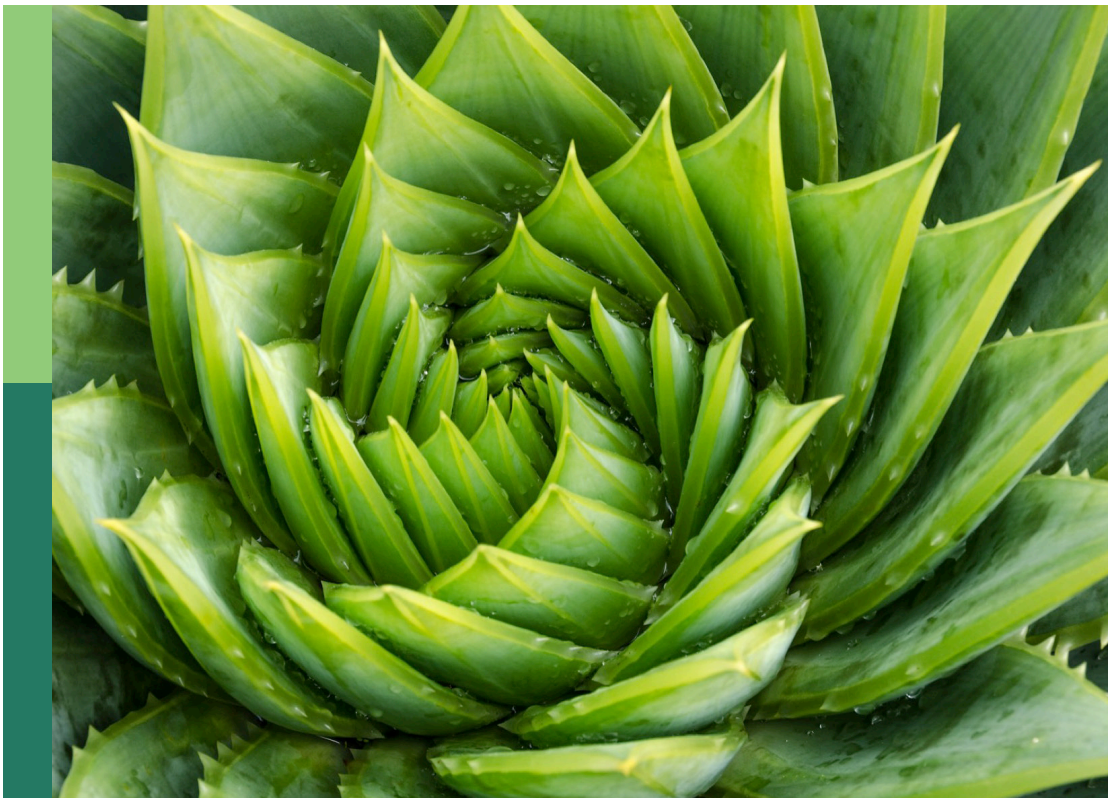
frontiers | Research Topics

Genetic and molecular mechanisms of important agronomic traits in forage grasses

Edited by
Wengang Xie, Linkai Huang and Mingshu Cao

Published in
Frontiers in Plant Science



ISSN 1664-8714
ISBN 978-2-8325-2469-5
DOI 10.3389/978-2-8325-2469-5

About Frontiers

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

Frontiers journal series

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

Dedication to quality

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

What are Frontiers Research Topics?

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

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

Genetic and molecular mechanisms of important agronomic traits in forage grasses

Topic editors

Wengang Xie — Lanzhou University, China

Linkai Huang — Sichuan Agricultural University, China

Mingshu Cao — AgResearch Ltd, New Zealand

Citation

Xie, W., Huang, L., Cao, M., eds. (2023). *Genetic and molecular mechanisms of important agronomic traits in forage grasses*. Lausanne: Frontiers Media SA.

doi: 10.3389/978-2-8325-2469-5

Table of contents

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 13 July 2022
doi: 10.3389/fpls.2022.919177

Comparative Physiological and Transcriptome Analysis Reveal the Molecular Mechanism of Melatonin in Regulating Salt Tolerance in Alfalfa (*Medicago sativa* L.)

Shuxia Li[1,2], *Yuan Wang*[1], *Xueqin Gao*[1,2], *Jian Lan*[1,2] *and Bingzhe Fu*[1,2]*

[1] School of Agriculture, Ningxia University, Yinchuan, China, [2] Ningxia Grassland and Animal Husbandry Engineering Technology Research Center, Yinchuan, China

As a high-quality legume forage, alfalfa is restricted by various abiotic stresses during its growth and development. Melatonin is a multifunctional signaling molecule that involves in plant defense against multiple stresses. However, little is known about its downstream signaling pathway and regulatory mechanisms in salt stress of alfalfa. In this study, we investigated the protective effects and key regulatory pathways of melatonin on alfalfa under salt tolerance. The results showed that melatonin promoted the growth of alfalfa seedlings under salt stress, as demonstrated by higher plant height, leaf area, and fresh weight. Melatonin treatment resulted in an increase in the photosynthetic capacity and starch content of alfalfa. Moreover, melatonin decreased cell membrane damage and reactive oxygen species (ROS) accumulation by enhancing antioxidant defense activity under salt stress conditions. Transcriptome sequencing (RNA-seq) analysis revealed that melatonin mainly induced the transcription of genes involved in Ca^{2+} signaling (cyclic nucleotide gated channel, *CNGCs*; cam modulin/calmodulin-like protein, *CAM/CMLs* and calcium-dependent protein kinase, *CDPKs*), starch and sucrose metabolism (α-amylase, *AMYs*; β-amylase, *BAMs*; starch synthase, *SSs* and sucrose synthase, *SUSs*), plant hormone signal transduction (auxin/indole acetic acid protein, *AUX/IAAs*; ABA receptor, *PYL4*; protein phosphatase 2C, *PP2Cs*; scarecrow-like protein, *SCLs* and ethylene-responsive transcription factor 1B, *ERF1B*), and key transcription factors (*C3Hs*, *MYBs*, *ERFs*, and *WRKYs*). Specifically, we focused on starch and sucrose metabolism and plant hormone signal transduction pathways. The interactions between melatonin and other phytohormones occurred *via* regulation of the expression of genes involved in hormone signaling pathways. In addition, melatonin increased the contents of endogenous melatonin, auxin, gibberellic acid (GA_3), salicylic acid, brassinosteroids, and ethylene, while decreasing the abscisic acid content under salt stress. In summary, this study established a regulatory network for melatonin-induced key signaling pathways and functional genes under salt stress and provided a theoretical basis for salt tolerance breeding in alfalfa.

Keywords: melatonin, salt stress, transcriptome, plant hormone, signal transduction, *Medicago sativa*

OPEN ACCESS

Edited by:
Wengang Xie,
Lanzhou University, China

Reviewed by:
Min Zhong,
South China Agricultural University, China
Gang Nie,
Sichuan Agricultural University, China

***Correspondence:**
Bingzhe Fu
Fbzhe19@163.com

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

Received: *13 April 2022*
Accepted: *08 June 2022*
Published: *13 July 2022*

Citation:
Li S, Wang Y, Gao X, Lan J and Fu B (2022) Comparative Physiological and Transcriptome Analysis Reveal the Molecular Mechanism of Melatonin in Regulating Salt Tolerance in Alfalfa (Medicago sativa L.). Front. Plant Sci. 13:919177. doi: 10.3389/fpls.2022.919177

INTRODUCTION

Soil salinization is a severely adverse environmental factor that threatens agricultural sustainability and global food security (Cheeseman, 2016). Salinization limits plant growth, development, productivity and quality, especially in arid, and semiarid regions (Pang et al., 2016). Unfortunately, salt stress influences more than 800 Mha of all irrigated lands worldwide, and this problem continues to worsen (Yang and Guo, 2018; FAO, 2019). The excessive salts in the soil solution cause osmotic, ionic, and secondary stresses on plants. Salt stress is commonly caused by high concentrations of sodium ions (Na^+) and chloride ions (Cl^-) in soil (Ismail and Horie, 2017). Plants have evolved diverse mechanisms to cope with salt stress. Currently, several regulatory components have been found to play important roles in signal transduction under salt stress, including ion balance regulation, reactive oxygen species (ROS) homeostasis modulation, and plant hormone metabolism (Horvath et al., 2015).

Melatonin (N-acetyl-5-methoxytryptamine) is a highly conserved and ubiquitous indoleamine molecule in the plant kingdom (Paredes et al., 2009). Numerous studies have revealed that melatonin, as a growth regulator or defense response biostimulator, is involved in regulating various biological processes in plants, including seed germination, seedling growth, root morphology, leaf senescence, nutrient absorption, floral transition, fruit ripening, as well as multiple abiotic and biotic stress responses (Fan et al., 2018; Arnao and Hernandez-Ruiz, 2019; Sun et al., 2021). The biosynthesis of melatonin in higher plants begins with tryptophan and is catalyzed by the following four successive enzyme reactions: Tryptophan decarboxylase (TDC), tryptamine 5-hydroxylase (T5H), serotonin N-acetyltransferase (SNAT), and N-acetylserotonin methyltransferase (ASMT)/caffeic acid O-methyltransferase (COMT) (Back et al., 2016).

Melatonin is a powerful antioxidant and free radical scavenger that suppresses peroxidative metabolism in plants under abiotic stresses (Arnao and Hernandez-Ruiz, 2019; Sharma et al., 2020). Experimental evidence proves that melatonin can exert its antioxidant capacity by scavenging excessive ROS and reactive nitrogen species (RNS) directly (Martinez et al., 2018). Melatonin increases a wide spectrum of plant stress tolerance indirectly by enhancing the activities of antioxidant enzymes (SOD, CAT, POD and GPX, etc.), improving photosynthesis and redox homeostasis, activating downstream signals and regulating the expression of stress-responsive genes (Zhang and Zhang, 2014; Liu et al., 2020). Melatonin interacts with various phytohormones such as auxin (IAA), ethylene (ETH), jasmonic acid (JA), salicylic acid (SA), abscisic acid (ABA), and brassinosteroids (BRs) to participate in stress responses (Arnao and Hernandez-Ruiz, 2019). Melatonin promotes ethylene biosynthesis by regulating the expression of *ACS1* gene; thus, enhancing the salt tolerance of grapevines (Xu et al., 2019). In the recent years, there has been much research on melatonin in salt stress and in a variety of plant species (Wei et al., 2015; Yan et al., 2019; Liu et al., 2020). Nevertheless, the signaling network of melatonin-mediated salt stress responses in plants is complex and remains largely obscure, especially in forages.

Alfalfa (*Medicago sativa* L.) is an important perennial legume species that is widely cultivated around the world. The planting area of alfalfa is approximately 32.2 Mha worldwide, 11% (∼3.77 Mha) of which are planted in China (Shi et al., 2017). Alfalfa provides high protein content and excellent palatability of forage for animals and improves soil fertility. It is therefore regarded as the "king of forages." However, alfalfa quality and yield are severely constrained by adverse environmental factors, such as soil salinity and limited water supplies in agriculture (Singer et al., 2018). Therefore, combining physiological, biochemical, and molecular approaches to improve the salt tolerance of alfalfa is significant for the production of high-quality alfalfa on saline-alkali land.

In this study, we investigated the physiological and molecular mechanisms of melatonin-mediated salt stress tolerance in alfalfa plants. The results reveal that melatonin could improve alfalfa salt tolerance by enhancing photosynthetic capacity and the antioxidant defense system and reducing membrane damage and ROS accumulation at the physiological level. Based on transcriptome data, we focused on the key genes involved in starch and sucrose metabolism and hormone signaling pathways. The results of this study aid in understanding the molecular mechanism underlying melatonin-mediated salt stress tolerance in alfalfa.

MATERIALS AND METHODS

Plant Materials and Treatments

Alfalfa (*Medicago sativa* L. cv. zhongmu 1#) seeds were surface-sterilized with 75% (v/v) ethanol and 5% sodium hypochlorite solution and then germinated on wet filter paper in Petri dishes for 6 days at 25°C. For the hydroponic experiment, seedlings with uniform growth were transferred into plastic containers filled with Hoagland solution in a growth chamber (25°C, 16-h light/8-h dark cycle, 60% relative humidity). For soil culture, the seedlings were transplanted into plastic pots that were 7 cm in diameter and 8.5 cm in depth with vermiculite and watered with Hoagland nutrient solution. After 3 weeks of incubation, the plants with uniform growth were treated with the following different solutions: (i) Control (CK), Hoagland nutrient solution alone; (ii) salt stress (S), Hoagland nutrient solution with 150-mM NaCl; (iii) melatonin (M), Hoagland nutrient solution plus melatonin; and (iv) salt stress with melatonin (SM), Hoagland nutrient solution with 150-mM NaCl plus melatonin. The NaCl treatment concentration was set according to the literatures (Benabderrahim et al., 2020; Yu et al., 2021). Each treatment contained 36 pots with one plant per pot. Alfalfa plants were treated with 10-μM melatonin in hydroponic experiments to observe the root phenotype. In soil culture pre-experiments, four different concentrations (0, 20, 50, and 100 μM) of melatonin were applied to choose the appropriate concentration of melatonin treatment. After treatment, the leaf samples were subsequently frozen in liquid nitrogen and stored at −80°C. The experiment was repeated 3 times.

Determination of Physiological Parameters and Hormone Contents

According to the growth phenotype of alfalfa, the representative individuals from each group were photographed on the 15th day of salt treatment, and the plant height, shoot fresh weight, and leaf area were measured. The net photosynthetic rate (Pn) of the third fully expanded leaf was determined using a Li-6400 portable photosynthesis system (Lincoln, NE, USA) according to the manufacturer's instructions. The photosynthetic photon flux density (PPFD) and the external CO_2 concentration were set at 1,000 μmol m^{-2} s^{-1} and 400 μmol mol^{-1}, respectively. Electrolyte leakage was detected by conductometer according to the methods of a previous study (Li et al., 2019a). The malondialdehyde (MDA) content was quantified by a thiobarbituric acid method (Puckette et al., 2007). The content of hydrogen peroxide (H_2O_2) and superoxide anion radical ($O_2^{\cdot-}$) were spectrophotometrically determined as described by Jiang and Zhang (2001). The activities of superoxide dismutase (SOD), peroxidase (POD), glutathione S-transferase (GST), ascorbate peroxidase (APX), as well as the levels of proline, total soluble sugar, and total starch were determined using the corresponding plant kits (Solarbio Science & Technology Co., Ltd. Beijing, China) according to the manufacturer's instructions.

The Na^+ and K^+ contents were measured following our previous work (Li et al., 2019b). Approximately 20–50 mg of dry powder sample from each treatment was weighed and dissolved with acetic acid solution. Then, the Na^+ and K^+ contents of the extract were determined using a flame photometer.

The content of endogenous hormones, including melatonin, ABA, IAA, gibberellic acid (GA_3), SA, BR, and ETH, in the alfalfa leaf samples were measured using HPLC–MS/MS analysis as described in the previous studies, with some modifications (Liu et al., 2018; Zhang et al., 2018).

The RNA Sequencing and Bioinformatics Analysis

The RNA samples for the transcriptome were extracted from the leaves of CK-, S-, M-, and SM-treated alfalfa plants. Each treatment was represented by three biological replicates of leaf samples. The total RNA was isolated and purified using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) following the manufacturer's procedure. The RNA amount and purity of each sample were quantified using a NanoDrop ND-1000 (NanoDrop, Wilmington, DE, USA). The integrity of RNA was assessed by a Bioanalyzer 2100 (Agilent, CA, USA) with RIN > 7.0 and confirmed by electrophoresis with denaturing agarose gel. Sequencing library construction and sequencing were performed using Illumina Novaseq™ 6000 (LC-Bio Technology CO., Ltd., Hangzhou, China) following the vendor's recommended protocol. The adaptor contamination reads were removed using Cutadapt software (version 1.9), and the low-quality bases and undetermined bases were removed from the raw data (Wang et al., 2012). HISAT2 software (version: 2.0.4) was used to map clean reads to the genome (Kim et al., 2015).

The mapped reads of each sample were assembled using StringTie with the default parameters. Then, all transcriptomes from all samples were merged to reconstruct a comprehensive transcriptome using gffcompare software. StringTie and Ballgown software were used to estimate the expression levels of all transcripts and perform gene expression level analysis by calculating the fragments per kilobase of transcripts per million (FPKM) mapped reads. Genes with fold change more than 2 or < 0.5 and $p < 0.05$ were defined as differentially expressed. The differential expression analysis was conducted using the DESeq2 R package (Pertea et al., 2015). The Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of differentially expressed genes (DEGs) were conducted using the R package (Minoru et al., 2008; Young et al., 2010).

Quantitative Real-Time PCR (qRT-PCR) Validation

Total RNA was isolated and purified using TRIzol reagent (Invitrogen, Carlsbad, CA, USA). The RNA samples were reverse-transcribed with HiScript II Q RT SuperMix for qPCR (+gDNA wiper) (Vazyme, Nanjing, China) according to the manufacturer's protocol. Twenty genes were randomly selected for the qRT-PCR assay to validate the RNA-seq data. The qRT-PCR was performed using SYBR Premix on a Bio-Rad CFX 96 RT-PCR System (Bio-Rad, Inc., CA, USA). Three independent biological replicates and three replicate reactions for each sample were performed in qRT-PCR assays. The alfalfa *β-actin* gene was used as an internal control for expression analysis (Long et al., 2014). The relative expression levels of selected genes were calculated by the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). The gene-specific primer pairs used in qRT-PCR are listed in **Supplementary Table 1**.

Statistical Analysis

Statistical analysis was performed using SPSS 20.0 statistical software (SPSS Inc., Chicago, IL, USA). Significant differences in the physiological parameters were analyzed by one-way ANOVA followed by Duncan's test. The differences between individual means were considered significant at $p < 0.05$.

RESULTS

Detection of the Alleviating Effects of Exogenous Melatonin on Salt Stress in Alfalfa

To evaluate whether melatonin can improve the salt stress tolerance of alfalfa, the seedlings were subjected to NaCl and melatonin treatment, and the phenotypic traits were observed (**Figure 1A** and **Supplementary Figures 1A,D**). Compared with the control plants, 10-μM melatonin treatment significantly increased the root length and root dry weight of alfalfa in hydroponic experiments. Salt stress inhibited the root length and root dry weight, while the application of melatonin significantly improved these parameters (**Supplementary Figures 1B,C**). In the soil culture pre-experiments, the plant height and Pn of the plants treated with 50-μM melatonin were higher than those of other treatments (0, 20 and 100 μM) under salt

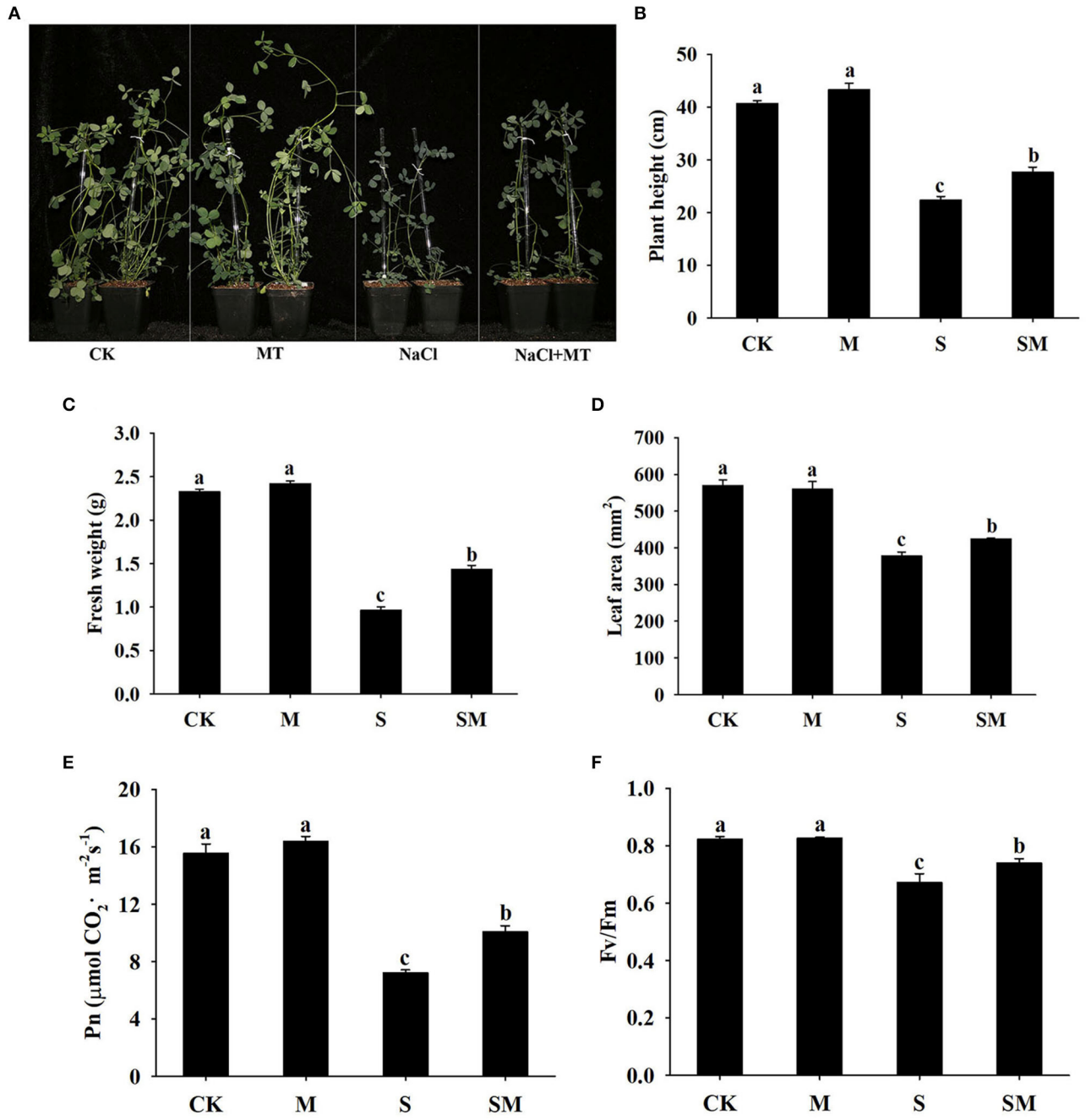


FIGURE 1 | Effect of exogenous melatonin on the phenotype traits and photosynthetic capacity of alfalfa seedlings after 15 day of 150-mM NaCl treatment. **(A)** The phenotype of a representative individual from each treatment. **(B)** Plant height. **(C)** Fresh weight. **(D)** Leaf area. **(E)** Net photosynthetic rate (Pn). **(F)** Fv/Fm. The data are means ± SE ($n = 6$) and different letters are significantly different ($p < 0.05$).

