 7B). 15 plants were selected for photography in each treatment, and the most representative photos were selected. The red circle represents the infected area. Bar=2 cm. **(B)** The expression levels of *CfALKBH5* in pTRV2-*CfALKBH5* and control plants. The relative expression levels are shown as fold change values. Asterisk indicates a significant difference (*, $P< 0.05$). **(C)** The levels of m6A in total RNA of pTRV2- and pTRV2- *CfALKBH5*-treated plants, respectively. **(D, E)** The total chlorophyll and carotenoid contents were measured in the leaves of pTRV2-*CfALKBH5* and control plants.

m6A methylation levels in different color leaves of Maiyuanjinqiu and *C. fargesii*. Secondly, we identified the m6A methylation sites in *C. fargesii* genome based on Illumina Novaseq sequencing. In addition, m6A methylation levels presented differences in the total RNA between Maiyuanjinqiu and *C. fargesii*, indicating functional diversity of m6A modification and providing new insights into early post-transcriptional regulation during leaf chlorosis.

The transcriptome-wide distribution of m6A methylation is divergent in eukaryotes, including transposable element gene exons and transposons (Hu et al., 2021; Song et al., 2021). In this research, m6A-seq revealed m6A distribution pattern mostly in the 3'-UTR of mRNAs both in Maiyuanjinqiu and *C. fargesii* (Figure 2C), which is consistent with previous reports that have demonstrated m6A peaks near the 3'-UTR in plants (Shen et al., 2016; Zhang et al., 2019; Zhou et al., 2019; Luo et al., 2020; Hu et al., 2021), Notably, we discovered highly conserved RRACH motif and URUAY (R = A/G, Y = U/A) sequence as a previously identified plant-specific motif (Wei et al., 2018; Zhou et al., 2019; Luo et al., 2020; Hu et al., 2021). In addition, RRACH-like motifs, which were found in humans, mice, Arabidopsis, and maize (Dominissini et al., 2013; Meyer et al., 2012; Linder et al., 2015; Duan et al., 2017; Miao et al., 2020) (Figure 2E). In addition, our findings revealed an overall negative association between mRNA m6A modification and transcripts abundance (Figure 3), which was consistent with the findings of the ripening of tomato fruits (Zhou et al., 2019). Moreover, we found several genes with m6A modification are related to photosynthesis, which is similar to those in *Arabidopsis* (Li et al., 2014). The above results indicated that m6A modifications were conservative in plants. Unexpectedly, there are more increased m6A-methylated genes in Y1 but more decreased m6A-methylated genes in Y2. We speculated that the formation mechanisms of the natural yellow leaves are complicated, and there might be other underlying compensation mechanism might