stress (**Supplementary Figures 1E,F**). Therefore, we selected 50-μM as the optimal concentration for melatonin treatment in subsequent experiments. Compared with the control, salt stress obviously inhibited the growth of alfalfa seedlings, as evidenced by the declines in plant height, leaf area, fresh weight, and photosynthetic parameters: Pn and Fv/Fm; however, the application of melatonin dramatically mitigated this effect in the soil culture experiments (**Figures 1B–F**). In addition, melatonin treatment significantly decreased the electrolyte leakage and MDA content under salt stress, but had no obvious effect on control plants (**Figures 2A,B**). Salt treatment resulted in a significant increase in H_2O_2 and $O_2^{\cdot-}$ contents, while the melatonin treatment reduced them by 0.8 and 0.7 times, respectively (**Figures 2C,D**). Plants have involved

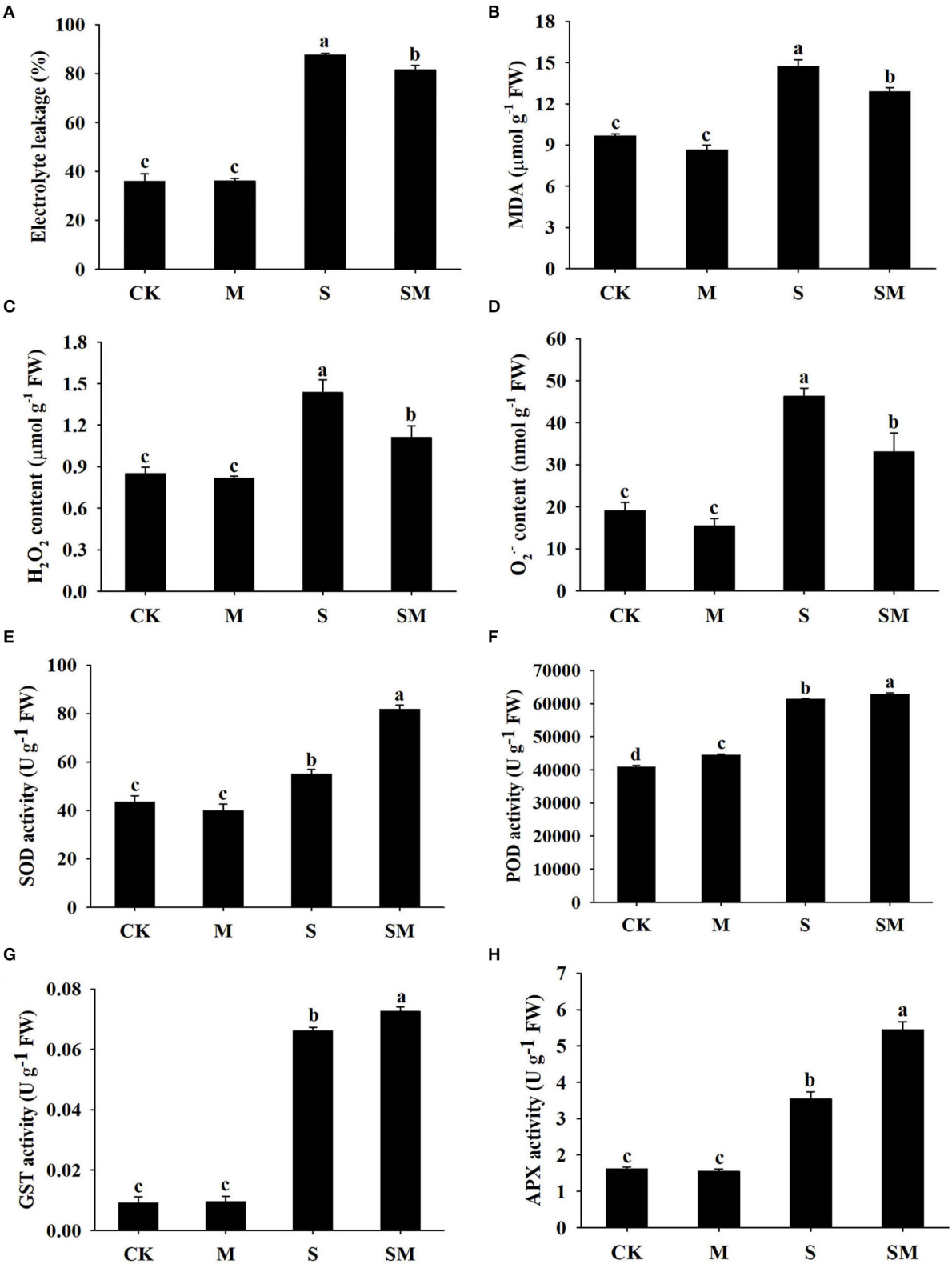


FIGURE 2 | Effect of exogenous melatonin on **(A)** electrolyte leakage, **(B)** MDA content, **(C)** H_2O_2 content, **(D)** $O_2^{\cdot-}$ content, **(E)** SOD, **(F)** POD, **(G)** GST, and **(H)** APX of alfalfa seedlings after 15 days of 150-mM NaCl treatment. Data are means ± SE ($n = 3$) and different letters are significantly different ($p < 0.05$).

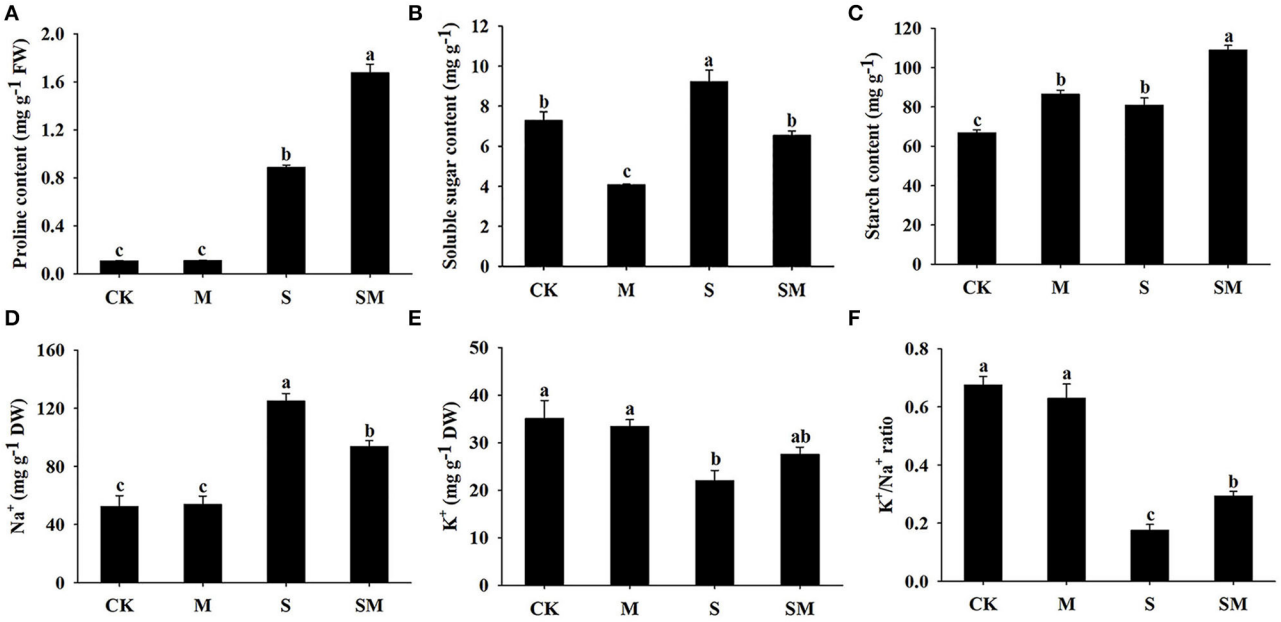


FIGURE 3 | Effect of exogenous melatonin on the contents of **(A)** proline, **(B)** soluble sugar, **(C)** starch, **(D)** Na^+, **(E)** K^+, and **(F)** K^+/Na^+ ratio of alfalfa seedlings after 15 days of 150-mM NaCl treatment. The data are means ± SE ($n = 3$) and different letters are significantly different ($p < 0.05$).

a complex antioxidant defense system to cope with abiotic stress-triggered oxidative damage, including some functionally correlated antioxidant enzymes. As shown in **Figures 2E–H**, salt stress significantly increased the SOD, POD, APX, and GST enzymatic activities in comparison with control plants. However, the activity of these enzymes was remarkably higher in melatonin-treated plants than in non-treated plants under salt stress.

Melatonin had no significant effect on the proline content under control conditions. When salt stress was applied, the proline content was significantly increased, while the melatonin-treated plants exhibited a higher level of proline accumulation (**Figure 3A**). The soluble sugar content of the seedlings was greatly increased under salt stress, while melatonin treatment decreased the soluble sugar content in both the control and salt-treated plants by 44.08 and 28.81%, respectively (**Figure 3B**). As shown in **Figure 3C**, melatonin treatment increased the starch content in alfalfa plants before and after salt stress. Salt stress resulted in significantly higher starch level than that in the control plants. The K^+/Na^+ ratio, a key indicator of salt stress mitigation, was significantly decreased under salt stress conditions due to elevated Na^+ and reduced K^+. However, melatonin treatment markedly increased the salt-induced K^+/Na^+ ratio in alfalfa (**Figures 3D–F**).

Transcriptome Sequencing (RNA-Seq) and DEGs Analysis

To further elucidate the molecular mechanism underlying melatonin-induced salt stress tolerance in alfalfa, transcriptome analyses of alfalfa leaves were performed on 12 samples (CK-1/-2/-3, M-1/-2/-3, S-1/-2/-3, and SM-1/-2/-3) using the Illumina Novaseq^{TM} 6000 sequencing platform. A total of 92.78 GB raw reads were obtained from all tested samples. More than 6.34 GB average clean data were obtained for each RNA-seq sample, and the valid data ratio was above 92% (Q20 > 99.94% and Q30 > 98.60%) (**Supplementary Table 2**). The valid data were mapped to the reference genome by HISAT software, and over 90.40% mapped reads were obtained. A total of 164,632 transcripts were generated in the alfalfa transcriptome data, and 126,412 and 49,751 transcripts were annotated in GO and KEGG, respectively (**Supplementary Table 3**). Three independent biological replicates of each treatment were clustered in a PCA, and the repeatability within each group (CK, M, S, and SM) and the discrimination between groups were good (**Supplementary Figure 2**). The transcriptome data were reliable enough to support further analysis.

Compared with the control plants, 2,485 (1,485 upregulated and 1,000 downregulated genes), 6,120 (3,267 upregulated genes and 2,853 downregulated genes) and 4,992 (2,264 upregulated and 2,728 downregulated genes) DEGs were identified in the "M *vs.* CK," "S *vs.* CK," and "SM *vs.* CK" comparisons, respectively. A total of 1,774 DEGs, including 732 upregulated and 1,042 downregulated genes, were obtained in "SM *vs.* S" (**Figure 4A** and **Supplementary Table 4**). A Venn diagram showed that 723 DEGs were affected by both melatonin and salt stress (the intersection of "M *vs.* CK" and "S *vs.* CK"). Melatonin affected 135 DEGs under both the control and salt stress conditions (the intersection of "M *vs.* CK" and "SM *vs.* S"). Additionally, 67 DEGs were found in the intersection of "M *vs.* CK," "S *vs.*

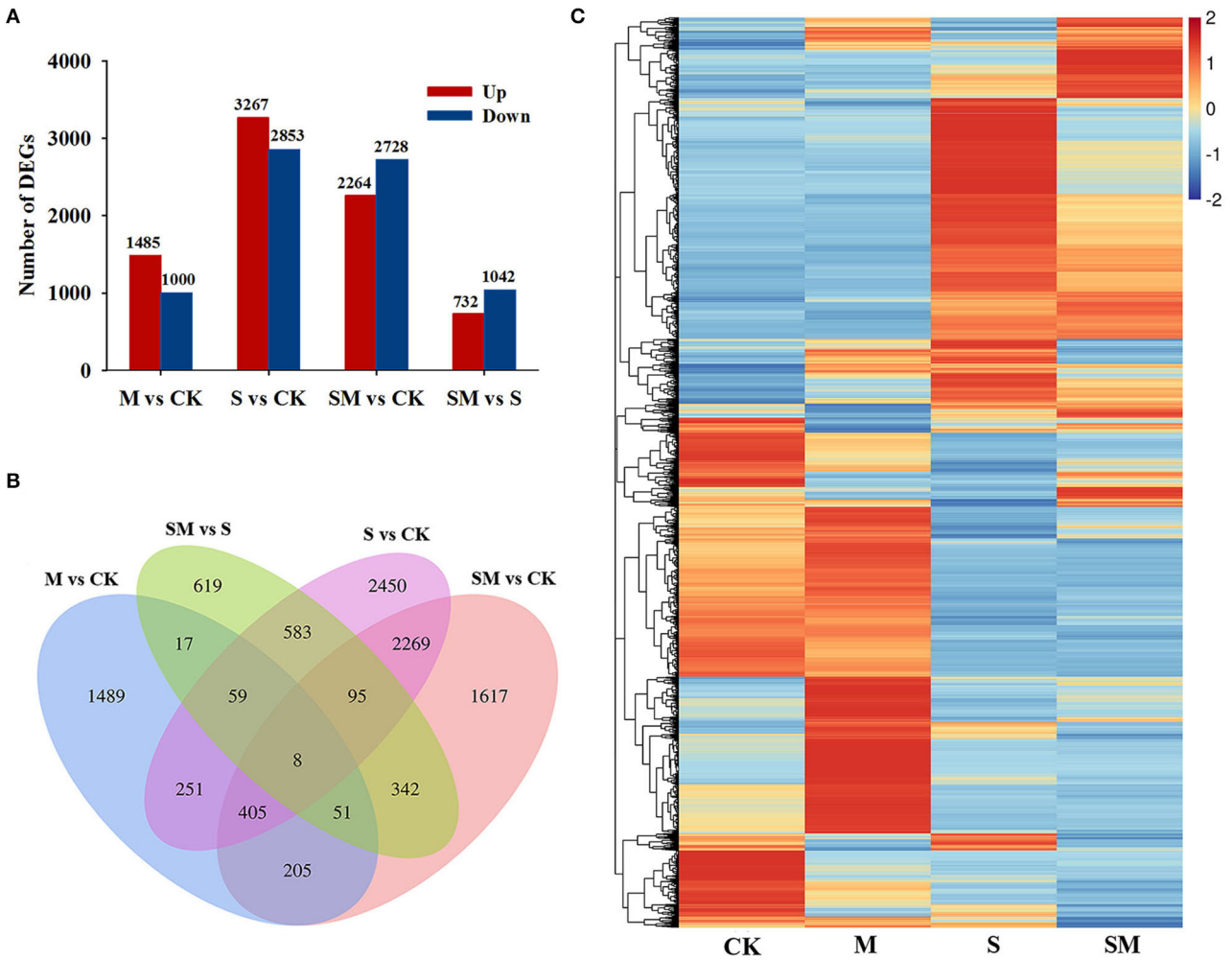


FIGURE 4 | Transcriptional profiles of alfalfa seedlings under different treatments. **(A)** Numbers of differentially expressed genes (DEGs) in the transcriptome data. **(B)** Venn diagram showing numbers of overlapping DEGs in the transcriptome data. **(C)** Hierarchical clustering (HCL) analysis of the DEGs under different treatments.

CK," and "SM *vs.* S" (**Figure 4B**), indicating that these genes may be involved in the melatonin-induced salt stress response. To determine the expression patterns of DEGs under different experimental conditions, the FPKM values of the DEGs were used for hierarchical cluster analysis (HCL) (**Figure 4C**).

Gene Ontology Enrichment Analysis of the DEGs

To further clarify the functional categories of the DEGs induced by melatonin treatment, GO enrichment analysis was performed. The DEGs from "M *vs.* CK," "S *vs.* CK," "SM *vs.* CK," and "SM *vs.* S" were assigned to 2,119, 2,958, 2,799, and 1,906 GO terms, respectively (**Supplementary Table 5**). As shown in **Supplementary Figure 3**, the identified DEGs were classified into the following three major GO categories: Biological process, cellular component, and molecular function. The top-25, top-15, and top-10 GO terms of biological process, cellular component and molecular function were selected for display and analysis according to the number of DEGs. The GO terms in the biological process category were mainly related to biological process, regulation of transcription, DNA-templated, transcription, DNA-templated, oxidation–reduction process, defense response, and protein phosphorylation. In the cellular component category, nucleus, plasma membrane, cytoplasm, integral component of membrane, chloroplast and cytosol were the main GO terms. In molecular function, the top enriched GO categories were involved in molecular function, protein binding, ATP binding, DNA binding transcription factor activity, metal ion binding and DNA binding. The GO terms in the three major categories were different among the comparison groups, which indicates that the salt resistance processes regulated by melatonin in alfalfa are complex.

Kyoto Encyclopedia of Genes and Genomes Pathway Enrichment Analysis of DEGs

Genes usually interact with each other to play roles in certain biological functions. To further understand the biological functions of genes involved in the melatonin-regulated salt resistance process, the enriched KEGG pathways were identified using the KEGG database (**Supplementary Table 6**). In total, 49 KEGG pathways were significantly enriched in at least one comparison ($p < 0.05$) (**Figure 5A**). A total of 93,232 and 10 KEGG pathways were significantly enriched in the "M *vs.* CK," "S *vs.* CK," "SM *vs.* CK" and "SM *vs.* S" comparisons, respectively. As shown in **Figure 5A**, aflatoxin biosynthesis, fatty acid elongation, benzoxazinoid biosynthesis, riboflavin metabolism, plant hormone signal transduction, glycine, serine and threonine metabolism, ABC transporters, brassinosteroid biosynthesis, phosphonate and phosphinate metabolism and N-glycan biosynthesis were the only significantly enriched metabolic pathways in "SM *vs.* CK." The pentose phosphate pathway, glycolysis/gluconeogenesis and carotenoid biosynthesis pathways were significantly enriched in the "M *vs.* CK," "S *vs.* CK" and "SM *vs.* CK" comparisons. Furthermore, carbon fixation in photosynthetic organism pathway was significantly changed in all comparisons and enriched in the top five (**Figure 5B**). Analysis of the top five KEGG pathway showed that the DEGs in the "S *vs.* CK" and "SM *vs.* CK" comparisons were enriched in photosynthesis and glyoxylate and dicarboxylate metabolism. The DEGs involved in photosynthesis–antenna proteins were observed in the "S *vs.* CK," "SM *vs.* CK" and "SM *vs.* S" comparisons. In addition, linoleic acid metabolism was enriched in "S *vs.* CK" and "SM *vs.* S" (**Figure 5B**). The results indicate that highly enriched pathways may be essential for melatonin-regulated salt resistance in alfalfa.

Detection of key Melatonin-Induced Genes and Transcription Factors (TFs) Under Salt Stress

Plants can trigger multiple biological processes to regulate gene transcription and physiological adaptation under unfavorable conditions (Deng et al., 2020). We examined the profiles of the genes involved in the plant-pathogen interaction pathway and starch and sucrose metabolism (**Figure 6**). In the plant–pathogen interaction pathway, a total of 87 DEGs were identified in "M *vs.* CK," including 66 upregulated and 21 downregulated DEGs. Under salt stress, most of the genes involved in Ca^{2+} signal transduction and WRKY TFs were downregulated in "S *vs.* CK" and "SM *vs.* CK," while the opposite was true in "M *vs.* CK" and "SM *vs.* S" (**Supplementary Table 7**). The expression profiles of 22 DEGs involved in the Ca^{2+} signaling pathway are shown in **Figure 6A**. Melatonin induced the expression of genes encoding cyclic nucleotide-gated ion channel (*CNGC20*), calmodulin/calmodulin-like protein (*CML7*, *CML11*, *CML24*, *CML45*, and *CML48*), calcium-dependent protein kinase (*CDPK1*), and respiratory burst oxidase (*RbohB*) in "M *vs.* CK" (**Supplementary Table 7**).