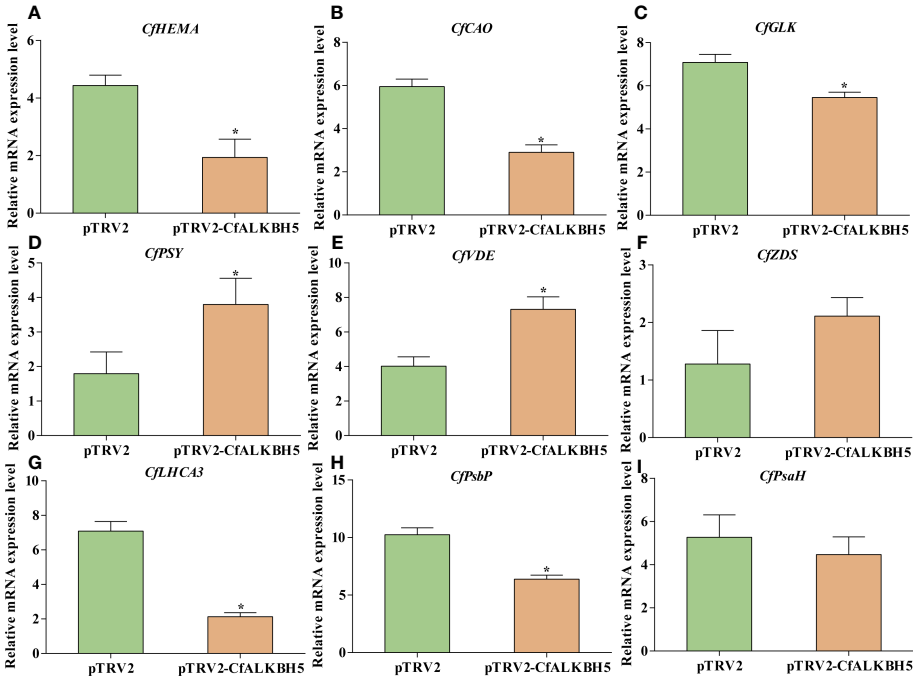


FIGURE 8
The expression comparison of important genes involved in chlorophyll metabolism, carotenoid biosynthesis and photosynthesis between pTRV2-*CfALKBH5*-and pTRV2-infected plants. **(A–C)** The expression levels of *CfHEMA*, *CfCAO*, *CfGLK* genes involved in chlorophyll metabolism in pTRV2-*CfALKBH5*-and pTRV2-infected plants. **(D–F)** The expression levels of *CfPSY*, *CfVDE*, *CfZDS* genes involved in carotenoid biosynthesis in pTRV2-*CfALKBH5*-and pTRV2-infected plants. **(G–I)** The expression levels of *CfLHCA3*, *CfPsbP*, *CfPsbH* genes involved in photosynthesis chlorophyll metabolism in pTRV2-*CfALKBH5*-and pTRV2-infected plants. Asterisks represent the significant difference, $P < 0.05$.

play a role in the light green leaf (Y2) of Maiyuanjinqiu. For example, in addition to demethyltransferase ALKBH5, whether there were other unknown m6A methyltransferases might be also involved in regulating the process deserve further investigation.

Several studies have demonstrated that RNA demethylases and methyltransferases may bind to and remove m6A marks, and play a vital role in regulating mRNA fate (Hu et al., 2019; Huang et al., 2020; Shao et al., 2021). In particular, several YT521-B homology (YTH) domain-containing proteins have been identified as m6A readers that regulate either negative or positive mRNA stability in animals and plants (Du et al., 2016; Shi et al., 2017; Huang et al., 2018; Wei et al., 2018; Baquero-Perez et al., 2019; Song et al., 2021). Moreover, several studies have demonstrated that the physiological functions of RNA methyltransferases were illustrated in different species. For example, the activity of mRNA methyltransferase METTL3 was influenced by SUMOylation in mammals (Du et al., 2018). In *Arabidopsis*, five potential RNA demethylases were identified, among which, ALKBH10B was involved in the floral transition and abiotic stress, while ALKBH9B participated in alfalfa mosaic virus infection (Duan et al., 2017; Martínez-Pérez, 2017; Han et al., 2023). In this study, an extensive search of the *Catalpa bungei* genome was performed and identified five putative RNA demethylases. Based on the transcriptome and qRT-PCR results, the expression levels of *CfALKBH5* reduced dramatically in yellow-green leaves of Maiyuanjinqiu compared with those of *C. fargesii*. In this study, CfALKBH5 nuclear localization suggested that CfALKBH5 is mainly responsible for nuclear RNA methylation rather than chloroplast or mitochondrial RNA methylation. In tomato, SlALKBH2 localized to the endoplasmic reticulum (ER) and regulated fruit ripening (Zhou et al., 2019). In *Arabidopsis*, CPSF30 localized to the nucleus and regulated the splicing of mRNA implicated in the salicylic acid pathway in response to external stimuli (Bruggeman et al., 2014). The localization of methylase might be related to its potential function, the functional analysis of *CfALKBH5* is very important for biological issues of concern.

In this study, the silencing of *CfALKBH5* significantly increased the abundance of m6A modification in total RNA in *C. fargesii*, indicating that CfALKBH5 indeed influences m6A modification. Moreover, CfALKBH5 suppression resulted in a chlorotic leaf phenotype and decreased chlorophyll contents. Moreover, qRT-PCR results showed that the expression levels of several critical genes involved in chlorophyll biosynthesis, carotenoid biosynthesis and photosynthesis had changed in pTRV2-*CfALKBH5*-infected plants. It is worth considering that although we detected changes of the expression levels of *CfPSY* and *CfVDE* genes in *pTRV2-CfALKBH5*-infected plants, there was no significant difference in carotenoid content. This means that more complex regulatory mechanisms might be involved in this process except for m6A methylation, which needed to be further studied. For example, whether these genes involved in chlorophyll metabolism, carotenoid biosynthesis, and photosynthesis are m6A modified and whether the m6A levels of these genes are altered in *CfALKBH5* knockdown plants remain unclear. In addition to m6A methylation, are there

other types of modification regulating leaf color of Maiyuanjinqiu also need to be studied.