The expression profiles of 32 DEGs involved in starch and sucrose metabolism are shown in **Figure 6B**. According to our results (**Supplementary Table 8**), 35 genes were differentially expressed in "M *vs.* CK," with 24 upregulated and 11 downregulated genes, while 25 DEGs were identified in "SM *vs.* S." In "M *vs.* CK," sucrose synthase (*SUSs*), glycoside hydrolase family 17 (*GH17*), and β-amylase 3 (*BAM3*) were downregulated, while trehalose-phosphate phosphatase (*TPP*), endoglucanase 6 (*EDGL6*), and glucan endo-1,3-beta-glucosidase 13 (*EGLC13*) were upregulated. Specifically, the gene encoding fructokinase (*FRK7*) was upregulated in both "M *vs.* CK" and "SM *vs.* S," while *EGLC14* was upregulated in "M *vs.* CK" and downregulated in "SM *vs.* CK." In response to salt stress, 63 DEGs (31 upregulated and 32 downregulated genes) were identified in both "S *vs.* CK" and "SM *vs.* CK." The genes encoding *EGLCs*, *TPP*, hexokinase 3 (*HXK3*), *SUSs*, α-amylase (*AMY*), beta-glucosidase 24 (*BGLU24*), and *FRK4* were downregulated, whereas *BAM*, starch synthase (*SSs*), *BGLUs*, sucrose-phosphate synthase 2 (*SPS2*), granule-bound starch synthase (*GBSSs*), and *AMY2* were upregulated under salt stress.

Transcription factors play a critical role in regulating upstream stress signal transduction to downstream gene transcription and specific biological processes. A total of 2,111 TFs from 52 families showing differential expression were identified (**Figure 7A** and **Supplementary Table 9**). The C3H, MYB, and ERF families were the three most abundant TF families in each comparison. The expression profiles of the three TF families in four comparisons are presented in **Figure 7B**. We found 373 and 278 TFs were differentially expressed in "M *vs.* CK" and "SM *vs.* S"; the largest proportion of the C3H family contained 69 and 55 DEGs, respectively. Interestingly, 147 TFs were differentially expressed in "SM *vs.* S," but not in "S *vs.* CK," which suggests that these TFs are participated in salt resistance regulated by melatonin.

Detection of Melatonin-Induced Genes Involved in Phytohormone Signal Transduction Under Salt Stress

To explore whether melatonin participates in regulating other plant hormones in alfalfa salt tolerance, the expression profiles of the genes related to plant hormone signal transduction were analyzed (**Figure 8** and **Supplementary Table 10**). Furthermore, we predicted the key protein–protein interaction (PPI) network of proteins in the hormone signaling pathway (**Supplementary Figure 4**).

For the auxin signal, nine genes were differentially expressed in "M *vs.* CK," two of which were upregulated, genes encoding auxin-induced protein (*AUX28*) and auxin response factor (*ARF19*). In addition, *AUX10*, *AUX28*, and *IAA6* appear to be central genes in "SM *vs.* S." Under salt stress, the expression of *IAA27*, SAUR-like auxin–responsive family protein, *IAA8*, and *IAA22D* was induced in "S *vs.* CK," and transport inhibitor response 1 (*TIR1*) and *ARF3* expression was induced in "SM *vs.* CK." For the ABA signal, two *PYR/PYLs* (*PYL4* and *PYR1*) were downregulated under salt stress in "S *vs.* CK" and "SM *vs.* CK," while *PYL9* and protein phosphatase 2C (*PP2C24* and

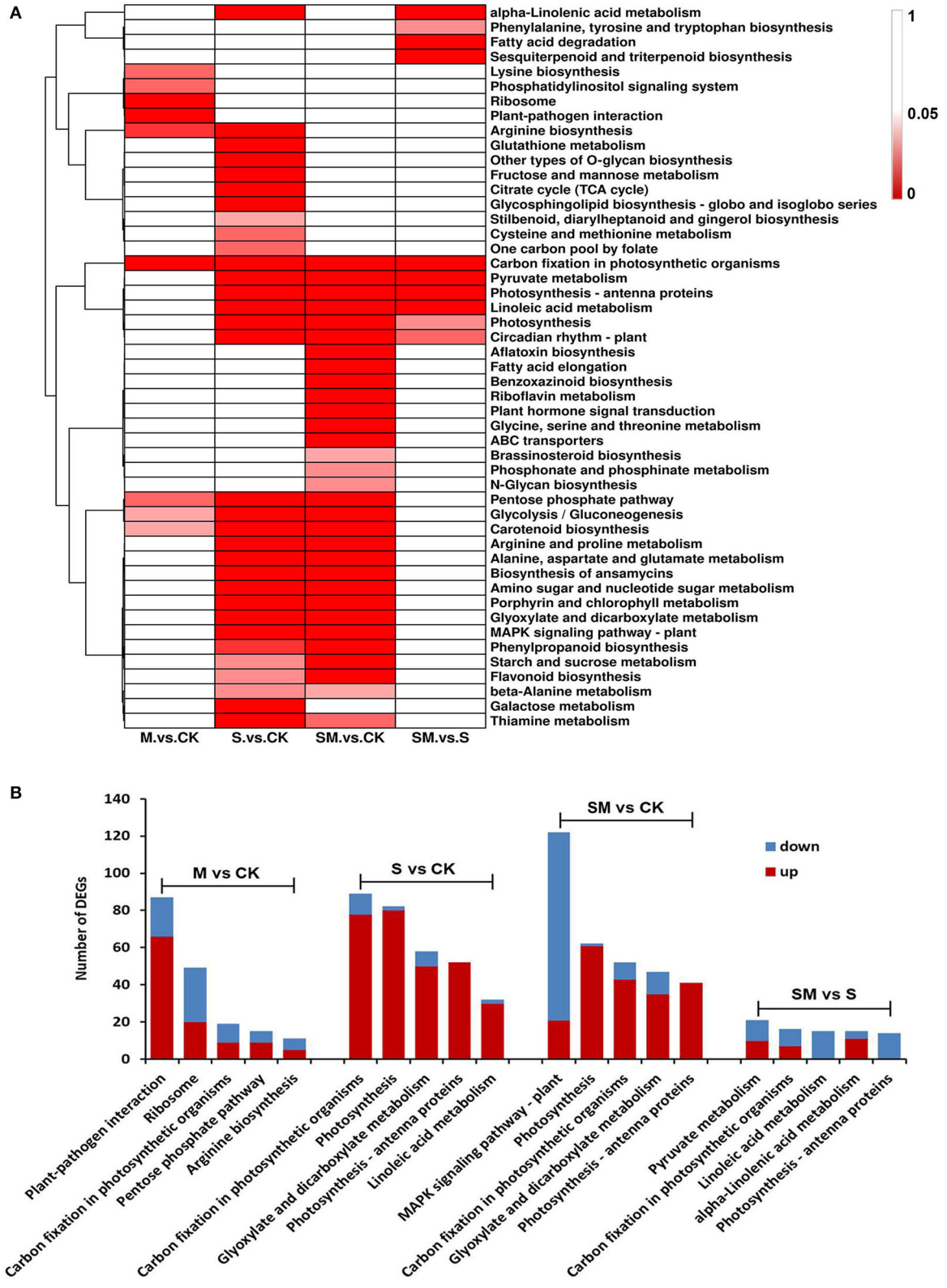


FIGURE 5 | KEGG pathway enrichment analysis of DEGs in different comparisons. **(A)** Heat map analysis of the significant *p*-values ($p < 0.05$) of KEGG term in different comparisons. **(B)** Top-5 KEGG pathways of DEGs in different comparisons.

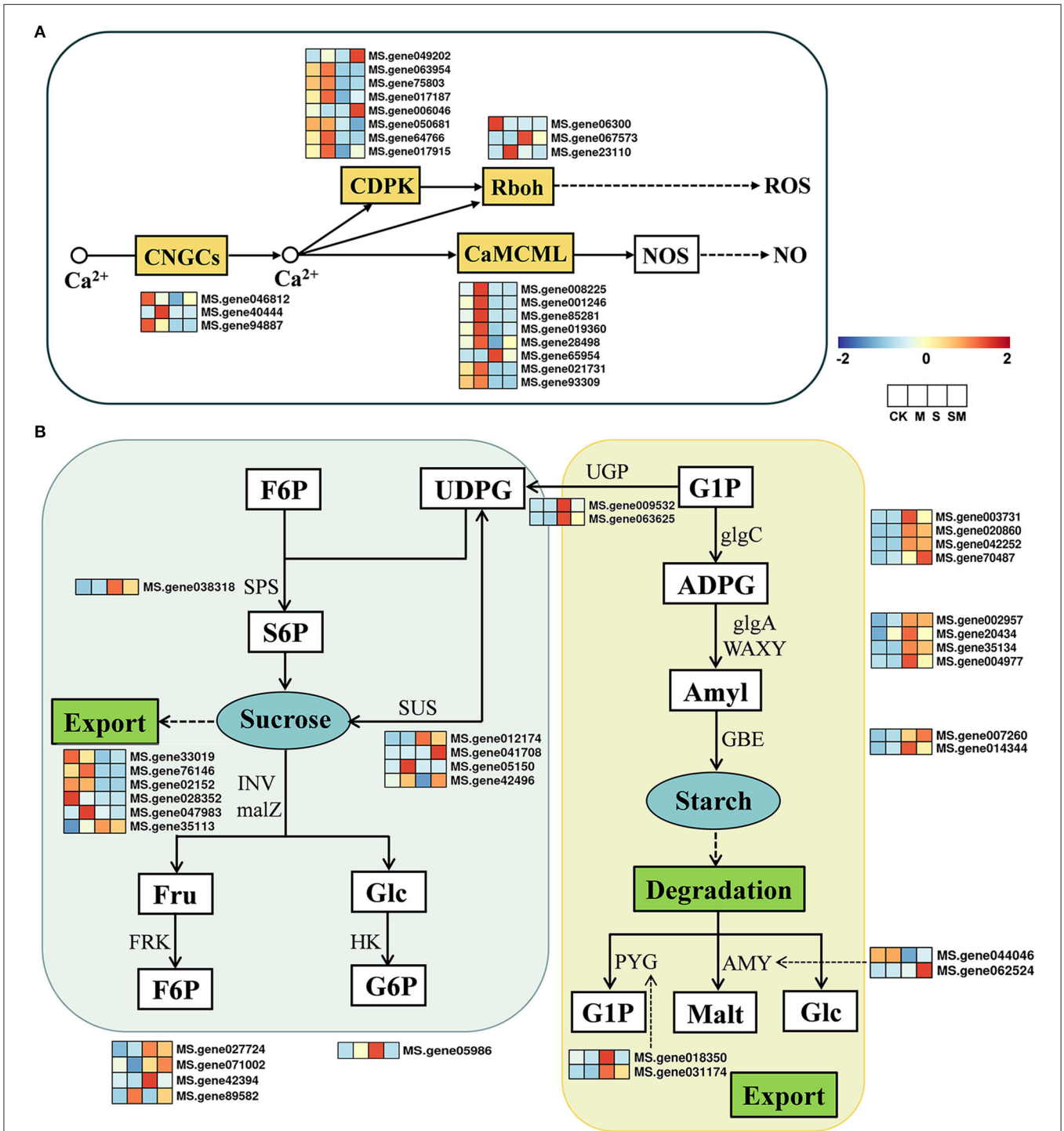


FIGURE 6 | Expression profiles of the key genes involved in **(A)** Ca^{2+} signal pathway and **(B)** starch and sugar metabolism in response to MT and salt treatment.

PP2C51) were upregulated. In addition, serine/threonine-protein kinase *SRK2A* was upregulated and *PP2C51* was downregulated in "M *vs.* CK." For the GA signal, the genes encoding F-box proteins (*GID2*), scarecrow-like proteins (*SCL14* and *SCL30*), and phytochrome interacting factor 3 (*PIF3*) were downregulated in "S *vs.* CK," whereas *SCL9* was upregulated in "S *vs.* CK" and "SM *vs.* CK." Moreover, the *PIF4* gene was upregulated in "M *vs.* CK" and "SM *vs.* CK," indicating that it was induced by melatonin. For the CTK signal, CTK receptor histidine kinases (*AHK4*) and response regulator and transcription factor RR-A-type family

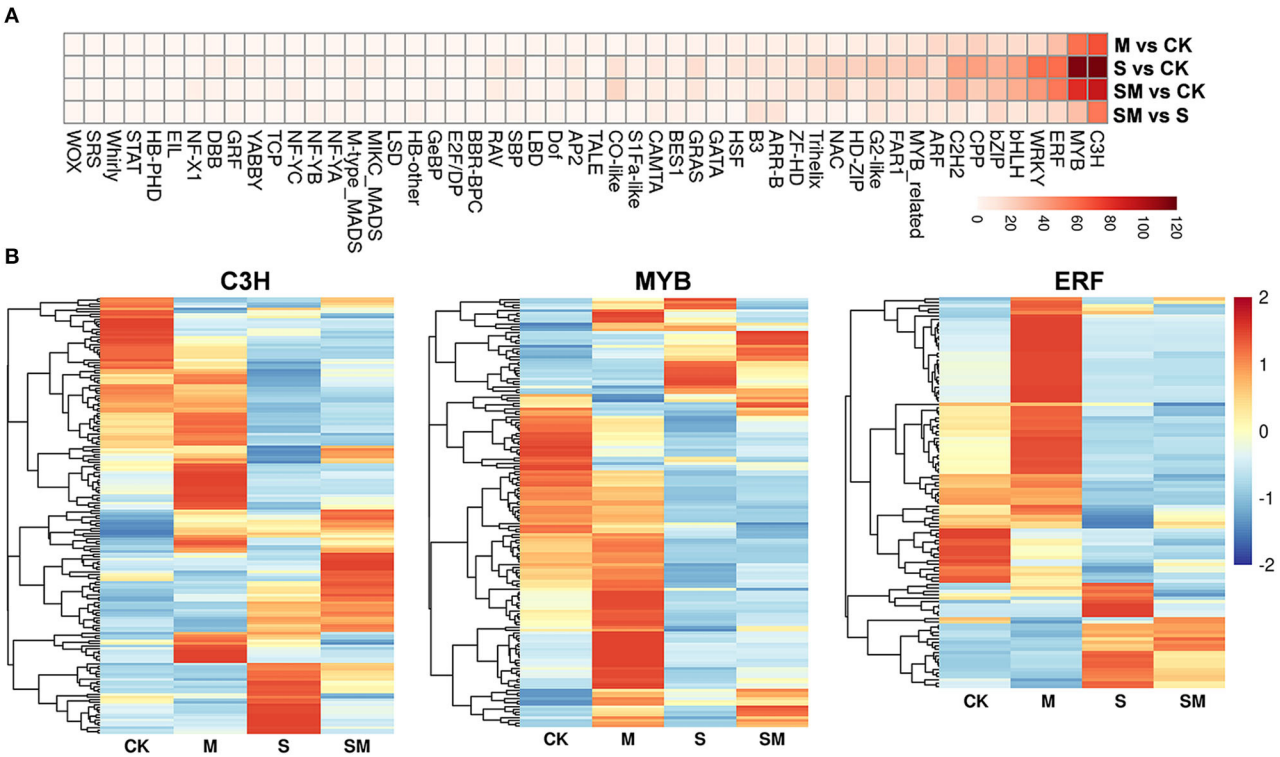


FIGURE 7 | Expression analysis of TF families in different comparisons. **(A)** Heatmap analysis of TF families. **(B)** The differential expression pattern of MYB, NAC, and WRKY families.

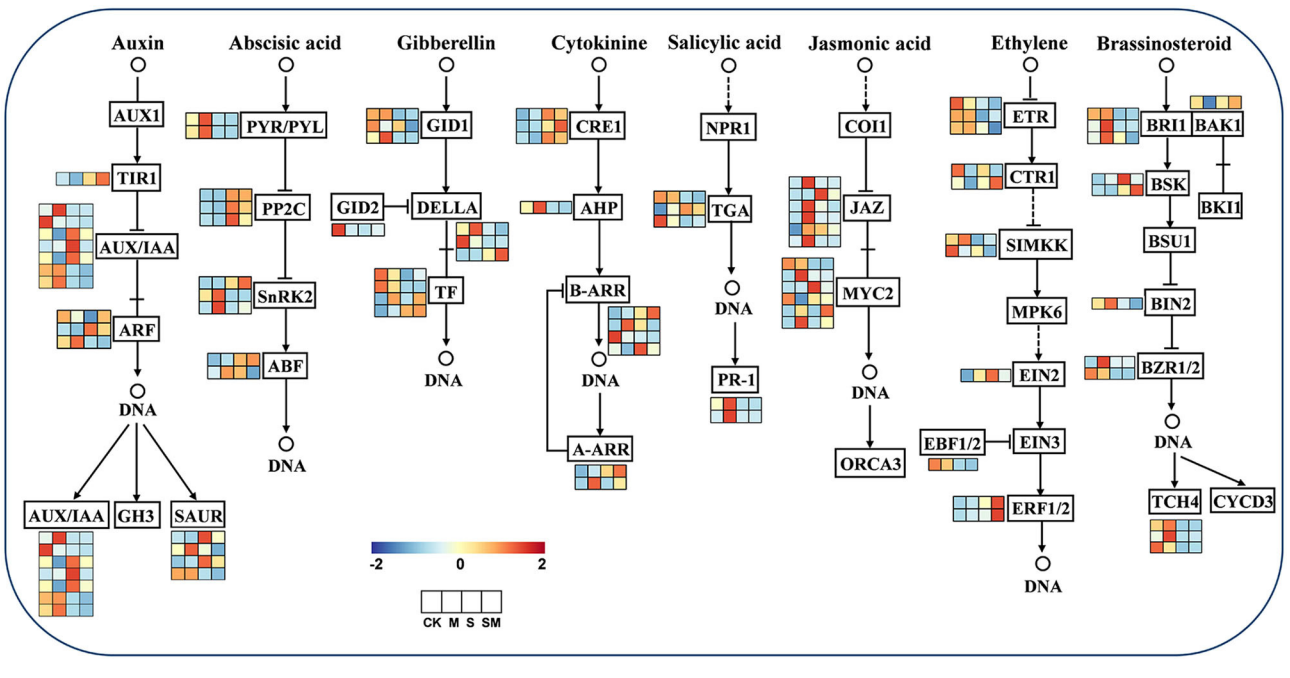


FIGURE 8 | Expression profiles of the plant hormone signal transduction pathway genes in response to MT and salt treatment.

genes were significantly upregulated in both "S *vs.* CK' and "SM *vs.* CK," while histidine-containing phosphotransfer proteins (*AHP1*) and two-component response regulators (*ARR12* and *ARR5*) were downregulated. However, melatonin upregulated the expression of *ARR5* in "M *vs.* CK" and "SM *vs.* S."

For the SA signal, the SA receptor gene *NPR3* and pathogenesis-related protein (*PR1*) were downregulated in "S *vs.* CK" and "SM *vs.* CK," while the downstream TGA TFs of NPR showed different expression patterns in SA signaling. For the JA signal, three *JAZs* (*TIFY3*, *TIFY10A* and *TIFY6B*) were upregulated in both "M *vs.* CK" and "S *vs.* CK," whereas melatonin also induced the expression of *TIFY5A*, *TIFY10B*, and *TIFY11B* in "M *vs.* CK." The *MYC2s* showed varied expression levels in JA signaling under salt and melatonin treatments, in which *bHLH18* was upregulated in "S *vs.* CK" and "SM *vs.* CK," but downregulated in "SM *vs.* S." For the ETH signal, salt stress inhibited the expression of ethylene receptor 2 (*ETR2*) and MEK map kinase kinase (*SIMKK*) in "S *vs.* CK" and "SM *vs.* CK." However, the ethylene receptor *EIN4* and *SIMKK* were significantly upregulated in "SM *vs.* S." Serine/threonine–protein kinase *CTR1* was downregulated in "M *vs.* CK" and "SM *vs.* CK," whereas it remained unchanged in "S *vs.* CK." In addition, ethylene-responsive transcription factor 1B (*ERF1B*) was downregulated in all comparisons except "SM *vs.* S." The EIN3-binding F-box protein 1 (*EBF1*) was downregulated in both "S *vs.* CK' and "SM *vs.* CK." For the BR signal, brassinazole-resistant 1 (*BZR1*), brassinosteroid LRR receptor kinase (*BRL*), and xyloglucan endotransglucosylase/hydrolase protein 23 (*XTH23*) were downregulated in "S *vs.* CK" and "SM *vs.* CK," whereas they were upregulated in "M *vs.* CK" comparison. In addition, serine/threonine–protein kinase *BSK3* was upregulated under both salt stress and melatonin treatment.