Conclusion

In summary, we first investigated the differences of global m6A methylation levels in different leaf color sectors in woody plants. We found that the m6A methylated sites were mainly identified around the 3'-untranslated regions (3'-UTR), which was slightly negatively correlated with the mRNA abundance. Furthermore, the m6A methylation levels were significantly enhanced in the yellow sectors compared with the green sectors based on m6A RNA methylation quantification detection and m6A-seq data. Crosstalk analyses between peak and differential genes were conducted, KEGG and GO analyses showed that m6A modification genes were associated with photosynthesis, pigments biosynthesis and metabolism, oxidation-reduction and response to stress, etc. In addition, the overall increase in m6A methylation is associated with the decreased expression of RNA demethylase gene *CfALKBH5*. Interestingly, the silencing of demethylase *CfALKBH5* caused a chlorotic phenotype and increased m6A methylation level, suggesting the role of *CfALKBH5* in the formation of yellow-green leaves of Maiyuanjinqiu. Our results revealed that m6A modification could be considered as a vital epigenomic mark and contribute to the naturally variations in plants.

Data availability statement

The data presented in the study are deposited in the National Genomics Data Center repository (https://ngdc.cncb.ac.cn/gsa/browse/CRA010119), the accession number is CRA010119. The datasets generated and /or analyzed during the current study are available in NCBI with the accession number OQ401832 (https://www.ncbi.nlm.nih.gov/search/all/?term=OQ401832). Further inquiries can be directed to the corresponding author.

Author contributions

YZ performed the experiments and wrote the original draft. NW conceived the experiment and revised the paper. PF, WM and NL contributed to data analysis. YY, LZ, JH and GQ provided help in the experiments. JW provided the project funds, supervised the entire experiment and writing process. All authors contributed to the article and approved the submitted.

Funding

This work was supported by funding of the National Key R&D Program of China (2021YFD2200301).

Acknowledgments

We thank Dr.Yuan Cao and Dr. Liuqiang Wang from the State Key Laboratory of Tree Genetics and Breeding for their excellent technical assistance on the fluorescence observation and professional suggestions. We also thank A. prof Chao Wang from Northeast Forestry University for giving us a lot of help in experiments.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1167789/full#supplementary-material

References

Amara, U., Hu, J. Z., Cai, J., and Kang, H. (2023). FLK is an mRNA m6 a reader that regulates floral transition by modulating the stability and splicing of FLC in arabidopsis. *Mol. Plant* 16 (5), 919–929. doi: 10.1016/j.molp.2023.04.005

Amara, U., Shoaib, Y., and Kang, H. (2022). ALKBH9C, a potential RNA m6A demethylase, regulates the response of arabidopsis to abiotic stresses and abscisic acid. *Plant Cell Environ.* 45 (12), 3566–3581. doi: 10.1111/pce.14447

Arribas-Hernandez, L., and Brodersen, P. (2020). Occurrence and functions of m6A and other covalent modifications in plant mRNA. *Plant Physiol.* 182 (1), 79–96. doi: 10.1104/pp.19.01156

Baquero-Perez, B., Antanaviciute, A., Yonchev, I. D., Carr, I. M., Wilson, S. A., and Whitehouse, A. (2019). The Tudor SND1 protein is an m6A RNA reader essential for replication of kaposi's sarcoma-associated herpesvirus. *Elife* 8, e47261. doi: 10.7554/eLife.47261

Boccaletto, P., and Bagiński, B. (2021). Modomics: an operational guide to the use of the RNA modification pathways database. *Methods Mol Biol.* 2284, 481–505. doi: 10.1007/978-1-0716-1307-8_26

Bruggeman, Q., Garmier, M., De Bont, L., Soubigou-Taconnat, L., Mazubert, C., Benhamed, M., et al. (2014). The polyadenylation factor subunit CLEAVAGE AND

POLYADENYLATION SPECIFICITY FACTOR30: a key factor of programmed cell death and a regulator of immunity in arabidopsis. *Plant Physiol.* 165 (2), 732–746. doi: 10.1104/pp.114.236083

Chen, G., Liu, H., Wei, Q., Zhao, H., and Yu, Y. (2017). The acyl-activating enzyme phaae13 is an alternative enzymatic source of precursors for anthocyanin biosynthesis in petunia flowers. *J. Exp. Bot.* 68 (3), 457–467. doi: 10.1093/jxb/erw426