Detection of the Phytohormone Contents of Alfalfa Under Salt Stress

To evaluate the effect of exogenous melatonin on the phytohormone levels of alfalfa, we detected the contents of melatonin, ABA, IAA, GA_3, SA, BR, and ETH. Herein, we observed that melatonin treatment significantly increased the endogenous melatonin content of alfalfa before and after salt treatment (**Figure 9A**). As shown in **Figure 9B**, salt stress significantly increased the endogenous ABA content in comparison with the control plants. However, the application of melatonin decreased the ABA content under salt stress. The contents of IAA, GA_3, SA, and BR showed similar trends under treatment conditions. Under salt stress conditions, the IAA, GA_3, SA, and BR contents were remarkably accumulated compared with the control, while melatonin markedly increased the contents in alfalfa (**Figures 9C–F**). Melatonin-treated plants exhibited obviously higher ETH content than the control plants. Salt stress resulted in a significant increase in the ETH content compared with the control, while melatonin application further enhanced ETH accumulation under salt stress (**Figure 9G**).

Validation of DEGs by qRT-PCR

To validate the accuracy and reproducibility of the RNA-seq data, 20 DEGs, including hub genes in starch and sucrose metabolism and plant hormone signal transduction, were randomly selected for qRT-PCR. The expression patterns of all genes were highly consistent with the RNA-seq data (**Supplementary Figure 5**), which confirms that the DEGs identified in this study are credible.

DISCUSSION

Salt stress is one of the main factors affecting sustainable agricultural development. Melatonin has been reported to be involved in plant adaptive responses to salt stress in various plants (Yan et al., 2019; Liu et al., 2020; Zhang et al., 2021). Here, we analyzed the key mechanisms by which melatonin enhanced salt tolerance in alfalfa at the physiological and molecular levels. Our results suggest that melatonin could effectively alleviate oxidative damage and ion toxicity by improving photosynthetic capacity, antioxidant defense system, proline content, and the K^+/Na^+ ratio. Combined with the changes in starch, sugar, and hormone contents, and the expression of key genes in the transcriptome, our study indicates that melatonin is highly involved in starch and sugar metabolism and plant hormone signal transduction to regulate salt stress.

In the current study, the application of melatonin mitigated the growth inhibition of alfalfa under salt stress, which was reflected in superior plant height, leaf area, and fresh weight (**Figures 1A–D**). Exogenous melatonin reduces the detrimental effects of salinity on photosynthetic capacity (Chen et al., 2018; Liu et al., 2020). Consistent with these findings, higher Pn and Fv/Fm were observed in melatonin-treated alfalfa under salt stress conditions (**Figures 1E,F**). Salt stress causes excessive accumulation of ROS, which leads to cell membrane damage and oxidative stress (Miller et al., 2010). Melatonin is known as a broad-spectrum antioxidant that scavenges ROS and increases antioxidant enzyme activity under salt stress (Reiter et al., 2016). Herein, we found that melatonin treatment reduced electrolyte leakage, MDA content, and ROS accumulation under salt stress, while the activities of SOD, POD, GST, and APX were increased (**Figure 2**). These results suggest that melatonin alleviates oxidative stress mainly by enhancing the activities of antioxidant enzymes in alfalfa plants. Proline acts as an osmolyte and participates in melatonin-regulated salt stress tolerance (Siddiqui et al., 2019). This is consistent with our results, as we found that melatonin led to a remarkable increase in the proline content of alfalfa under salt stress (**Figure 3A**).

Melatonin, RBOH-mediated ROS, and Ca^{2+} exhibit complex signaling crosstalk to regulate plant responses to salt stress (Wei et al., 2018; Liu et al., 2020). The elevation of intracellular Ca^{2+} induced by salt stress sequentially activates CNGCs, CDPKs, or CAM/CMLs (Srivastava et al., 2013; Gao et al., 2020), while RBOHs can be activated by CDPKs (Drerup et al., 2013; Dubiella et al., 2013). In our study, melatonin activated the expression of *CNGCs* (such as *CNGC1* and *CNGC20*), *CaM/CMLs* (such as *CAM7*, *CAM11*, and *CAM24*), and *CDPKs* (such as *CDPK1*, *CDPK11*, and *CDPK28*) in the Ca^{2+} signaling pathway, which in turn downregulated *RBOHs* (*RBOHB* and *RBOHD*) (**Supplementary Table 7**). A recent study by Liu et al.

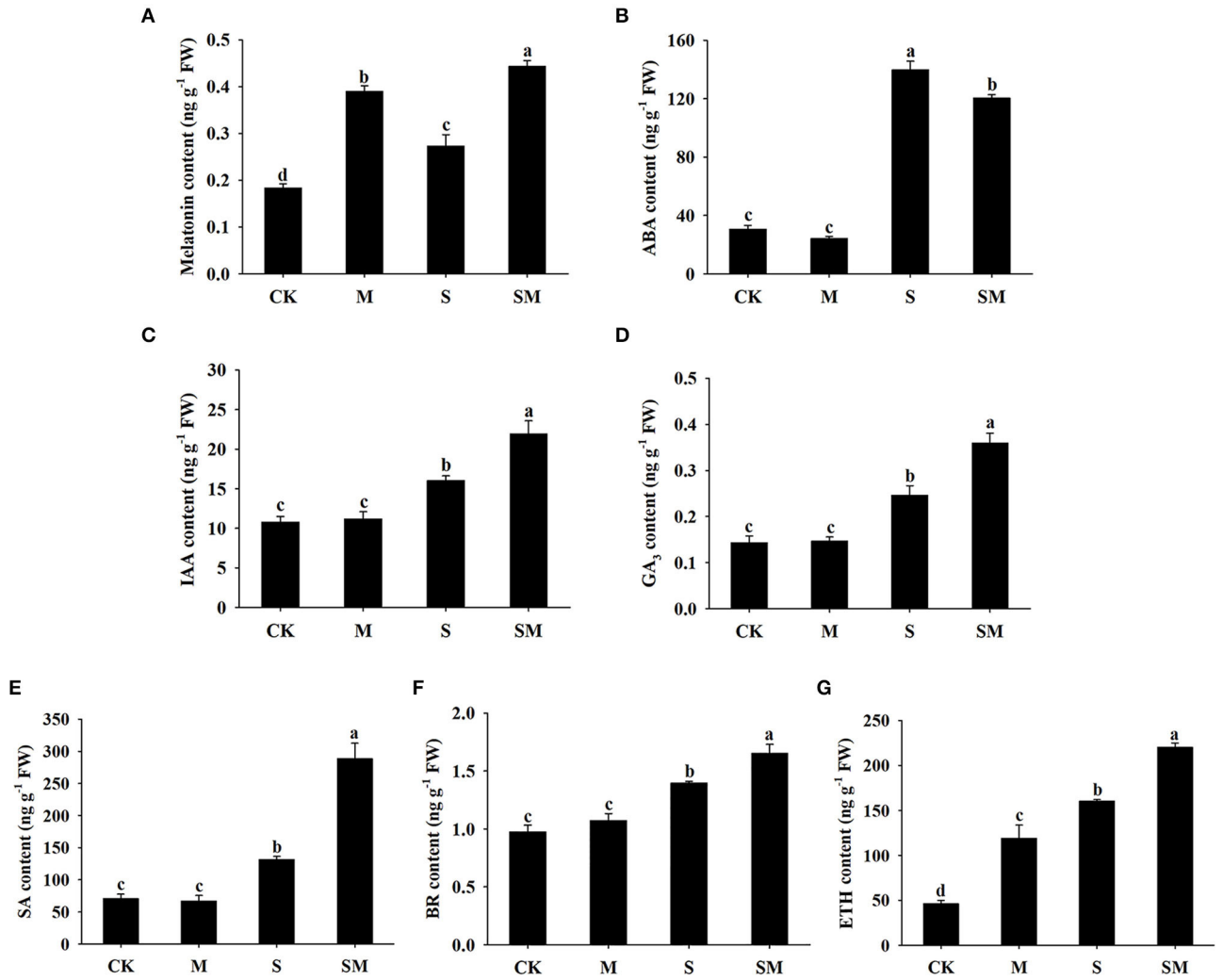


FIGURE 9 | Effect of exogenous melatonin on the endogenous hormone levels of alfalfa seedlings after 15 days of 150-mM NaCl treatment. **(A)** Melatonin content, **(B)** ABA content, **(C)** IAA content, **(D)** GA_3 content, **(E)** SA content, **(F)** BR content, and **(G)** ETH content. The data are means ± SE ($n = 3$) and different letters are significantly different ($p < 0.05$).

(2020) revealed that melatonin activated Ca^{2+} signaling to mediate RBOH, which is critical for changes in transcriptional profiles and high- and low-affinity K^+ transporter activity, thereby improving the salt tolerance of rice. This may be related to the higher K^+/Na^+ in melatonin-treated plants under salt stress conditions (**Figure 3F**).

Sugars act as osmoprotectants to participate in the abiotic stress response in plants (Wang et al., 2016; Yang et al., 2019). Stress-induced starch-to-sugar conversion promotes the accumulation of sugar, which provides osmoprotection and energy supplies for plants (Dong and Beckles, 2019). In sucrose biosynthesis and degradation, SPS and SUS are the key enzymes (Winter and Huber, 2000). Salt reduces the activities of GBSS and AGPase; thus, inhibiting the synthesis of starch (Libalweksler et al., 1994; Chen et al., 2008). BAM1 and AMY3 were also activated by osmotic stress to degrade starch into sugar (Thalmann et al., 2016). However, Yin et al. (2010) reported that salinity increased AGPase activity and enhanced carbohydrate accumulation as starch during the early development stages in tomato. In our study, the contents of sugar and starch in alfalfa increased obviously under salt stress (**Figures 3B,C**). Interestingly, salt stress upregulated α-glucan phosphorylase, *BAM*, *AMY2*, *SS*, *AGPs*, *GBSS*, and *SPS2*, while it downregulated *AMY* and *SUS* (**Supplementary Table 8**). This result indicates that, during a certain stage of salt treatment, the contents of starch and sugar in alfalfa could be simultaneously maintained at high levels to resist stress.

The previous reports have revealed that melatonin could enhance abiotic stress resistance by maintaining a higher accumulation of carbohydrates in bermudagrass (Shi et al., 2015).

In our study, melatonin treatment increased the accumulation of starch, which was a major reason for the higher photosynthetic capacity of melatonin-treated plants under salt stress. The sugar and starch content increased in both salt-tolerant and salt-susceptible tomato leaves, but the tolerant genotype had higher starch accumulation (Balibrea et al., 2000). Plants convert a portion of sugars to starch to minimize the physiological damage of excess sugar in the source leaves (Dong and Beckles, 2019). According to our data, the genes *EDGL6*, *EGLC13*, *EGLC14, GBSS2*, *BAM*, and *FRK7* were upregulated in "M *vs.* CK," while six *SUSs* were downregulated (**Figure 6B** and **Supplementary Table 8**). The melatonin-induced expression changes of most genes were more conducive to the accumulation of starch, thus reducing the content of sucrose. Higher starch is proposed to increase starch statoliths and gravitropic response, and to direct root growth for the acquisition of nutrients, minerals, or water under salt stress (Baldwin et al., 2013; Thitisaksakul et al., 2017). The isoforms of BGLU located to various organelles catalyze the single-step hydrolysis of ABA-glucose ester to produce ABA, and they become activated upon stress (Han et al., 2020). Wang et al. (2021) have revealed that melatonin regulates *BGLU18* mediated ABA-glucose ester hydrolysis to modulate ABA homeostasis and abiotic stress responses. We found that *BGLU12* and *BGLU46* were downregulated in the "SM *vs.* S" comparison, indicating that they may be associated with the ABA signaling pathway. Moreover, a high Na^+ concentration stimulates the accumulation of starch in common reed (*Phragmites australis*), which leads to the deposition of Na^+ in starch granules, thus achieving "ion-trapping" to maintain cellular osmotic balance (Kanai et al., 2007). Melatonin treatment may be beneficial to this process, which needs to be further verified. Taken together, these results suggest that salt stress activates the starch and sucrose metabolism to maintain a high and balanced level; moreover, the application of melatonin was beneficial to the maintenance of the starch-sucrose ratio to improve the salt tolerance of alfalfa.

Transcription factors have been reported to play important roles in salt stress responses, which uniquely regulate and modify different stress-responsive genes. Many TFs have been identified to be involved in melatonin-mediated stress tolerance. Most of these were stress-related TFs, including MYBs, WRKYs, NACs, and zinc finger-related TFs (Zhang et al., 2014; Zhan et al., 2021). In this study, various TFs were significantly differentially expressed in different comparisons, among which C3H, MYB, and ERF were the three most abundant families (**Figure 7**). The largest proportion of the C3H family contained 69 and 55 DEGs in "M *vs.* CK" and "SM *vs.* S," respectively, suggesting that these TFs participated in salt resistance regulated by melatonin. In tomato, *DREB1α* and *IAA3* are key downstream transcription factors of melatonin-induced sodic alkaline stress tolerance at the genetic level (Yan et al., 2019). These results indicate that TFs might contribute to enhancing the salt tolerance of melatonin-treated alfalfa.

It has been shown that phytohormones, act as secondary messengers, participate in stress sensing and signal transduction through antagonistic or synergistic action, and activate downstream transcription factors to regulate stress responses (Long and Benfey, 2006; Yan et al., 2019). In this study, we focused on the major hormone signaling pathways under salt and melatonin treatments and their possible crosstalk. Salt stress induces changes in endogenous melatonin levels in different plant species (Li et al., 2019c; Zhang et al., 2021). It has been reported that *COMT* and *SNAT* are pivotal genes for melatonin biosynthesis (Arnao and Hernandez-Ruiz, 2014). Herein, we observed a significant increase in endogenous melatonin levels in both the control and salt-treated plants after melatonin treatment (**Figure 9A**). In addition, the expression of *COMT* was upregulated under salt stress, and its extent of upregulation in "SM *vs.* CK" was higher than that in "S *vs.* CK" (**Supplementary Figure 5**). The results suggest that melatonin is closely related to the salt tolerance of alfalfa, and consistent results were found in cotton and tomato (Yan et al., 2019; Zhang et al., 2021).

Auxins play key roles in regulating plant growth and development, and they can govern the growth response of plants to stress (Eyidogan et al., 2012). Auxin-responsive genes have been separated into three major classes: Aux/IAA, GH3 and SAUR (Hagen and Guilfoyle, 2002). There are at least two auxin receptors reported in the literature: nuclear TIR1/AFB Aux/IAA coreceptor and auxin-binding protein 1 (ABP1) (Grones and Friml, 2015). Stress pathways interact with the auxin gene regulatory network through the transcription of *Aux/IAA* genes (Shani et al., 2017). In this study, melatonin suppressed almost all *Aux/IAA* and *SAUR* genes under salt stress, but induced *TIR1* and *ARF3* expression (**Supplementary Table 10**). Chen et al. (2017) reported that overexpression of *CsTIR* enhanced salt tolerance in transgenic Arabidopsis. It has been reported that melatonin treatment induces a slight increase in endogenous IAA in *Brassica juncea* and tomato (Chen et al., 2009; Wen et al., 2016). Our results showed that endogenous IAA content in alfalfa was significantly increased after melatonin application under salt stress (**Figure 9C**). The results indicate that melatonin may alleviate salt stress by modulating the expression of auxin coreceptors and response genes and increasing auxin content.

Abscisic acid has been proposed as a stress hormone because it acts as an important internal signal mediating plant responses to stress, and stress tends to induce ABA synthesis (Fahad et al., 2015). In the ABA biosynthetic pathway, when PYR/PYL/RCAR binds to ABA, the complex interacts with PP2C to reduce the inhibition of SNF1-related kinases (SnRKs), thus activating their downstream transcription factors (Yang et al., 2019). ABA oxidase (AAO) and 9-cis-epoxycarotenoid dioxygenases (NCED) are key enzymes in the ABA biosynthetic pathway (Mcadam et al., 2015). In this study, salt stress induced *PYL9*, *PP2Cs*, *AAO1*, and *NCEDs* expression and ABA accumulation, while repressing *PYL4* and SnRK2s expression. However, melatonin treatment increased SRK2A expression in the "M *vs.* CK" comparison (**Supplementary Table 10**). Di et al. (2018) reported that *BnPYL9-1* and *BnPYL9-2* expression was inhibited by salinity stress, which may be due to negative feedback regulation caused by high ABA accumulation. Also, SnRK2s are key regulators governing plant adaptive responses to osmotic stresses. Soma et al. (2020) reported that subclass I SnRK2s (SRKs) are rapidly activated by osmotic stress prior to ABA accumulation, implying

that SRK2A may not be activated by ABA. Melatonin treatment upregulated the expression of ABA catabolism gene *CYP707A1* and reduced ABA accumulation under salt stress. This result is consistent with that in tomato reported by Hu et al. (2021). In addition, melatonin treatment resulted in downregulated expression of *PP2C51*. The previous studies have revealed that *PYL4* and *PP2Cs* are considered to be central genes that interact with ERF and GA metabolic genes, and that melatonin plays a regulatory role upstream of the ABA signaling pathway (Wang et al., 2021). In this study, we speculate that PP2Cs interacted with ETH and IAA metabolic genes in the "S *vs.* CK" and "SM *vs.* CK" comparisons (**Supplementary Figure 4**). These results suggest that melatonin might enhance the salt tolerance of alfalfa by mediating *PP2C51* and *CYP707A1* expression and crosstalk with ETH and IAA signals.

As a key messenger and integrator of intrinsic growth responses, GA is also involved in regulating plant responses to salt stress. Gibberellic acid signaling is mediated by its receptor GID1, the repressor DELLA, and the F-box protein GID2; thus, triggering downstream responses (Kohli et al., 2013). The repressor, DELLA, controls the GA signaling pathway by antagonizing the GA-positive regulator *SCL3* promoter sequence and blocks PIF transcriptional regulation activity to orchestrate GA homeostasis (Hirano et al., 2008). The findings of our study showed that salt stress inhibited *SCLs* and *PIF3* expression, while melatonin treatment resulted in upregulation of *PIF4* and *SCL13* to regulate GA signaling. Zhang et al. (2014) reported that melatonin increased GA content by promoting the expression of the GA synthesis genes *GA20ox* and *GA3ox* under salt stress conditions. Similarly, our results also showed that *GA2ox2* was upregulated after melatonin application. The repressor, DELLAs, orchestrate the crosstalk between GA and other plant hormones, such as ABA and ETH, to participate in the salt stress response (Achard et al., 2006). In our study, *PIFs* and *SCLs* were hub genes in the GA signaling pathway, and *SCLs* interacted with the genes involved in ABA and SA signal transduction to alleviate salt stress in alfalfa.

Ethylene is considered a stress hormone required in various abiotic stress responses in plants. Salt stress promotes ethylene production in plants by regulating the activities of 1-aminocyclopropane-1-carboxylate (ACC) synthase (ACS) and ACC oxidase (ACO) (Achard et al., 2006; Dong et al., 2011). Melatonin enhances salt tolerance by promoting *MYB108A*-mediated ethylene biosynthesis in grapevines (Xu et al., 2019). In our study, salt stress and melatonin treatment increased the ETH content in alfalfa (**Figure 9G**), indicating that melatonin could defend against salt stress by regulating ethylene levels. The ERFs are the major downstream regulatory factors of the ETH signaling pathway in stress responses. Although *ERF1B* was identified as a positive regulator of salt stress tolerance, its expression was repressed under salt treatment in *Chrysanthemum* (Gao et al., 2015). Similarly, *ERF1B* was downregulated by melatonin under salt stress (**Supplementary Table 10**). The regulatory effect of melatonin on ethylene synthesis may be related to complex hormone signal crosstalk (Xu et al., 2019). Ethylene-responsive transcription factor can combine with ABA to affect stomatal opening under drought stress and through the combination of ERFs, GA, and CIPK under low-oxygen stress to play a specific role (Kohli et al., 2013). In this study, *ERF1B* was associated with other hormonal signaling components, such as *ARF2A*, *TIFY10A*, and *MYC2*, indicating that hormonal signals regulate salt stress through a crosstalk relationship.

Salicylic acid participates in defense responses to a variety of environmental stresses including salinity. Salicylic acid signaling leads to the activation of the NPR1, which is thought to be recruited to numerous downstream PRs by transcription factors such as TGAs (Jin et al., 2018). Li et al. (2019a) reported that overexpression of *TGA17* enhanced the salt tolerance of soybean. Our data showed that melatonin induced the upregulation of *TGA3* and *TGA7* under salt stress. Moreover, cytokinin-activated transcription factor ARR2 in Arabidopsis promotes plant immunity through a TGA3/NPR1-dependent salicylic acid signaling pathway (Choi et al., 2010). Here, the hub gene *TGA* was associated with the *ETR1*, *SCL14*, and *MYC2* genes in the ETH, GA and JA signaling pathways, which suggested that crosstalk occurred between these phytohormones. Taken together, these results reveal the crosstalk relationship of plant hormones and provide new insights into the involvement of hormone signals in salt stress responses.

Jasmonic acid is known to play major roles in mediating the plant defense response against salt stress. The transcription factor MYC2 and repressor protein JAZ play crucial roles in the JA response under stress conditions (Verma et al., 2016). Under JA-stimulated conditions, JA-Ile (bioactive JA) binds to its receptor, an F-box protein CORONATINE INSENSITIVE1 (COI1), and leads to 26S proteasome-mediated degradation of JAZ, thereby allowing for MYC2 to upregulate the expression level of JA target genes (Chini et al., 2007). In this study, salt stress and melatonin treatment induced the expression of *TIFY3*, *TIFY10A*, *TIFY6B*, and *MYC2* (*bHLH18*), indicating that they might play important regulatory roles against salt stress in alfalfa (**Supplementary Table 10**). This is consistent with the regulation of JA signaling by melatonin in loquat under drought stress (Wang et al., 2021).

The previous studies have demonstrated that melatonin and BRs synergistically regulate plant morphogenesis (Hwang and Back, 2018; Fu et al., 2022). However, there are still divergent views on the crosstalk of melatonin and BR signaling in abiotic stress (Hwang and Back, 2019). In this study, melatonin treatment increased the BR content under salt stress (**Figure 9F**). In addition, melatonin induced the transcription levels of genes related to BR signaling including *BZR1*, *BRL*, *BSK3*, and *XTH23* in alfalfa (**Supplementary Table 10**), indicating that melatonin and BR function together in response to salt stress. These findings are consistent with the study of Fu et al. (2022), who reported that melatonin-induced cold and drought tolerance is regulated by BR in perennial ryegrass. BR-mediated stress tolerance in Arabidopsis was associated with the ABA, SA, and ETH pathways (Divi et al., 2010). In this study, the BR LRR receptor kinase CURL3 interacted with EIN3, NPR1, IAA6, PP2C51, and GID2 under salt stress, indicating that there may be a crosstalk relationship between BR and ETH, SA, IAA, ABA, and GA signaling (**Supplementary Figure 4**).

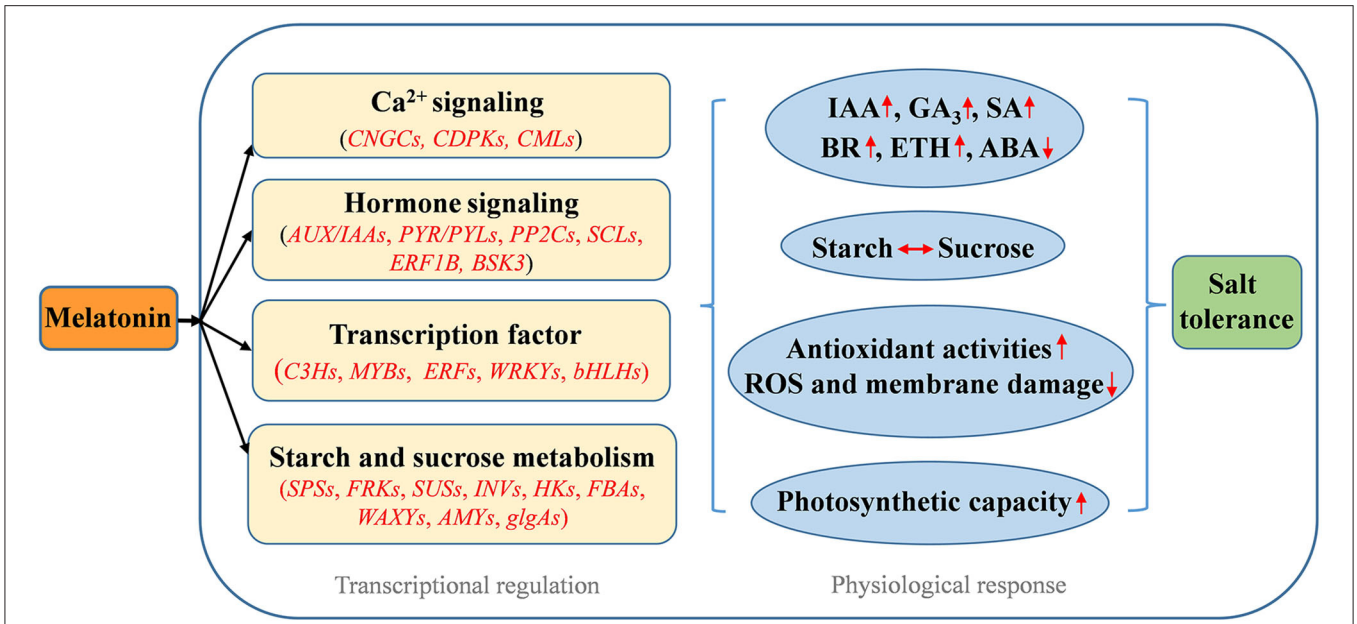


FIGURE 10 | Schematic model for melatonin-induced salt stress tolerance in alfalfa.

In conclusion, this study provides new insight into the protective roles of melatonin against salt stress in alfalfa. Based on physiochemical and transcriptomic data, a schematic model for the regulation of the alfalfa salt stress response by melatonin is presented in **Figure 10**. Melatonin alleviated salt stress by increasing starch accumulation to maintain high photosynthetic capacity and enhancing the antioxidant defense system to scavenge excess ROS accumulation. Melatonin improved the salt stress tolerance of alfalfa mainly by mediating the profiles of the genes involved in Ca^{2+} signaling and starch and sucrose metabolism. Melatonin participated in mediating other plant hormone signal transduction pathways and affected endogenous hormone levels by regulating related genes under salt stress. Moreover, the crosstalk relationship among important phytohormone signaling pathways in alfalfa under salt stress is evidenced. Our results lay a foundation for further research on the molecular mechanisms of the melatonin-induced salt stress response.

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: NCBI Gene Expression Omnibus (GEO) database, accession no: GSE199945.

AUTHOR CONTRIBUTIONS

SL and BF designed the experiments and wrote the manuscript. SL, YW, XG, and JL performed the experiment. YW and XG analyzed the data and prepared the figures. JL and BF provided ideas and revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (32101426), Ningxia Natural Science Foundation (2022AAC03124), and Ningxia Hui Autonomous Region Agricultural Breeding Special Project (2019NYYZ0403).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.919177/full#supplementary-material

REFERENCES

Achard, P., Cheng, H., Grauwe, L. D., Decat, J., Schoutteten, H., Moritz, T., et al. (2006). Integration of plant responses to environmentally activated phytohormonal signals. *Science* 311, 91–94. doi: 10.1126/science.1118642

Arnao, M. B., and Hernandez-Ruiz, J. (2014). Melatonin: plant growth regulator and/or biostimulator during stress? *Trends Plant Sci.* 19, 789–797. doi: 10.1016/j.tplants.2014.07.006

Arnao, M. B., and Hernandez-Ruiz, J. (2019). Melatonin: A new plant hormone and/or a plant master regulator? *Trends Plant Sci.* 24, 38–48. doi: 10.1016/j.tplants.2018.10.010

Back, K., Tan, D., and Reiter, R. J. (2016). Melatonin biosynthesis in plants: multiple pathways catalyze tryptophan to melatonin in the cytoplasm or chloroplasts. *J. Pineal Res.* 61, 426–437. doi: 10.1111/jpi.12364

Baldwin, K. L., Strohm, A. K., and Masson, P. H. (2013). Gravity sensing and signal transduction in vascular plant primary roots. *Am. J. Bot.* 100, 126–142. doi: 10.3732/ajb.1200318

Balibrea, M. E., Dell'Amico, J., Bolarin, M. C., and Perez-Alfocea, F. (2000). Carbon partitioning and sucrose metabolism in tomato plants growing under salinity. *Physiol. Plant.* 110, 503–511. doi: 10.1111/j.1399-3054.2000.1100412.x

Benabderrahim, M. A., Guiza, M., and Haddad, M. (2020). Genetic diversity of salt tolerance in tetraploid alfalfa (*Medicago sativa* L.). *Acta Physiol. Plant.* 42:5. doi: 10.1007/s11738-019-2993-8

Cheeseman, J. (2016). 7-Food security in the face of salinity, drought, climate change, and population growth. *Halophytes Food Security Dry Lands* 2016, 111–123. doi: 10.1016/B978-0-12-801854-5.00007-8

Chen, H., Chen, J., and Wang, S. (2008). Molecular regulation of starch accumulation in rice seedling leaves in response to salt stress. *Acta Physiol. Plant.* 30, 135–142. doi: 10.1007/s11738-007-0101-y

Chen, Q., Qi, W., Reiter, R. J., Wei, W., and Wang, B. (2009). Exogenously applied melatonin stimulates root growth and raises endogenous IAA in roots of etiolated seedling of *Brassica juncea*. *J. Plant Physiol.* 166, 324–328. doi: 10.1016/j.jplph.2008.06.002

Chen, Y., Mao, J., Sun, L., Huang, B., Ding, C., Gu, Y., et al. (2018). Exogenous melatonin enhances salt stress tolerance in maize seedlings by improving antioxidant and photosynthetic capacity. *Physiol. Plant.* 164, 349–363. doi: 10.1111/ppl.12737

Chen, Z., Li, M., Yuan, Y., Hu, J., Yang, Y., Pang, J., et al. (2017). Ectopic expression of cucumber (*Cucumis sativus* L.) *CsTIR/AFB* genes enhance salt tolerance in transgenic *Arabidopsis*. *Plant Cell Tiss. Organ Cult.* 131, 107–118. doi: 10.1007/s11240-017-1267-7

Chini, A., Fonseca, S., Fernandez, G., Adie, B., Chico, J. M., Lorenzo, O., et al. (2007). The JAZ family of repressors is the missing link in jasmonate signalling. *Nature* 448, 666–671. doi: 10.1038/nature06006

Choi, J., Huh, S. U., Kojima, M., Sakakibara, H., Paek, K. H., and Hwang, I. (2010). The cytokinin-activated transcription factor ARP2 promoters plant immunity via TGA3/NPR1-development salicylic acid signaling in Arabidopsis. *Dev. Cell.* 19, 284–295. doi: 10.1016/j.devcel.2010.07.011

Deng, C., Zhang, Z., Yan, G., Wang, F., Zhao, L., Liu, N., et al. (2020). Salt-responsive transcriptome analysis of triticale reveals candidate genes involved in the key metabolic pathway in response to salt stress. *Sci. Rep.* 10:20669. doi: 10.1038/s41598-020-77686-8

Di, F., Jian, H., Wang, T., Chen, X., Ding, Y., Du, H., et al. (2018). Genome-wide analysis of the *PYL* gene family and identification of *PYL* genes that respond to abiotic stress in *Brassica napus*. *Genes* 9:156. doi: 10.3390/genes9030156

Divi, U. K., Rahman, T., and Krishna, P. (2010). Brassinosteroid-mediated stress tolerance in *Arabidopsis* shows interaction with absicisic acid, ethylene and salicylic acid pathways. *BMC Plant Biol.* 10, 151–165. doi: 10.1186/1471-2229-10-151

Dong, H., Zhen, Z., Peng, J., Chang, L., Gong, Q., and Wang, N. (2011). Loss of *ACS7* confers abiotic stress tolerance by modulating ABA sensitivity and accumulation in Arabidopsis. *J. Exp. Bot.* 62, 4875–4887. doi: 10.1093/jxb/err143

Dong, S., and Beckles, D. M. (2019). Dynamic changes in the starch-sugar interconversion within plant source and sink tissues promote a better abiotic stress response. *J. Plant Physiol.* 234–235, 80–93. doi: 10.1016/j.jplph.2019.01.007

Drerup, M., Schlücking, K., Hashimoto, K., Manishankar, P., Steinhorst, L., Kuchitsu, K., et al. (2013). The Calcineurin B-like calcium sensors CBL1 and CBL9 together with their interacting protein kinase CIPK26 regulate the Arabidopsis NADPH oxidase RBOHF. *Mol. Plant* 6, 559–569. doi: 10.1093/mp/sst009

Dubiella, U., Seybold, H., Durian, G., Komander, E., Lassig, R., Witte, C. P., et al. (2013). Calcium-dependent protein kinase/NADPH oxidase activation circuit is required for rapid defense signal propagation. *Proc. Natl. Acad. Sci. U.S.A.* 110, 8744–8749. doi: 10.1073/pnas.1221294110

Eyidogan, F., Oz, M. T., Yucel, M., and Oktem, H. A. (2012). "Signal transduction of phytohormones under abiotic stresses," in: *Phytohormones and Abiotic Stress Tolerance in Plants*, eds N. A. Khan, R. Nazar, N. Iqbal, and N. A. Anjum (Berlin: Springer). doi: 10.1007/978-3-642-25829-9_1

Fahad, S., Hussain, S., Matloob, A., Khan, F. A., and Huang, J. (2015). Phytohormones and plant responses to salinity stress: a review. *Plant Growth Regul.* 75, 391–404. doi: 10.1007/s10725-014-0013-y

Fan, J., Xie, Y., Zhang, Z., and Chen, L. (2018). Melatonin: a multifunctional factor in plants. *Int. J. Mol. Sci.* 19, 1–14. doi: 10.3390/ijms19051528

FAO (2019) *Saline Soils and Their Management, Food and Agricultural Organization of the United Nations*. Available online at: http://www.fao.org/3/x5871e/x5871e04.htm (accessed December 1, 2019).

Fu, J., Zhang, S., Jiang, H., Zhang, X., Gao, H., Yang, P., et al. (2022). Melatonin-induced cold and drought tolerance is regulated by brassinosteroids and hydrogen peroxide signaling in perennial ryegrass. *Environ. Exp. Bot.* 196:104815. doi: 10.1016/j.envexpbot.2022.104815

Gao, C., Li, P., Song, A., Wang, H., Wang, Y., Ren, L., et al. (2015). Isolation and characterization of six *AP2/ERF* transcription factor genes in *Chrysanthemum nankingense*. *Int. J. Mol. Sci.* 16, 2052–2065. doi: 10.3390/ijms16012052

Gao, T., Zhang, Z., Liu, X., Wu, Q., Chen, Q., Liu, Q., et al. (2020). Physiological and transcriptome analyses of the effects of exogenous dopamine on drought tolerance in apple. *Plant Physiol. Biochem.* 148, 260–272. doi: 10.1016/j.plaphy.2020.01.022

Grones, P., and Friml, J. (2015). ABP1: finally docking. *Mol. Plant.* 8, 356–358. doi: 10.1016/j.molp.2014.12.013

Hagen, G., and Guilfoyle, T. (2002). Auxin-responsive gene expression: genes, promoters and regulatory factors. *Plant Mol. Biol.* 49, 373–385. doi: 10.1023/A:1015207114117

Han, Y., Watanabe, S., Shimada, H., and Sakamoto, A. (2020). Dynamics of the leaf endoplasmic reticulum modulate β-glucosidase-mediated stress-activated ABA production from its glucosyl ester. *J. Exp. Bot.* 71, 2058–2071. doi: 10.1093/jxb/erz528

Hirano, K., Ueguchi-Tanaka, M., and Matsuoka, M. (2008). GID1-mediated gibberellin signaling in plants. *Trends Plant Sci.* 13, 192–199. doi: 10.1016/j.tplants.2008.02.005

Horvath, E., Csiszar, J., Galle, A., Poor, P., Szepesi, A., and Tari, I. (2015). Hardening with salicylic acid induces concentration-dependent changes in abscisic acid biosynthesis of tomato under salt stress. *J. Plant Physiol.* 183, 54–63. doi: 10.1016/j.jplph.2015.05.010

Hu, E., Liu, M., Zhou, R., Jiang, F., and Wu, Z. (2021). Relationship between melatonin and abscisic acid in response to salt stress of tomato. *Sci. Hortic.* 285:110176. doi: 10.1016/j.scienta.2021.110176

Hwang, O. J., and Back, K. (2018). Melatonin is involved in skotomorphogenesis by regulating brassinosteroid biosynthesis in rice plants. *J. Pineal Res.* 65:e12495. doi: 10.1111/jpi.12495

Hwang, O. J., and Back, K. (2019). Melatonin deficiency confers tolerance to multiple abiotic stresses in rice via decreased brassinosteroid levels. *Int. J. Mol. Sci.* 20:5173. doi: 10.3390/ijms20205173

Ismail, A. M., and Horie, T. (2017). Genomics, physiology, and molecular breeding approaches for improving salt tolerance. *Annu. Rev. Plant Biol.* 68, 405–434. doi: 10.1146/annurev-arplant-042916-040936

Jiang, M., and Zhang, J. (2001). Effect of abscisic acid on active oxygen species, antioxidative defence system and oxidative damage in leaves of maize seedlings. *Plant Cell Physiol.* 42, 1265–1273. doi: 10.1093/pcp/pce162

Jin, H., Choi,. S. M., Kang, M., Yun, S., Kwon, D. J., Noh, Y. S., et al. (2018). Salicylic acid-induced transcriptional reprogramming by the HAC-NPR1-TGA histone acetyltransferase complex in Arabidopsis. *Nucleic Acids Res.* 46, 11712–11725. doi: 10.1093/nar/gky847

Kanai, M., Higuchi, K., Hagihara, T., Konishi, T., Ishii, T., Fujita, N., et al. (2007). Common reed produces starch granules at the shoot base in response to salt stress. *New Phytol.* 176, 572–580. doi: 10.1111/j.1469-8137.2007.02188.x

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

Kohli, A., Sreenivasulu, N., Lakshmanan, P., and Kumar, P. (2013). The phytohormone crosstalk paradigm takes center stage in understanding how plants respond to abiotic stresses. *Plant Cell Rep.* 32, 945–957. doi: 10.1007/s00299-013-1461-y

Li, B., Liu, Y., Cui, X., Fu, J., Zhou, Y., Zheng, W., et al. (2019a). Genome-wide characterization and expression analysis of soybean TGA transcription factors

identified a novel TGA gene involved in drought and salt tolerance. *Front. Plant Sci.* 10:549. doi: 10.3389/fpls.2019.00549

Li, C., He, Q., Zhang, F., Yu, J., Li, C., Zhao, T., et al. (2019b). Melatonin enhances cotton immunity to *Verticillium* wilt via manipulating lignin and gossypol biosynthesis. *Plant J.* 100, 784–800. doi: 10.1111/tpj.14477

Li, S., Liu, J., An, Y., Cao, Y., Liu, Y., Zhang, J., et al. (2019c). *MsPIP2;2*, a novel aquaporin gene from *Medicago sativa*, confers salt tolerance in transgenic Arabidopsis. *Environ. Exp. Bot.* 165, 39–52. doi: 10.1016/j.envexpbot.2019.05.020

Libalweksler, Y., Nir, M., Benhayyim, G., and Telor, E. (1994). Starch metabolism in salt-tolerant and salt-sensitive shamouti callus. *Plant Physiol. Biochem.* 32, 655–659.