Dominissini, D., Moshitch-Moshkovitz, S., Salmon-Divon, M., Amariglio, N., and Rechavi, G. (2013). Transcriptome-wide mapping of N6 -methyladenosine by m6A-seq based on immunocapturing and massively parallel sequencing. *Nat. Protocol.* 8, 176. doi: 10.1038/nprot.2012.148

Du, H., Zhao, Y., He, J., Zhang, Y., Xi, H., Liu, M., et al. (2016). YTHDF2 destabilizes m6A-containing RNA through direct recruitment of the CCR4-NOT deadenylase complex. *Nat. Commun.* 7 (1), 12626. doi: 10.1038/ncomms12626

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

Fire, A., Xu, S., Montgomery, M., KOSTAS, S. A., DRIVER, S. E., and MELLO, C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391 (6669), 806–811. doi: 10.1038/35888

Frye, M., Jaffrey, S. R., Pan, T., Rechavi, G., and Suzuki, T. (2016). RNA Modifications: what have we learned and where are we headed? *Nat. Rev. Genet.* 17 (6), 365–372. doi: 10.1038/nrg.2016.47

Han, R., Shoaib, Y., Cai, J., and Kang, H. (2023). ALKBH10B-mediated m6A demethylation is crucial for drought tolerance by affecting mRNA stability in arabidopsis. *Environ. Exp. Bot.* 209, 105306. doi: 10.1016/j.envexpbot.2023.105306

He, Y. J., Li, L. L., Yao, Y. X., Li, Y. L., Zhang, H. Q., and Fan, M. (2021). Transcriptome-wide N6-methyladenosine (m6A) methylation in watermelon under CGMMV infection. *BMC Plant Biol.* 21 (1), 1–14. doi: 10.1186/s12870-021-03289-8

Hou, N., Li, C., He, J., Liu, Y., Yu, S., Malnoy, M., et al. (2022). MdMTA-mediated m6A modification enhances drought tolerance by promoting mRNA stability and translation efficiency of genes involved in lignin deposition and oxidative stress. *New Phytol.* 234 (4), 1294–1314. doi: 10.1111/NPH.18069

Hou, Y., Sun, J., Wu, B., Gao, Y., Nie, H., Nie, Z., et al. (2021). CPSF30-l-mediated recognition of mRNA m6A modification controls alternative polyadenylation of nitrate signaling-related gene transcripts in arabidopsis. *Mol. Plant* 14 (4), 688–699. doi: 10.1016/j.molp.2021.01.013

Hu, J. Z., Manduzio, S., and Kang, H. (2019). Epitranscriptomic RNA methylation in plant development and abiotic stress responses. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.00500

Hu, J., Cai, J., Park, S. J., Lee, K., Li, Y., Chen, Y., et al. (2021). N6-methyladenosine mRNA methylation is important for salt stress tolerance in arabidopsis. *Plant J.* 106 (6), 1759–1775. doi: 10.1111/tpj.15270

Huang, H., Weng, H., Sun, W., Qin, X., Shi, H., Wu, H., et al. (2018). Recognition of RNA n 6-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. *Nat. Cell Biol.* 20 (3), 285–295. doi: 10.1038/s41556-018-0045-z

Huang, H., Weng, H., and Chen, J. (2020). The biogenesis and precise control of RNA mA methylation. *Trends Genet.* 36 (1), 44–52. doi: 10.1016/j.tig.2019.10.011

Huong, T. T., Yang, Z., Ngoc, L. N. T., and Kang, H. (2022). ALKBH8B, a putative RNA demethylase, plays a role in the response of arabidopsis to salt stress and abscisic acid. *J. Plant Biol.* 65 (4), 319–330. doi: 10.1007/s12374-022-09351-8

Kierzek, E., and Kierzek, R. (2003). The thermodynamic stability of RNA duplexes and hairpins containing N6-alkyladenosines and 2-methylthio-N6-alkyladenosines. *Nucleic Acids Res.* 31 (15), 4472–4480. doi: 10.1093/nar/gkg633

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12 (4), 357–360. doi: 10.1038/nmeth.3317

Li, Y., Wang, X., Li, C., Hu, S., Yu, J., and Song, S. (2014). Transcriptome-wide N6-methyladenosine profiling of rice callus and leaf reveals the presence of tissue-specific competitors involved in selective mRNA modification. *RNA Biol.* 11 (9), 1180–1188. doi: 10.4161/rna.36281