Liu, J., Shabala, S., Zhang, J., Ma, G., Chen, D., Shabala, L., et al. (2020). Melatonin improves rice salinity stress tolerance by NADPH oxidase-dependent control of the plasma membrane K^+ transporters and K^+ homeostasis. *Plant Cell Environ.* 43, 2591–2605. doi: 10.1111/pce.13759

Liu, J., Zhai, R., Liu, F., Zhao, Y., Wang, H., Liu, L., et al. (2018). Melatonin induces parthenocarpy by regulating genes in gibberellin pathways of 'Starkrimson' Pear (*Pyrus communis* L.). *Front. Plant Sci.* 9:946. doi: 10.3389/fpls.2018.00946

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

Long, R., Wang, H., Shen, Y., Kang, J., Zhang, T., Sun, Y., et al. (2014). Molecular cloning and functional analysis of a salt-induced gene encoding an RNA-binding protein in alfalfa. *Mol. Breed.* 34, 1465–1473. doi: 10.1007/s11032-014-0130-3

Long, T. A., and Benfey, P. N. (2006). Transcription factors and hormones: new insights into plant cell differentiation. *Curr. Opin. Cell Biol.* 18, 710–714. doi: 10.1016/j.ceb.2006.09.004

Martinez, V., Nieves-Cordones, M., Lopez-Delacalle, M., Rodenas, R., Mestre, T., Garcia-Sanchez, F., et al. (2018). Tolerance to stress combination in tomato plants: new insights in the protective role of melatonin. *Molecules* 23, 1–20. doi: 10.3390/molecules23030535

Mcadam, S. A. M., Sussmilch, F. C., Brodribb, T. J., and Ross, J. J. (2015). Molecular characterization of a mutation affecting abscisic acid biosynthesis and consequently stomatal responses to humidity in an agriculturally important species. *AoB Plants* 7, 634–639. doi: 10.1093/aobpla/plv091

Miller, G., Suzuki, N., Ciftci-Yilmaz, S., and Mittler, R. (2010). Reactive oxygen species homeostasis and signalling during drought and salinity stresses. *Plant Cell Environ.* 33, 453–467. doi: 10.1111/j.1365-3040.2009.02041.x

Minoru, K., Michihiro, A., Susumu, G., Masahiro, H., Mika, H., Masumi, I., et al. (2008). KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* 36, 480–484. doi: 10.1093/nar/gkm882

Pang, Q., Zhang, A., Zang, W., Wei, L., and Yan, X. (2016). Integrated proteomics and metabolomics for dissecting the mechanism of global responses to salt and alkali stress in *Suaeda corniculata*. *Plant Soil.* 402, 379–394. doi: 10.1007/s11104-015-2774-0

Paredes, S. D., Korkmaz, A., Manchester, L. C., Tan, D. X., and Reiter, R. J. (2009). Phytomelatonin: a review. *J. Exp. Bot.* 60, 57–69. doi: 10.1093/jxb/ern284

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

Puckette, M. C., Weng, H., and Mahalingam, R. (2007). Physiological and biochemical responses to acute ozone-induced oxidative stress in *Medicago truncatula*. *Plant Physiol. Biochem.* 45, 70–79. doi: 10.1016/j.plaphy.2006.12.004

Reiter, R. J., Mayo, C., Tan, D. X., Sainz, R., Alatorre-Jimenez, M., and Qin, L. (2016). Melatonin as an antioxidant: Under promises but over delivers. *J. Pineal Res.* 61, 253–278. doi: 10.1111/jpi.12360

Shani, E., Salehin, M., Zhang, Y., Sanchez, S. E., Doherty, C., Wang, R., et al. (2017). Plant stress tolerance requires auxin-sensitive AUX/IAA transcriptional repressors. *Curr. Biol.* 27, 437–444. doi: 10.1016/j.cub.2016.12.016

Sharma, A., Wang, J., Xu, D., Tao, S., Chong, S., Yan, D., et al. (2020). Melatonin regulates the functional components of photosynthesis, antioxidant system, gene expression, and metabolic pathways to induce drought resistance in grafted *Carya cathayensis* plants. *Sci. Total Environ.* 713:136675. doi: 10.1016/j.scitotenv.2020.136675

Shi, H., Jiang, C., Ye, T., Tan, D., Russel, J. R., Zhang, H., et al. (2015). Comparative physiological, metabolomic, and transcriptomic analyses reveal mechanisms of improved abiotic stress resistance in bermudagrass [*Cynodon dactylon* (L). Pers.] by exogenous melatonin. *J. Exp. Bot.* 66, 681–694. doi: 10.1093/jxb/eru373

Shi, S., Nan, L., and Smith, K. (2017). The current status, problems, and prospects of alfalfa (*Medicago sativa* L.) breeding in China. *Agronomy* 7:1. doi: 10.3390/agronomy7010001

Siddiqui, M., Alamri, S., Al-Khaishany, M., Khan, M., Al-Amri, A., Ali, H., et al. (2019). Exogenous melatonin counteracts NaCl-induced damage by regulating the antioxidant system, proline and carbohydrates metabolism in tomato seedlings. *Int. J. Mol. Sci.* 20:353. doi: 10.3390/ijms20020353

Singer, S. D., Hannoufa, A., and Acharya, S. (2018). Molecular improvement of alfalfa for enhanced productivity and adaptability in a changing environment. *Plant Cell Environ.* 41, 1955–1971. doi: 10.1111/pce.13090

Soma, F., Takahashi, F., Suzuki, T., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2020). Plant Raf-like kinases regulate the mRNA population upstream of ABA-unresponsive SnRK2 kinases under drought stress. *Nat. Commun.* 11:1373. doi: 10.1038/s41467-020-15239-3

Srivastava, A., Rai, A., Patade, V., and Suprasanna, P. (2013). "Calcium signaling and its significance in alleviating salt stress in plants," in *Salt Stress in Plants*, eds P. Ahmad, M. M. Azooz, and M. N. V. Prasad (New York, NY: Springer). doi: 10.1007/978-1-4614-6108-1_9

Sun, C., Liu, L., Wang, L., Li, B., Jin, C., and Lin, X. (2021). Melatonin: A master regulator of plant development and stress responses. *J. Integr. Plant Biol.* 63, 126–145. doi: 10.1111/jipb.12993

Thalmann, M., Pazmino, D., Seung, D., Horrer, D., Nigro, A., Meier, T., et al. (2016). Regulation of leaf starch degradation by abscisic acid is important for osmotic stress tolerance in plants. *Plant Cell* 28, 1860–1878. doi: 10.1105/tpc.16.00143

Thitisaksakul, M., Dong, S., and Beckles, D. M. (2017). How rice Glycogen Synthase Kinase-like 5 (OsGSK5) integrates salinity stress response to source-sink adaptation: a proposed model. *Plant Signal. Behav.* 12:e1403708. doi: 10.1080/15592324.2017.1403708

Verma, V., Ravindran, P., and Kumar, P. P. (2016). Plant hormone-mediated regulation of stress responses. *BMC Plant Biol.* 16:86. doi: 10.1186/s12870-016-0771-y

Wang, D., Chen, Q., Chen, W., Guo, Q., Xia, Y., Wang, S., et al. (2021). Physiological and transcription analyses reveal the regulatory mechanism of melatonin in inducing drought resistance in loquat (*Eriobotrya japonica* Lindl.) seedlings. *Environ. Exp. Bot.* 181:104291. doi: 10.1016/j.envexpbot.2020.104291

Wang, L., Wang, S., and Wei, L. (2012). RSeQC: quality control of RNA-seq experiments. *Bioinformatics* 28, 2184–2185. doi: 10.1093/bioinformatics/bts356

Wang, Y., Zhang, Z., Zhang, P., Cao, Y., Hu, T., and Yang, P. (2016). Rhizobium symbiosis contribution to short-term salt stress tolerance in alfalfa (*Medicago sativa* L.). *Plant Soil.* 402, 247–261. doi: 10.1007/s11104-016-2792-6

Wei, J., Li, D., Zhang, J., Shan, C., Rengel, Z., Song, Z., et al. (2018). Phytomelatonin receptor PMTR1-mediated signaling regulates stomatal closure in *Arabidopsis thaliana*. *J. Pineal Res.* 65:e12500. doi: 10.1111/jpi.12500

Wei, W., Li, Q., Chu, Y., Reiter, R., Yu, X., Zhu, D., et al. (2015). Melatonin enhances plant growth and abiotic stress tolerance in soybean plants. *J. Exp. Bot.* 66, 695–707. doi: 10.1093/jxb/eru392

Wen, D., Gong, B., Sun, S., Liu, S., Wang, X., Wei, M., et al. (2016). Promoting roles of melatonin in adventitious root development of *Solanum lycopersicum* L. by regulating auxin and nitric oxide signaling. *Front. Plant Sci.* 7:718. doi: 10.3389/fpls.2016.00718

Winter, H., and Huber, S. C. (2000). Regulation of sucrose metabolism in higher plants: localization and regulation of activity of key enzymes. *Crit. Rev. Biochem. Mol. Biol.* 35, 253–289. doi: 10.1080/10409230008984165

Xu, L., Xiang, G., Sun, Q., Ni, Y., Jin, Z., Gao, S., et al. (2019). Melatonin enhances salt tolerance by promoting *MYB108A*-mediated ethylene biosynthesis in grapevines. *Hortic. Res.* 6:114. doi: 10.1038/s41438-019-0197-4

Yan, Y., Jing, X., Tang, H., Li, X., Gong, B., and Shi, Q. (2019). Using transcriptome to discover a novel melatonin-induced sodic alkaline stress resistant pathway in *Solanum lycopersicum* L. *Plant Cell Physiol.* 9, 2051–2064. doi: 10.1093/pcp/pcz126

Yang, Y., Gao, S., Su, Y., Lin, Z., Guo, J., Li, M., et al. (2019). Transcripts and low nitrogen tolerance: regulatory and metabolic pathways in

sugarcane under low nitrogen stress. *Environ. Exp. Bot.* 163, 97–111. doi: 10.1016/j.envexpbot.2019.04.010

Yang, Y., and Guo, Y. (2018). Elucidating the molecular mechanisms mediating plant salt-stress responses. *New Phytol.* 217, 523–539. doi: 10.1111/nph.14920

Yin, Y., Kobayashi, Y., Sanuki, A., Kondo, S., Fukuda, N., Ezura, H., et al. (2010). Salinity induces carbohydrate accumulation and sugar-regulated starch biosynthetic genes in tomato (*Solanum lycopersicum* L. cv. 'Micro-Tom') fruits in an ABA- and osmotic stress-independent manner. *J. Exp. Bot.* 61, 563–574. doi: 10.1093/jxb/erp333

Young, M. D., Wakefield, M. J., Smyth, G. K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 11:R14. doi: 10.1186/gb-2010-11-2-r14

Yu, R., Wang, G., Yu, X., Li, L., Li, C., Song, Y., et al. (2021). Assessing alfalfa (*Medicago sativa* L.) tolerance to salinity at seedling stage and screening of the salinity tolerance traits. *Plant Biol.* 23, 664–674. doi: 10.1111/plb.13271

Zhan, Y., Wu, T., Zhao, X., Wang, Z., and Chen, Y. (2021). Comparative physiological and full-length transcriptome analyses reveal the molecular mechanism of melatonin-mediated salt tolerance in okra (*Abelmoschus esculentus* L.). *BMC Plant Biol.* 21:180. doi: 10.1186/s12870-021-02957-z

Zhang, H., and Zhang, Y. (2014). Melatonin: a well-documented antioxidant with conditional pro-oxidant actions. *J. Pineal Res.* 57, 131–146. doi: 10.1111/jpi.12162

Zhang, N., Zhang, H., Zhao, B., Sun, Q., Cao, Y., Li, R., et al. (2014). The RNA-seq approach to discriminate gene expression profiles in response to melatonin on cucumber lateral root formation. *J. Pineal Res.* 56, 39–50. doi: 10.1111/jpi.12095

Zhang, X., Feng, B., Wang, H., Xu, X., Shi, Y., He, Y., et al. (2018). A substitution mutation in *OsPELOTA* confers bacterial blight resistance by activating the salicylic acid pathway. *J. Integr. Plant Biol.* 60, 160–172. doi: 10.1111/jipb.12613

Zhang, Y., Fan, Y., Rui, C., Zhang, H., Xu, N., Dai, M., et al. (2021). Melatonin improves cotton salt tolerance by regulating ROS scavenging system and Ca^{2+} signal transduction. *Front. Plant Sci.* 12:693690. doi: 10.3389/fpls.2021.693690

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

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

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 02 September 2022
DOI 10.3389/fpls.2022.938859

OPEN ACCESS

EDITED BY
Wengang Xie,
Lanzhou University,
China

REVIEWED BY
Guangdi Li,
NSW Government,
Australia
Shufeng Zhou,
Sichuan Agricultural University,
China

*CORRESPONDENCE
Ruyu He
243535990@qq.com

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

RECEIVED 08 May 2022
ACCEPTED 08 August 2022
PUBLISHED 02 September 2022

CITATION
Cheng M, Pan Z, Cui K, Zheng J, Luo X, Chen Y, Yang T, Wang H, Li X, Zhou Y, Lei X, Li Y, Zhang R, Iqbal MZ and He R (2022) RNA sequencing and weighted gene co-expression network analysis uncover the hub genes controlling cold tolerance in *Helictotrichon virescens* seedlings. *Front. Plant Sci.* 13:938859. doi: 10.3389/fpls.2022.938859

RNA sequencing and weighted gene co-expression network analysis uncover the hub genes controlling cold tolerance in *Helictotrichon virescens* seedlings

Mingjun Cheng[1,2], Zeyang Pan[3], Kuoshu Cui[4], Junjun Zheng[3], Xuan Luo[5], Youjun Chen[1], Tao Yang[3], Hui Wang[1], Xiaofeng Li[3], Yang Zhou[3], Xiong Lei[6], Yingzheng Li[3], Ruizhen Zhang[2], Muhammad Zafar Iqbal[3,7] and Ruyu He[3]*

[1]Institute of Qinghai-Tibetan Plateau, Southwest Minzu University, Chengdu, China, [2]Sichuan Grass Industry Technology Research and Promotion Center, Chengdu, China, [3]Maize Research Institute, Sichuan Agricultural University, Chengdu, China, [4]Sichuan Agricultural Technology Extension Station, Chengdu, China, [5]Institute of Agricultural Information and Rural Economy, Sichuan Academy of Agricultural Sciences, Chengdu, China, [6]Sichuan Academy of Grassland Science, Chengdu, China, [7]College of Grassland Science and Technology, Sichuan Agricultural University, Chengdu, China

Helictotrichon virescens is a perennial herbaceous plant with a life expectancy of about 10 years. It has high cold and heat resistance and can successfully survive over winter in the habitats with a temperature range of −25 to 25°C. Therefore, this study aimed to identify the key genes regulating low-temperature stress responses in *H. virescens* and analyze cold tolerant at molecular level. This study used RNA sequencing (RNA-Seq) and weighted gene co-expression network analysis (WGCNA) to identify the hub genes associated with cold tolerance in *H. virescens*. RT-PCR was conducted, homologous genes were identified, and related bioinformatics were analyzed to verify the identified hub genes. Moreover, WGCNA analysis showed that only the brown module had the highest correlation with the active-oxygen scavenging enzymes [peroxide (POD), superoxide dismutase (SOD), and catalase (CAT)]. The expression levels of three hub genes in the brown module (Cluster-37118.47362, cluster-37118.47713, and cluster-37118.66740) were significantly higher under low-temperature stress than those under control conditions. Furthermore, gene ontology (GO) and KEGG annotations showed that the three hub genes were mainly enriched in the metabolism pathways of sphingolipids, selenocompounds, glyoxylate, and dicarboxylate, carotenoids biosynthesis, and other biological pathways. The results of this study also showed that the subcellular localization prediction results showed that the cold tolerance hub genes were all localized to the plasma membrane. By constructing a protein interaction network, it was found that the hub gene Cluster-37118.66740 interacted with Sb09g003460.1 and Sb04g020180.1 proteins in *Sorghum bicolor*. By constructing phylogenetic trees of the four

species of *H. virescens*, *Sorghum bicolo*, *Oryza sativa Japonica*, and *Arabidopsis thaliana*, the results showed that, the hub gene Cluster 37118.66740 (of *H. virescens*) and Os03g0340500 (of *Oryza sativa Japonica*) belonged to the same ancestral branch and were in the same subfamily. Thus, this study provides methodology and guidance to identify the cold tolerance genes for other herbage and their cold tolerant molecular mechanisms at molecular level.

KEYWORDS

***Helictotrichon virescens*, perennial herbs, cold tolerance, WGCNA, active-oxygen scavenging enzymes, homologous gene identification**

Introduction

Qinghai-Tibet Plateau region has a cold continental climate with daily temperature fluctuations. In this region, the winter is long and cold, with extremely low temperatures. Cold stress and other disasters occur from time to time, resulting in short growing period with small plants and low biomass of grassland vegetation, and even unable to survive over winter normally which leads to plant death. Low-temperature stress severely restricts plant growth and grassland renovation in Qinghai-Tibet Plateau.

Helictotrichon virescens is a Tetraploid perennial herb with about 100 species distributed across Asia, Europe, and North America. More than 20 species are distributed across the provinces of China; for example, there are about 11 species and one variant in Sichuan, and other nine species and one variant in the high-altitude areas of Ganzi, Aba, and Liangshan. These species are mainly distributed in the high-altitude areas of 2,000–4,500 m, where they serve as an important part of the wild pastures. *Helictotrichon virescens* has many leaves, soft blades, lower crude fiber content (than the old Mangosteen and Elysia), high nutritional value, and good palatability for ruminants. It has high economic values in terms of output and quality and has been widely promoted in the northwest plateau of Sichuan. The grass has a lifespan of about 10 years with high cold and heat resistance and can successfully overwinter in low-temperature habitats (−25 to 25°C; Cheng et al., 2022). *Helictotrichon virescens* has been playing an irreplaceable role in promoting animal husbandry in the alpine grassland, and has become an important source of high-quality forage for livestock supplement during winter and spring seasons in the alpine pastoral areas of the Qinghai-Tibet Plateau.

The process of crops adapting to low-temperature stress can be divided into several levels, such as changes in biofilm structure, cell osmotic regulators, antioxidant enzymes, and changes in photosynthesis parameters. Low-temperature stress affects the stability of the plant cell membrane systems (Hoffman et al., 2010), which is considered as the main target of low-temperature damage. The imbalance of ions and osmolates inside and outside the cell destroys the stability of the cell membrane system, which in turn affects the photosynthesis and respiration of plants. Change in the cell membrane system is the initial response mechanism of turfgrass chilling injury or freezing injury response (Levitt, 1980). Turfgrass produces the large amounts of reactive oxygen species (ROS) under low-temperature stress, including hydrogen peroxide (H_2O_2), superoxide radical (O^{2-}), hydroxyl radical (-OH), etc., which in turn induce antioxidant enzyme systems activation, including peroxides Enzyme (POD), superoxide dismutase (SOD), catalase (CAT), and other anti-stress reactions, which collectively remove ROS and free radicals from plants, thereby alleviate the damage of turfgrass caused by low-temperature stress (Sigaud-Kutner et al., 2010). Under the stress of low temperature, plants can also reduce the cold damage by adjusting the osmotic balance of cells with the accumulation or decomposition of osmotic regulating substances. In recent years, a large number of studies have demonstrated that the contents of free proline, soluble sugar, and soluble protein are closely related to the stress tolerance mechanism in plants.

Physiological, transcriptomic, proteomic, and metabolomic changes during cold acclimation of perennial herbs have been extensively studied (Hoffman et al., 2010; Bocian et al., 2011, 2015; Paina et al., 2014; Abeynayake et al., 2015; Augustyniak et al., 2018). For example, a major quantitative trait locus QTL in *LpCBFIIIc* was associated with low-temperature tolerance, which was identified in 109 perennial ryegrass plants (Hulke et al., 2012). Candidate gene association mapping revealed that *LpCBF1b* was closely associated with winter survival, while *LpLEA3*, *LpMn SOD*, *LpCAT*, and *LpChl*, *Cu-Zn SOD* were mainly associated with spring plant regeneration (Yu et al., 2015). Heterologous overexpression of *AtDREB1A/CBF3* increased cold stress tolerance in ryegrass plants (Li et al., 2011). Transcriptomic analysis revealed that 60 cytochrome P450s transcripts were upregulated and 59 transcripts were downregulated in tall fescue under low-temperature stress, of which 26 and 17 transcripts were involved in the metabolism of flavonoids and brassinosteroids, respectively. Cold stress regulation suggests that P450s play a

crucial role in the cold stress response of tall fescue (Tao et al., 2017). Zhao et al. (2020) studied cold-tolerant material "10–122" and low-temperature sensitive material "09–126" of *Poa pratensis* in native grassland of Qinghai Province by transcriptome sequencing technology to analyze differentially expressed genes between low-temperature stress and normal temperature control. The enrichment analysis of differentially expressed genes in the two materials showed that the differentially expressed genes were significantly enriched in photosynthesis, REDOX reaction, carbohydrate metabolism, cell membrane system, transporter, and sub-biological metabolism under low-temperature stress. In addition, some genes of calcium signaling regulation, hormone metabolism, signal transduction, antioxidant system, carbohydrate metabolism, and other pathways were only upregulated in cold-tolerant germplasm "10–122," which have the potential to be used as cold-tolerant, such as *CML*, *CPK*, *CALM*, *DHAR*, *GST*, *NCED*, *SNRK2*, *BSK*, *CKX*, *BIN2*, *ARF*, and *PEK*, etc.

Due to the lack of genomic information of *H. virescens*, the current research on cold tolerance of *H. virescens* is mainly limited to physiological indicators. The molecular mechanism and regulatory network of *H. virescens* in response to low-temperature stress are still unclear. Therefore, in this study, RNA sequencing was performed on *H. virescens* after treating it under low temperature, and three hub genes significantly relating to protective enzymes were mined through the analysis strategy of WGCNA. RT-PCR validation, homologous gene mining, and related bioinformatics analysis were performed. This study lays the foundation for an in-depth understanding of the response mechanism of perennial herbs to low-temperature stress.

Materials and methods

Plant materials and measurements

Uniformly sized and plump *H. virescens* seeds were selected and sown into pots (21 cm in diameter and 16 cm in height; 50–60 seeds per pot). The pot mix contained a mixture of peat, pine needles, and yellow clay in a volume ratio of 3:1:1, according to Cheng et al. (2022). The seedlings with thinned to 20 plants per pot after the first leaf had fully expanded. The plants were watered daily using a half-strength Hogland nutrient solution [2.5 mmol L^{-1} $Ca(NO_3)_2$, 2.5 mmol L^{-1} KNO_3, 1 mmol L^{-1} $MgSO_4$, 0.5 mmol L^{-1} KH_2PO_4, 45 μmol L^{-1} Fe-EDTA, 23 μmol L^{-1} H_3BO_3, 4.55 μmol L^{-1} $MnCl_2$, 0.16 μmol L^{-1} $CuSO_4$, 0.38 μmol L^{-1} $ZnSO_4$, and 0.28 μmol L^{-1} Na_2MoO_4]. After 5-week growth, seedlings were placed in a constant temperature and light incubator (MLR-352H-PC) simulating low-temperature stress = (0°C), with the illumination of 3000 Lx. The control temperature was 25°C. The third leaf of each plant was collected after 12, 36, and 60 h of the low-temperature treatment and immediately frozen in liquid nitrogen, and stored in a −80°C ultra-low temperature refrigerator for the determination of related indicators. Each treatment had three biological replicates.

Phenotypic data analysis

At the end of the experiment, 30 seedlings with relatively consistent growth were selected from the treated and control groups for the relative leaf conductivity using assay developed by Cheng et al. (2022). Briefly, the middle section (2 cm × 4 cm) of the first fully expanded leaf were obtained from each seedling, and then mixed and cut into 1 cm pieces. The pieces were subsequently divided into three parts placed in 10 ml EP tubes, which were then filled with distilled water. After soaking for 3 h, the Ec1 of the leaves was measured using a conductivity meter, while the Ec2 was measured using the same instrument after the leaves being incubated in a boiling water bath for 10 min and cooled to room temperature. The relative conductivity was calculated using formula: REC = EC1/EC2 × 100%.