Li, Y., Zhang, Z., Wang, P., Wang, S. A., Ma, L., Li, L., et al. (2015). Comprehensive transcriptome analysis discovers novel candidate genes related to leaf color in a *Lagerstroemia indica* yellow leaf mutant. *Genes Genom.* 37, 851–863. doi: 10.1038/nrm.2016.132

Liang, Z., Riaz, A., Chachar, S., Ding, Y., Du, H., and Gu, X. (2020). Epigenetic modifications of mRNA and DNA in plants. *Mol. Plant* 13 (1), 14–30. doi: 10.1016/j.molp.2019.12.007

Lichtenthaler, H. K., and Wellburn, A. R. (1983). Determinations of total carotenoids and chlorophylls a and b of leaf extracts in different solvents. *Analysis* 11 (5), 591–592. doi: 10.1042/bst0110591

Linder, B., Grozhik, A. V., Olarerin-George, A. O., Meydan, C., Mason, C. E., and Jaffrey, S. R. (2015). Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat. Methods* 12 (8), 767–772. doi: 10.1038/nmeth.3453

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* 25 (4), 402–408. doi: 10.1006/meth.2001.1262

Luo, J. H., Wang, Y., Wang, M., Zhang, L. Y., Peng, H. R., Zhou, Y. Y., et al. (2020). Natural variation in RNA m6A methylation and its relationship with translational status. *Plant Physiol.* 182 (1), 332–344. doi: 10.1104/pp.19.00987

Martin, M. (2011). Cut adapt removes adapter sequences from high-throughput sequencing reads, embnet. 17, 10-12. *EMBnet J.* 17 (1), 10–12. doi: 10.14806/ej.17.1.200

Martínez-Pérez, M., Aparicium, 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. *PNAS* 114 (40), 10755–10760. doi: 10.1073/pnas.1703139114

Meng, J., Lu, Z. L., Liu, H., Zhang, L., Zhang, S. W., Chen, Y. D., et al. (2014). A protocol for RNA methylation differential analysis with MeRIP-seq data and exomePeak R/Bioconductor package. *Methods* 69 (3), 274–281. doi: 10.1016/j.ymeth.2014.06.008

Meyer, K. D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C. E., and Jaffrey, S. R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* 149 (7), 1635–1646. doi: 10.1016/j.cell.2012.05.003

Meyer, K. D., and Jaffrey, S. R. (2014). The dynamic epitranscriptome: N6? methyladenosine and gene expression control. *Nat. Rev. Mol. Cell Biol.* 15 (5), 313–326. doi: 10.1038/nrm3785

Miao, Z., Zhang, T., Qi, Y., Song, J., Han, Z., and Ma, C. (2020). Evolution of the RNA N6-methyladenosine methylome mediated by genomic duplication. *Plant Physiol.* 182 (1), 345–360. doi: 10.1104/pp.19.00323

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616

Roundtree, I. A., Luo, G. Z., Zhang, Z., Wang, X., Zhou, T., Cui, Y., et al. (2017). YTHDC1 mediates nuclear export of N6-methyladenosine methylated mRNAs. *ELife* 6, e31311. doi: 10.7554/eLife.31311.001

Scutenaire, J., Deragon, J. M., Jean, V., Benhamed, M., Raynaud, C., Favory, J. J., et al. (2018). The YTH domain protein ECT2 is an m6A reader required for normal trichome branching in arabidopsis. *Plant Cell.* 30 (5), 986–1005. doi: 10.1105/tpc.17.00854

Shao, Y., Wong, C. E., Shen, L., and Yu, H. (2021). N6-methyladenosine modification underlies messenger RNA metabolism and plant development. *Curr. Opin. Plant Biol.* 63, 102047. doi: 10.1016/j.pbi.2021.102047

Shen, L., Liang, Z., Gu, X., Chen, Y., Teo, Z. W., Hou, X., et al. (2016). N6-methyladenosine RNA modification regulates shoot stem cell fate in arabidopsis. *Dev. Cell.* 38 (2), 186–200. doi: 10.1016/j.devcel.2016.06.008

Shi, H., Wang, X., Lu, Z., Zhao, B. S., Ma, H., Hsu, P. J., et al. (2017). YTHDF3 facilitates translation and decay of N6-methyladenosine-modified RNA. *Cell Res.* 27 (3), 315–328. doi: 10.1038/cr.2017.15