Chlorophyll a and b were extracted from 0.2 g of fresh leaves using 95% ethanol at room temperature, as described by Cheng et al. (2022). Briefly, the homogenate was centrifuged at 10,000 *g* for 10 min, and the chloroplast pigment extract was aliquoted into 1 cm cuvette to measure the absorbance at wavelengths 663 and 646 nm using 95% ethanol as blank. The chlorophyll concentration was calculated using formula:

$$\text{Chlorophyll a} = 12.21 \times OD_{663} - 2.81 \times OD_{646}$$

$$\text{Chlorophyll b} = 20.13 \times OD_{646} - 5.03 \times OD_{663}$$

Pro, SOD, POD, catalase (CAT), and ROS activities were measured using kits purchased from China Quanzhou Ruixin Biological Technology Co., LTD, and the type is Ruixinbio (Quanzhou, China).

The mean, standard error (SE), maximum and minimum values, and coefficient of variation, kurtosis, and skewness of each trait were calculated in Microsoft Excel. ANOVA of the collected phenotypic traits was conducted using SPSS (Statistical Product and Service Solutions, version 21.0, IBM, Armonk, NY, United States). The correlation analysis of phenotypic traits was carried out using the PerformanceAnalytics package in R software.

RNA-seq, WGCNA, and hub gene identification

Three biological replicates of the leaf samples (third leaf of each plant collected after 12, 36, and 60 h of low-temperature treatment) stored at −80°C were processed for RNA-seq. The RNA extraction, detection, cDNA library construction, and sequencing were conducted at Beijing Nuohezhiyuan Technology Co., Ltd. Utilizing Oligo(dT) magnetic beads enriched mRNA with polyA tail procedures. Briefly, mRNA was randomly fragmented by divalent cations in NEB Fragmentation Buffer (NEB, Ipswich, MA, United States). The first strand of cDNA was prepared using random oligonucleotides as a primer, and the second strand was synthesized using DNA polymerase I. The

ends of purified double-stranded cDNA were repaired by adding a tail and sequencing connector and screened to 250–300 bp cDNA for PCR amplification. The PCR product was purified again to obtain the final library, and after qualifying the library. High-quality sequencing libraries were sequenced on the Illumina HiseqTM 4,000 sequencing platform, statistical Q20 (Phred = −10log10 [e]), Q30 (Phred = −10log10 [e])], GC content and sequencing error (< 6%), and other indicators of sequencing quality control were used for these procedures (Cheng et al., 2022), and Trinity software (Grabherr et al., 2011) was used to splice clean reads into transcripts (Unigenes), which were used as reference sequences for subsequent analysis. The splicing transcript was sequenced based on its length from long to short, and the length of transcript was added to the length of splicing transcript, so that it was not <50/90% of the total length, namely N50/N90, to measure the continuity of *de novo* assembly. Its numerical value can be used to evaluate the quality of the assembly. The obtained Unigenes were functionally annotated using Blast (Altschul, 2012) on seven databases, including non-redundant (NR), nucleotide (NT), gene ontology (GO), EuKaryotic Orthologous Groups (KOG), KEGG Orthology (KO), Swiss-Prot, and FPKM. Thereafter, clean reads from each sample were aligned to the reference sequence (obtained by Trinity splicing), and the read counts of the aligned genes were obtained using Bowtie2. The number of read counts was converted to FPKM values (expected number of Fragments per Kilobase of transcript sequence per Millions base pairs sequenced; Trapnell et al., 2010) and used to evaluate the gene expression levels (FPKM > 0.3 was regarded as gene expression). Quantitative gene expression analysis was conducted using the RSEM method (Dewey and Li, 2011).

Differentially expressed genes were determined from the different samples using statistical analysis based on gene expression levels, and their original read counts were normalized with DEG-Seq software (Love et al., 2014). Negative binomial distribution (*p* value) was used for hypothesis testing, followed by multiple hypothesis test correction using Benjamini-Hochberg (BH) method to obtain the FDR value (False Discovery Rate, or padj). The obtained FDR values were then used to screen for the differential genes using $|\log_2$ (fold change)$| > 1$&padj < 0.05 as the standard.

The differentially expressed genes were constructed according to the normalized FPKM values of obtained transcriptome data from 18 cold stress treatments (including control) for weighted gene co-expression network analysis (WGCNA). The R package used for this analysis is presented in Supplementary Paper 1. The WGCNA program parameter settings were: variance data expression >0; no missing data expression <0.1; soft threshold = 10 (estimated value); maximum block size = 2,000; depth split = 4; minimum block size = 50; and Merge cut height = 0.1. Moreover, hub genes were screened in each module using the connection value (|KME|) > 0.95, module membership (MM) > 0.9, and gene significance (GS) > 0.9 of eigengenes, and functionally annotated. Cytoscape.v3.9.1 was used to draw the local transcriptional regulatory network.

GO and KEGG enrichment hub genes

The GO enrichment analysis of the hub genes was performed using GOseq method (Young et al., 2010), while KOBAS method (Kanehisa et al., 2008) was used for the KEGG metabolism and signal transduction pathway enrichment analysis of the hub genes. In both analyses, padJ < 0.05 was used as the threshold for significant enrichment.

Real-time fluorescence quantitative analysis of the hub genes

The RNA was extracted from the samples (stored at −80°C) using the HiPure Plant RNA Mini Kit (Magen), according to the manufacturer's instructions. The concentration and quality (OD_{260}/OD_{280} value) of the RNA samples were determined using the NanoVue plus Spectrophotometer. Subsequently, total RNA was used as a template for the first-strand cDNA synthesis using the RevertAid First Strand cDNA SynthesisKit (TaKaRa), which also contains recombinant endonuclease (DNase I) for removing genomic DNA contamination from RNA samples. Briefly, DNase I was added to the RNA samples and incubated at 37°C for 30 min for DNA removal. Thereafter, 1 µl of 50 mM EDTA was added, and the mixture was incubated at 65°C for 10 min to inactivate the DNase I. Reverse transcription was then performed by adding the reverse transcriptase to the mixture and incubating at 42°C for 60 min. The samples were then stored at −20°C for later use. The reverse transcription process is shown in Supplementary Paper 2.

The real-time fluorescence quantitative (qRT-PCR) reaction system (CFX96 real-time PCR system; Bio-Rad) was sterilized before configuration, and the samples were mixed and loaded into the instrument. The real-time fluorescence quantitative reaction system is shown in Supplementary Paper 3. The reaction involved three biological and four technical replicates, as shown in the procedure presented in Supplementary Paper 4. The qRT-PCR results were analyzed by the $2^{-\triangle\triangle Ct}$ method with GAPDH as the internal reference; it is widely present in many organisms and is abundant in cells, accounting for 10–20% of the total protein. The GAPDH gene has a highly conserved sequence, and the protein expression level in the same cell or tissue is generally constant. Therefore, this gene has been widely used as an internal reference gene in qPCR for a long time. In this study, GAPDH is the internal reference gene of naked oats (from the Wheat Research Institute of Sichuan Agricultural University).

Homologous EST label identification of hub gene

The Open Reading Frame Finder tool from NCBI was used to align the RNA-Seq data to the open reading frames (ORFs) of the transcripts obtained from the reference genomes of *Avena barbata* and *Hordeum vulgare* subsp. The functions of the expressed genes

TABLE 1 Statistical description of *Helictotrichon virescens* phenotypic data at 25°C.

Trait	Range	Mean	SE	CV/%	Kurtosis	Skewness	Sig
Pro	0.0115–0.4040	0.2635	0.0631	0.2397	−1.7140	−0.8400	ns
Rec	0.0371–0.0398	0.0385	0.0004	0.0100	−1.7140	0.0130	ns
Chla	0.0028–0.0029	0.0029	0.0000	0.0013	−1.7140	−0.7480	ns
Chlb	0.0016–0.0018	0.0017	0.0000	0.0170	−1.7140	0.0000	ns
POD	1372.48–1392.48	1382.4800	2.8868	0.0021	−1.7140	0.0000	ns
SOD	698.47–811.64	749.7333	16.5500	0.0221	−1.7140	0.3990	ns
CAT	467.93–518.90	493.6233	7.3576	0.0149	−1.7140	−0.0360	ns
ROS	3507.94–3755.02	3601.4067	38.7033	0.0107	−1.7140	0.7980	ns

ns, not significant.

are determined by comparing and analyzing the obtained Expressed Sequence Tag (EST) with the known sequences in various public databases. The ORFs corresponding to the previously screened hub gene were used to search the EST and Genbank databases on the NCBI official website[1] to obtain the hub genes. Thereafter, the matched base sequences were screened for homologous genes, using the alignment thresholds of identity>85% and Query Coverage>30%.

Bioinformatics analysis of hub genes and construction of protein interaction network

First, extract the hub genes sequence from the non-redundant gene database obtained by RNA-seq sequencing and splicing of *H. virescens*. Then, the hub genes sequence was aligned to the NR database and the Swissprot protein database, the ORF coding frame information corresponding to the transcript was extracted from the alignment result, and the coding region sequence was translated into the amino acid sequence according to the standard codon table (according to the 5′- >3′ order). Finally, the ORF of the hub gene was predicted by ESTSCAN[2] software based on PSM and codon preference, there by predicting the nucleic acid sequence and amino acid sequence encoded by the hub genes.

The online analysis website ExPASy[3] was used to predict the protein physicochemical properties of the hub gene; Subcellular localization prediction was performed using WoLF PSORT[4]; Protein transmembrane domain prediction was performed by the online tool TMHMM 2.0[5]; The online analysis website STRING[6] was used, the monocot genome annotation information was selected as the background file for comparison, and the protein interaction network was constructed for the hub genes.

Identification and phylogenetic analysis of homologous genes of hub genes

The predicted amino acid sequence of the hub genes was used as a probe, the *Sorghum bicolor*, *Oryza sativa Japonica*, and *Arabidopsis thaliana* were identified by alignment and domain alignment in pfam, CDD and SMART databases, and homologous genes were located with similarity >50 and E-value ≤ 1e−10. Among them, data files were downloaded from the National Center for Biotechnology Information (National Center for Biotechnology Information) database. After multiple sequence alignment using mafft software, the phylogenetic tree of the gene family was constructed by JTT + G4 method using IQ-tree software, and the phylogenetic tree was beautified using EvolView online website platform.

Results

Evaluation of cold-tolerance phenotypes of *Helictotrichon virescens*

The cold tolerance of *H. virescens* was evaluated at its seedling stage in a simulated cold-stress environment in the laboratory. The results showed that there were no significant differences in Pro, Rec, Chla, Chlb, POD, SOD, CAT, and ROS levels in *H. virescens* seedlings grown at 25°C (Table 1). The coefficient of variation of Pro was the largest (0.2397), while POD was the smallest (0.0021); however, the kurtosis of each indicator was similar. Moreover, ROS exhibited the largest skewness (0.7980), and Chlb and POD had the smallest (0.0000). As shown in Figure 1A, Pro had a significant negative correlation with Chlb ($p < 0.001$), but positively correlated with SOD ($p < 0.001$) and CAT ($p < 0.0001$). Meanwhile, Chlb significantly negatively correlated with SOD ($p < 0.001$) and CAT ($p < 0.001$). POD and SOD showed a significant positive correlation ($p < 0.05$), while SOD and CAT

1 http://www.ncbi.nlm.nih.gov/dbEST/dbEST_summary.html
2 https://myhits.sib.swiss/cgi-bin/estscan
3 https://web.expasy.org/protparam/
4 https://wolfpsort.hgc.jp/
5 https://services.healthtech.dtu.dk/service.php?TMHMM-2.0
6 https://cn.string-db.org/

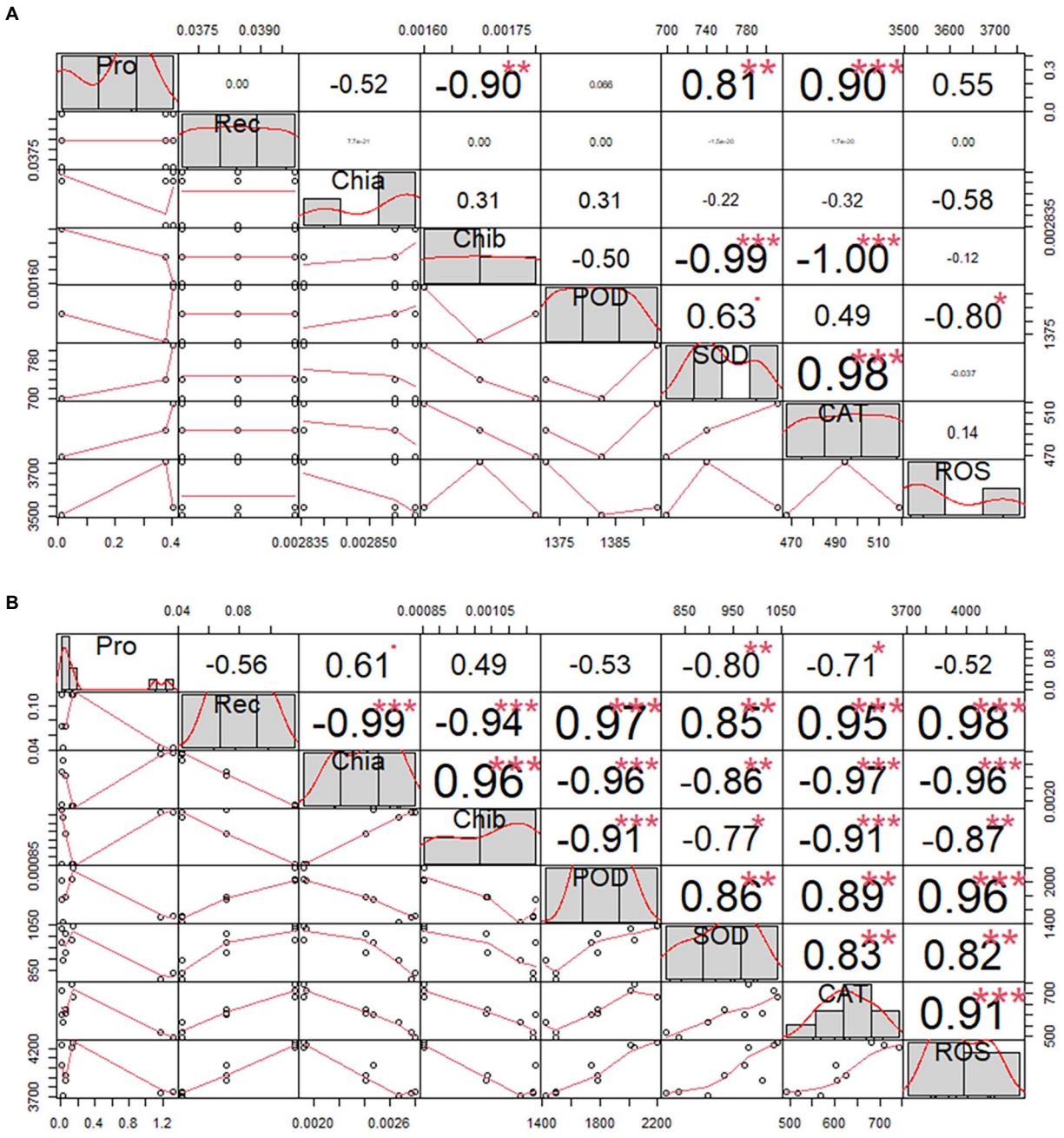


FIGURE 1
Correlation analysis of *Helictotrichon virescens* phenotypic indicators. Morphological data obtained from samples under **(A)** the control temperature (25°C) and **(B)** low-temperature stress (0°C). * represents $p<0.05$,** represents $p<0.01$ and *** represents $p<0.001$.

showed a highly significant positive correlation ($p<0.001$). There was no correlation between the other indicators.

The phenotypic data obtained after 12, 36, and 60 h under low-temperature stress (0°C) are shown in Table 2. Rec, Chia, Chib, POD, CAT, and ROS showed significant differences in treatment time gradient ($p<0.01$), except for Pro and SOD levels. In addition to, pro had the largest coefficient of variation between the treatment time gradients, while ROS exhibited the smallest. Moreover, the kurtosis was the largest (0.8250) for Pro, but the smallest (−1.8650) for ChIb. Pro also had the largest skewness (1.6090), while SOD had the smallest (−0.5600). Furthermore, there was no correlation between Pro and POD, and Pro and ROS, except for Pro and Rec, as shown in Figure 1B.

Table 3 shows that Pro, Chia, Chib, POD, SOD, CAT, and ROS were significantly different between the treated and control groups ($p<0.01$), the results of the above studies indicated that *H. virescens* had less low temperature damage and stronger resistance.

TABLE 2 Statistical description of *H. virescens* phenotypic data at 0°C.

Trait	Range	Mean	SE	CV/%	Kurtosis	Skewness	Sig
Pro	0.0056–1.3260	0.3233	0.1763	0.5451	0.8250	1.6090	ns
Rec	0.0426–0.1155	0.0765	0.0106	0.1386	−1.7140	0.3080	**
Chla	0.0019–0.0028	0.0024	0.0001	0.0510	−1.6560	−0.3000	**
Chlb	0.0009–0.0012	0.0010	0.0000	0.0459	−1.8650	−0.3990	**
POD	1423.07–2195.00	1774.6511	89.9860	0.0507	−1.2050	0.1680	**
SOD	811.64–1043.80	947.4278	27.8088	0.0294	−1.0020	−0.5600	ns
CAT	494.04–743.17	616.8478	28.0549	0.0455	−0.9540	0.0610	**
ROS	3707.94–4269.36	3972.1533	75.2806	0.0190	−1.8280	0.2500	**

ns, not significant.
**$p < 0.01$.

TABLE 3 ANOVA of *H. virescens* phenotypic data (Treatment vs. Control).

	SS	df	MS	F	Sig
Pro	1.2350229	5	0.2470	2.2707	0.0000
Rec	0.0146118	5	0.0029	-	-
Chla	2.098E-06	5	0.0000	386.4508	0.0000
Chlb	2.157E-06	5	0.0000	69.7016	0.0000
POD	1250555.2	5	250111.0380	119.3142	0.0000
SOD	222368.21	5	44473.6421	18.4626	0.0000
CAT	119992.54	5	23998.5080	32.3438	0.0000
ROS	1009380.7	5	201876.1399	19.3729	0.0000

RNA-seq, WGCNA, and hub gene identification

High-throughput illumina sequencing of 18 *H. virescens* samples generated 24,000,000 bp of raw data, from which 22,730,000 bp of high-quality sequence data (clean data) were obtained after assembly and de-redundancy. The clean reads were over 6.5 G per sample, and the base error rate of the sequences from each sample was 0.02–0.03%. Additionally, the Q20 and Q30 of the clean reads were over 98.00 and 94.00%, respectively, and their GC contents were 50.21–55.21%. Trinity splicing generated 396,649 transcript sequences from the clean reads, which were used as reference sequences for subsequent analyses. These transcript sequences were assembled into 112,775 Unigenes after Corset hierarchical clustering (Cheng et al., 2022). The alignment efficiency of the clean reads from the 18 *H. virescens* samples with the reference genome was high (>70%).

The expression modules were divided according to the standard of mixed dynamic shearing, and 38,921 transcripts exhibiting high expression levels were selected for subsequent analyses. The results showed that 27 gene co-expression modules were identified in this study (Figure 2).

The association analysis was conducted using the phenotypic traits based on the previously divided modules. The results showed that the brown, magenta, and tan modules had significant positive correlations ($p < 0.01$) with ROS and the main ROS scavenging enzymes, such as SOD, POD, and CAT. Among these, the brown module had the highest correlation with ROS and the main ROS scavenging enzymes, which were 0.94 ($p = 4e{-}9$), 0.89 ($p = 6e{-}7$), 0.93 ($p = 1e{-}08$), and 0.9 ($p = 3e{-}07$) for SOD, POD, CAT, and ROS, respectively. Furthermore, the turquoise and blue modules exhibited significant positive correlations with the chlorophyll a (Chla) and b (Chlb) contents, and the relative electrical conductivity (Rec) of *H. virescens* leaves. The other modules negatively correlated with the measured plant parameter (Figure 3). Therefore, this study selected three modules (turquoise, blue and brown modules) that exhibited the highest correlation with cold tolerance phenotype for further analysis. Among them, the brown module had 108 hub genes, while the blue and turquoise modules had 96 and 21 hub genes, respectively.

GO enrichment hub genes

The GO enrichment analysis was performed on the hub genes obtained in the brown, blue, and turqouise modules. As shown in Figure 4A, the 108 hub genes of the brown module were mainly enriched in sucrose synthase, translation initiation factor, and biological pathways activities, such as UDP-glucosyltransferase, nitric-oxide synthase, glucosyltransferase, translation factor, structural molecule activities, and structural constituent of ribosome. Conversely, the 96 hub genes of the blue module were mainly enriched in glycine reductase, proline-tRNA ligase, oxidoreductase, and other activities of the biological pathways, such as sodium: proton antiporter, phosphatase, and carbohydrate derivative transporter activities (Figure 4B). Figure 4C shows that the GO enrichment analysis of the 21 hub genes of the turquoise module was significant in glucose-6-phosphate dehydrogenase, peptide-methionine (R)-S-oxide reductase, glutamate synthase, nitronate monooxygenase, and oxidoreductase, and other activities of biological pathways.

As shown in Supplementary Table 1, clusters 37118.66740, 37118.7361, 37118.42165, 37118.46066, 37118.37348, 37118.47362, −37118.48621, 37118.14130, 37118.46066, 37118.14533, and 37118.28125 of the brown module were significantly enriched in sucrose synthase, Translation Initiation Factor,

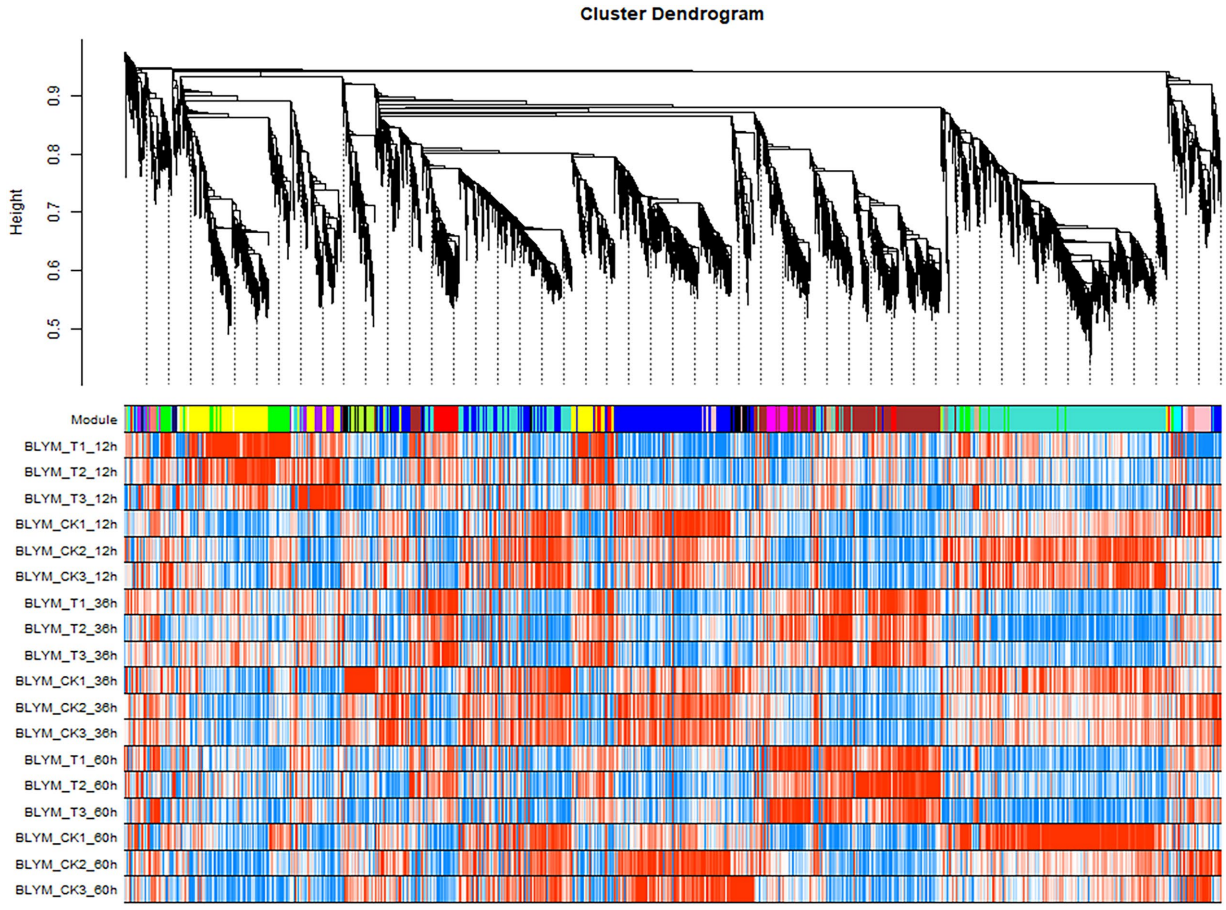


FIGURE 2
Helictotrichon virescens transcript clustering and module identification. The different colors represent different modules, whereby the grey module denotes a collection of differential genes not assigned to other modules. The turquoise module gathered the most genes (2,714 differential genes), while orange, dark-orange, dark-grey, and dark-turquoise modules gathered the least genes (13 genes), among which each module gathered 366 genes on average.

UDP-Glucosyltransferase, UDP-glycosyltransferase, and other activities of the biological pathways. The expression levels of the above 11 hub gene clusters were significantly higher under low-temperature stress than those under normal temperature (Figure 5A), indicating the important roles of these hub genes in the coercion process.

Clusters −7118.46642, 37118.47341, 37118.45930, 37118.46409, 37118.52628, 37118.52427, 37118.42974, and 37118.41399 of the blue module were significantly enriched in glycine reductase, proline-TrNA ligase, oxidoreductase, sodium: proton antiporter, and other activities of the biological pathways (Supplementary Table 2). As shown in Figure 5B, the expression levels of these hub gene clusters were significantly lower under the low-temperature treatment than those under control conditions.

As shown in Supplementary Table 3, the turquoise module had eight clusters (37118.55857, 37118.71123, 37118.46249, 37118.38354, 37118.47615, 37118.46437, 37118.55857, 37118.71123, 37118.46249, 37118.38354, 37118.47615, 37118.46437, 37118.45675, and 37118.45412), which were significantly enriched in army-6-phosphate dehydrogenase and glutamate synthase, nitronate monooxygenase, xidoreductase, and other activities of the biological pathways. The expression levels of these hub gene clusters were significantly lower under the low-temperature stress than those under control conditions (Figure 5C).

Cheng et al. (2022) showed that *H. virescens* has a very high cold tolerance and can successively survive a cold environment of up to −25°C. The 11 hub genes of the brown module were the only ones with significantly higher expression under low-temperature stress, while those from the other two modules had lower expression levels under low-temperature stress than those under control conditions. Therefore, the 11 hub genes of the brown module could be used as key genes responsible for cold tolerance in *H. virescens*, and might need to be studied further.

KEGG enrichment hub genes

The KEGG enrichment analysis was performed for the hub genes in the brown, blue, and turquoise modules. As shown in

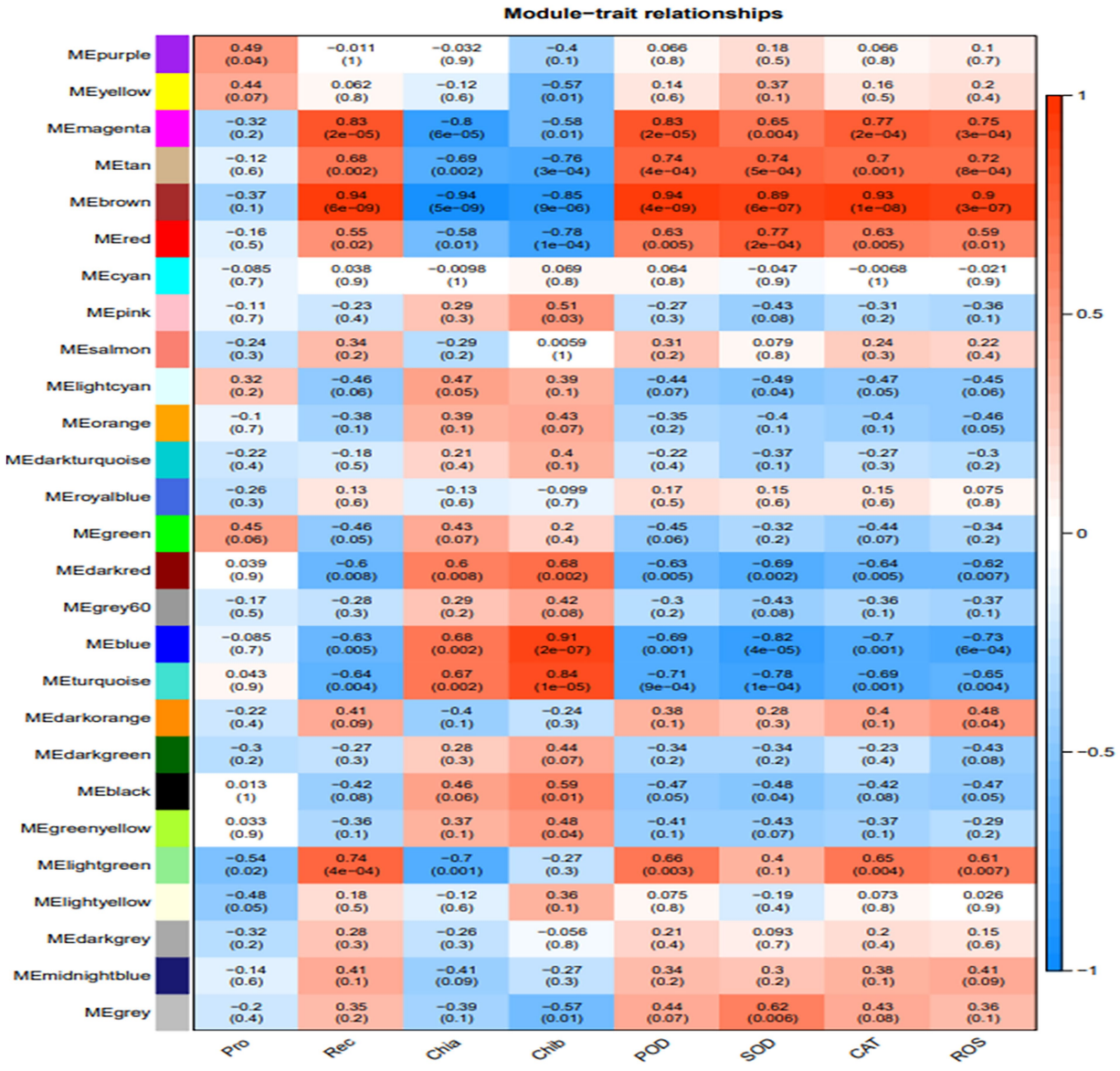


FIGURE 3
Association modules of *Helictotrichon virescens* phenotypic modules. Each row represents a module, while each column represents a trait. The numbers in the rectangular boxes represent the correlation coefficients and corresponding *p* values between the modules and traits.

Figure 6A and Supplementary Table 4, the eight hub genes in the brown module were mainly enriched in sphingolipid, selenocompound, sphingolipid, starch, sucrose, pentose phosphate, fructose, mannose, and galactose metabolism pathways, and other biological pathways. Notably, these biological pathways are all related to glucose metabolism, showing that monosaccharide or polysaccharide metabolism is involved in *H. virescens* response to low-temperature stress.

Contrarily, two out of the 96 hub genes in the blue module were mainly enriched in the endocytosis pathway (Figure 6B and Supplementary Table 5), while four out of the 21 hub genes in the turquoise module were mainly enriched in the carotenoid biosynthesis, and glyoxylate and dicarboxylate metabolism pathways (Figure 6C and Supplementary Table 6).

As shown in Figure 6D, the 14 hub genes enriched by KEGG analysis from the three modules were further subjected to cluster analysis based on their expression levels. The expression levels of clusters 37118.66740 and s37118.47362 were significantly higher under the low-temperature treatment than those under control conditions. However, the expression levels of the remaining hub genes, of which eight belonged to the brown module, were significantly lower under the low-temperature treatment than those under control conditions. Figure 3 shows that the brown module has the highest correlation with the main ROS scavenging

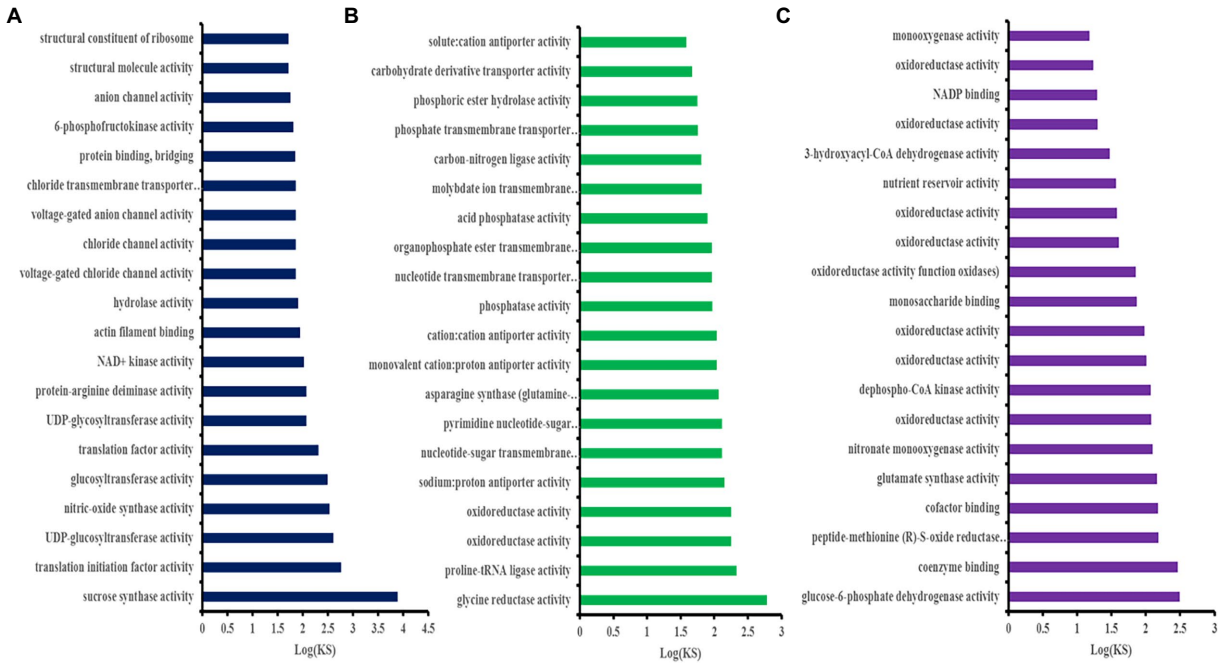


FIGURE 4
Gene ontology (GO) enrichment analysis of the *Helictotrichon virescens* hub genes (MF of the top 20 clusters) from the brown **(A)**, blue **(B)**, and turquoise **(C)** modules.

enzymes, indicating that these enzymes play important roles in *H. virescens* response to low-temperature stress.

Real-time fluorescence quantitative analysis of hub gene

In this study, the RT-PCR analysis of the four hub genes, which exhibited higher expression in the brown module (Figure 7), showed that, the expression of GAPDH was stable at different time points. However, after 12 h of low-temperature treatment, clusters 37118.47362, 37118.47713, and 37118.66740 had no significant difference between the control and treated groups. However, after 36 and 60 h of low-temperature treatment, the clusters exhibited significant differences between the control and treated groups.

Local regulatory network of the hub genes

Cytoscape_v3.9.1 was used to draw a local regulatory network map of cold tolerance-related pathways based on the four key regulatory genes obtained by GO and KEGG enrichment analyses from the brown module (Figure 8). The results showed that clusters 37118.28125 and 37118.47362 are related to cold tolerance and had higher weights and more interactions with multiple genes.

Homologous EST label identification of the hub genes

The comparison analysis of the generated hub gene transcripts shows that the four hub genes are closely related to *Triticum aestivum*, *Hordeum vulgare* subsp. *vulgare*, and the ESTs of *Avena barbata*. Gene Cluster-37118.28125 matched 21 homologous ESTs, among which it had a higher homology with cDNA sequence DK620133.1 obtained from the shoot tips of low-temperature barley subspecies seedlings. Meanwhile, gene cluster-37118.66740 matched 52 homologous ESTs, and had higher homology with DK600269.1 sequence from the cDNA library of the stem segments of low-temperature barley subspecies. The cDNA sequences obtained from callus cloning of the low-temperature barley subspecies also had high homology with gene cluster-37118.66740. The gene cluster-37118.47713 identified five homologous ESTs from oat, while gene cluster-37118.47362 matched 92 homologous ESTs, of which 11 were homologous sequences related to abiotic stress. The ESTs were deposited in the GenBank database (NCBI), and their sequences are shown in Supplementary Table 7.

Bioinformatics analysis of hub genes

The physicochemical properties of the protein encoded by the hub genes were analyzed. The results showed that the number of amino acids in Cluster-37118.47362 was relatively small, and the

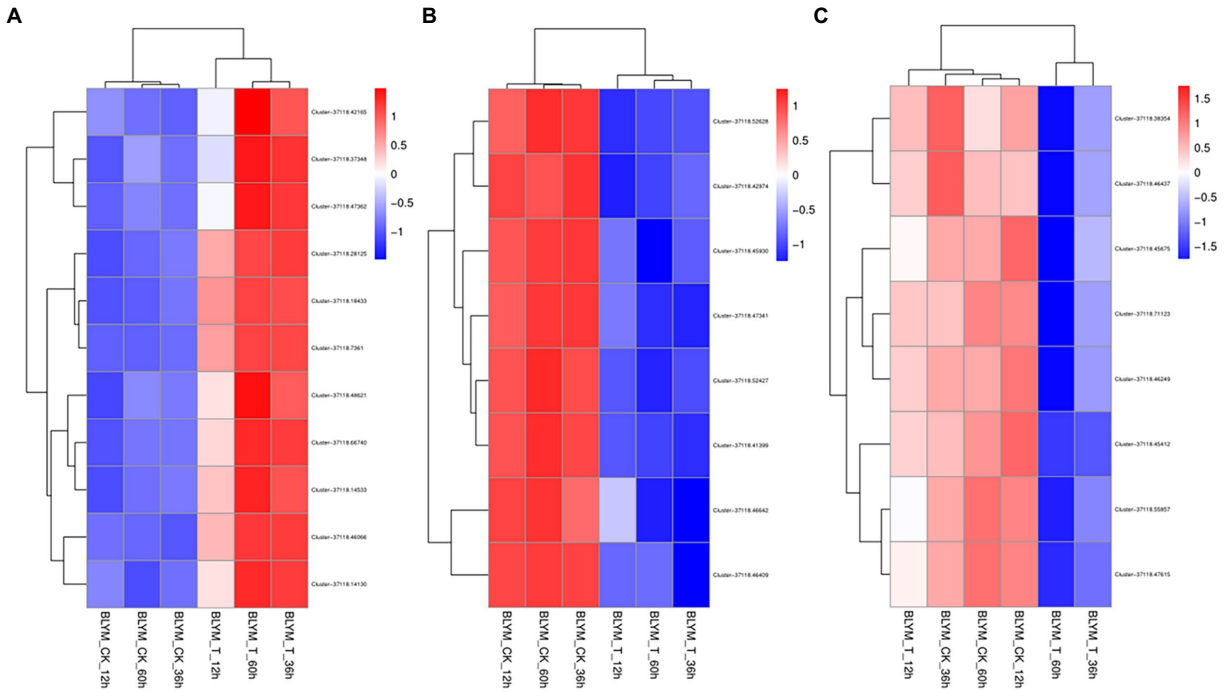


FIGURE 5
Cluster analysis of the *Helictotrichon virescens* hub genes from the brown **(A)**, blue **(B)**, and turquoise **(C)** modules.

number of amino acids in Cluster-37118.66740 was relatively large. In addition, the protein isoelectric points of the four hubs were all greater than 5, and they were all hydrophobic proteins (Table 4). The subcellular localization prediction results showed that, the hub genes were all localized to the plasma membrane.

Construction of protein interaction network of hub genes

According to the protein function annotation information of the *Sorghum bicolor* genome by STRING software, it was found that Cluster-37118.28125, Cluster-37118.47362, and Cluster-37118.47713 could not be compared to monocotyledonous crops. However, the predicted results of Cluster-37118.66740 protein were relatively accurate and the protein was highly similar to Sb01g035890.1, a key regulatory protein in the sucrose metabolism pathway in *Sorghum bicolor*. By constructing a protein interaction network, it was found that the hub gene Cluster-37118.66740 interacted with Sb09g003460.1 and Sb04g020180.1 proteins in *Sorghum bicolor* (Figure 9A).

Identification and phylogenetic analysis of homologous genes of hub genes

A total of five sucrose synthase activity (SUS) family members were identified in *Sorghum bicolor*, 6 in *Oryza sativa Japonica*, and six *in Arabidopsis thaliana* using the protein sequence of the cold hub gene Cluster-37118.66740 as a probe (Supplementary Table 8). By constructing phylogenetic trees of the four species of *H. virescens*, *Sorghum bicolo*, *Oryza sativa Japonica*, and *Arabidopsis thaliana*, the results showed that, the hub gene Cluster 37118.66740 (of *H. virescens*) and Os03g0340500 (of *Oryza sativa Japonica*) belonged to the same ancestral branch and were in the same subfamily (Figure 9B). In addition, monocotyledonous and dicotyledonous plants form different branches, and *Oryza sativa Japonica*, *Sorghum bicolo* and *H. virescens* are more closely related. The results of this study also showed that the clustering distribution of species in the subfamily was obvious, indicating that the differentiation time of the sucrose synthase family genes was earlier.

Discussion

Relationship between physiological indices and response to low-temperature of *Helictotrichon virescens*

Abiotic stresses affect plant growth and development at different levels. Excessive production and accumulation of ROS (such as OH^-, H_2O_2, and O^{2-}) in plant cells leads to the damage of macromolecules, including proteins, RNA, and DNA. Superoxide radicals (O^{2-}) are disproportionated by SOD to H_2O_2 and further scavenged by CAT and peroxides such as POD by conversion to