Shi, H., Wei, J., and He, C. (2019). Where, when, and how: context-dependent functions of RNA methylation writers, readers, and erasers. *Mol. Cell* 74 (4), 640–650. doi: 10.1016/j.molcel.2019.04.025

Song, P., Yang, J., Wang, C., Lu, Q., Shi, L., Tayier, S., et al. (2021). Arabidopsis N6-methyladenosine reader CPSF30-l recognizes FUE signals to control polyadenylation site choice in liquid-like nuclear bodies. *Mol. Plant* 14 (4), 571–587. doi: 10.1016/j.molp.2021.01.014

Su, T., Fu, L., Kuang, L., Chen, D., Zhang, G., Shen, Q., et al. (2022). Transcriptome-wide m6A methylation profile reveals regulatory networks in roots of barley under cadmium stress. *J. Hazard. Mater.* 423, 127140. doi: 10.1016/j.jhazmat.2021.127140

Tang, J., Chen, S., and Jia, G. (2023). Detection, regulation, and functions of RNA N6 methyladenosine modification in plants. *Plant Commun.* 4, 100546. doi: 10.1186/s12864-019-5776-0

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 (3), 562–578. doi: 10.1038/nprot.2012.016

Wang, X., Lu, Z., Gomez, A., Hon, G. C., Yue, Y., Han, D., et al. (2014). N6 -methyladenosine dependent regulation of messenger RNA stability. *Nature* 505 (7481), 117–120. doi: 10.1038/nature12730

Wang, L., Wang, C., Wang, D., and Wang, Y. (2014). Molecular characterization and transcript profiling of *NAC* genes in response to abiotic stress in *Tamarix hispida*. *Tree Genet. Genomes* 10 (1), 157–171. doi: 10.1007/s11295-013-0672-2

Wang, X., Zhao, B. S., Boxuan, S., Roundtree, I. A., Lu, Z. K., Han, D. L., et al. (2015). N6-methyladenosine modulates messenger RNA translation efficiency. *Cell* 161 (6), 1388–1399. doi: 10.1016/j.cell.2015.05.014

Wang, N., Zhu, T., Lu, N., Wang, Z., Yang, G., Wang, J., et al. (2019). Quantitative phosphoproteomic and physiological analyses provide insights into the formation of the variegated leaf in *Catalpa fargesii*. *IJMS* 20 (8), 1895. doi: 10.3390/ijms20081895

Wang, Z., Zhu, T., Ma, W., Wang, N., Qu, G., Zhang, S., et al. (2018). Genome-wide analysis of long non-coding RNAs in *Catalpa bungei* and their potential function in floral transition using high-throughput sequencing. *BMC Genet.* 19 (1), 1–16. doi: 10.1186/s12863-018-0671-2

Wei, L. H., Song, P., Wang, Y., Lu, Z., Tang, Q., Yu, Q., et al. (2018). The m6A reader ECT2 controls trichome morphology by affecting mRNA stability in arabidopsis. *Plant Cell.* 30 (5), 968–985. doi: 10.1105/tpc.17.00934

Wu, Z., Zhang, X., He, B., Diao, L., Sheng, S., Wang, J., et al. (2007). A chlorophyll-deficient rice mutant with impaired chlorophyllide esterification in chlorophyll biosynthesis. *Plant Physiol.* 145 (1), 29–40. doi: 10.1104/pp.107.100321

Xiao, W., Adhikari, S., Dahal, U., Chen, Y. S., Hao, Y. J., Sun, B. F., et al. (2016). Nuclear m6A reader YTHDC1 regulates mRNA splicing. *Mol. Cell.* 61 (6), 507–519. doi: 10.1016/j.molcel.2016.03.004

Yang, Y., Hsu, P. J., Chen, Y. S., and Yang, Y. G. (2018). Dynamic transcriptomic m6 a decoration: writers, erasers, readers and functions in RNA metabolism. *Cell Res.* 28 (6), 616–624. doi: 10.1038/s41422-018-0040-8

Yang, W., Meng, J., Liu, J., Ding, B., Tan, T., Wei, Q., et al. (2020). The N1-methyladenosine methylome of petunia mRNA. *Plant Physiol.* 183 (4), 1710–1724. doi: 10.1104/pp.20.00382

Zhang, G., Lv, Z., Diao, S., Liu, H., Duan, A., He, C., et al. (2021). Unique features of the m6A methylome and its response to drought stress in sea buckthorn (*Hippophae rhamnoides* linn.). *RNA Biol.* 18 (sup2), 794–803. doi: 10.1080/15476286.2021.1992996

Zhang, C., Samanta, D., Lu, H., Bullen, J. W., Zhang, H., Chen, I., et al. (2016). Hypoxia induces the breast cancer stem cell phenotype by HIF-dependent and ALKBH5-mediated m6A-demethylation of NANOG mRNA. *PNAS* 113 (14), E2047–E2056. doi: 10.1073/pnas.160288311

Zhang, F., Zhang, Y. C., Liao, J. Y., Yu, Y., Zhou, Y. F., Feng, Y. Z., et al. (2019). The subunit of RNA N6-methyladenosine methyltransferase OsFIP regulates early degeneration of microspores in rice. *PloS Genet.* 15 (5), e1008120. doi: 10.1371/journal.pgen.1008120

Zhao, B. S., Roundtree, I. A., and He, C. (2017). Post-transcriptional gene regulation by mRNA modifications. *Nat. Rev. Mol. Cell Biol.* 18 (1), 31–42. doi: 10.1038/nrm.2016.132

Zhao, X., Yang, Y., Sun, B. F., Shi, Y., Yang, X., Xiao, W., et al. (2014). FTO-dependent demethylation of N6-methyladenosine regulates mRNA splicing and is required for adipogenesis. *Cell Res.* 24 (12), 1403–1419. doi: 10.1038/cr.2014.151

Zheng, H., Camacho, L., Wee, E., Batoko, H., Legen, J., Leaver, C. J., et al. (2005). A rab-e GTPase mutant acts downstream of the rab-d subclass in biosynthetic membrane traffic to the plasma membrane in tobacco leaf epidermis. *Plant Cell.* 17 (7), 2020–2036. doi: 10.1105/tpc.105.031112

Zheng, H., Li, S., Zhang, X., and Sui, N. (2020). Functional implications of active N6-methyladenosine in plants. *Front. Cell Dev. Biol.* 8. doi: 10.3389/fcell.2020.00291

Zhou, L., Tang, R., Li, X., Tian, S., Li, B., and Qin, G. (2021). N6-methyladenosine RNA modification regulates strawberry fruit ripening in an ABA-dependent manner. *Genome Biol.* 22 (1), 1–32. doi: 10.1186/s13059-021-02385-0

Zhou, L., Tian, S., and Qin, G. (2019). RNA Methylomes reveal the m6A-mediated regulation of DNA demethylase gene SlDML2 in tomato fruit ripening. *Genome Biol.* 20 (1), 1–23. doi: 10.1186/s13059-019-1771-7

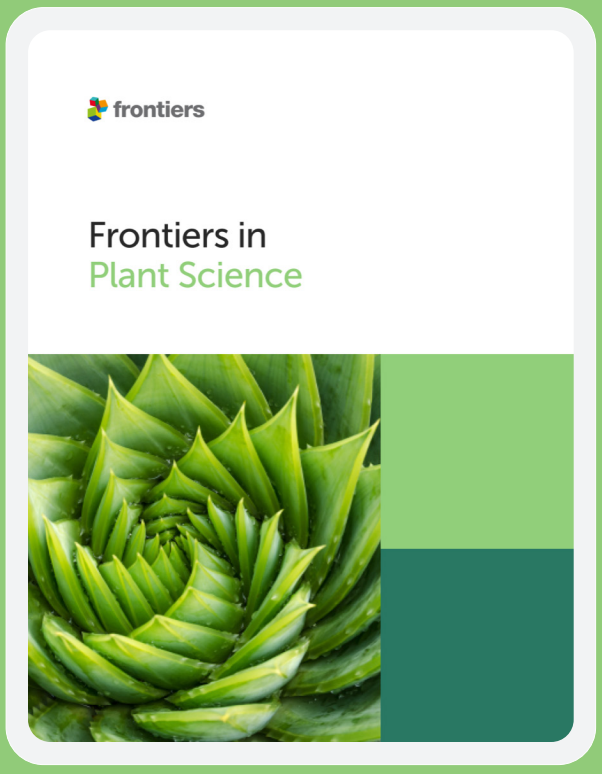
Frontiers in
Plant Science

Cultivates the science of plant biology and its applications

The most cited plant science journal, which advances our understanding of plant biology for sustainable food security, functional ecosystems and human health.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact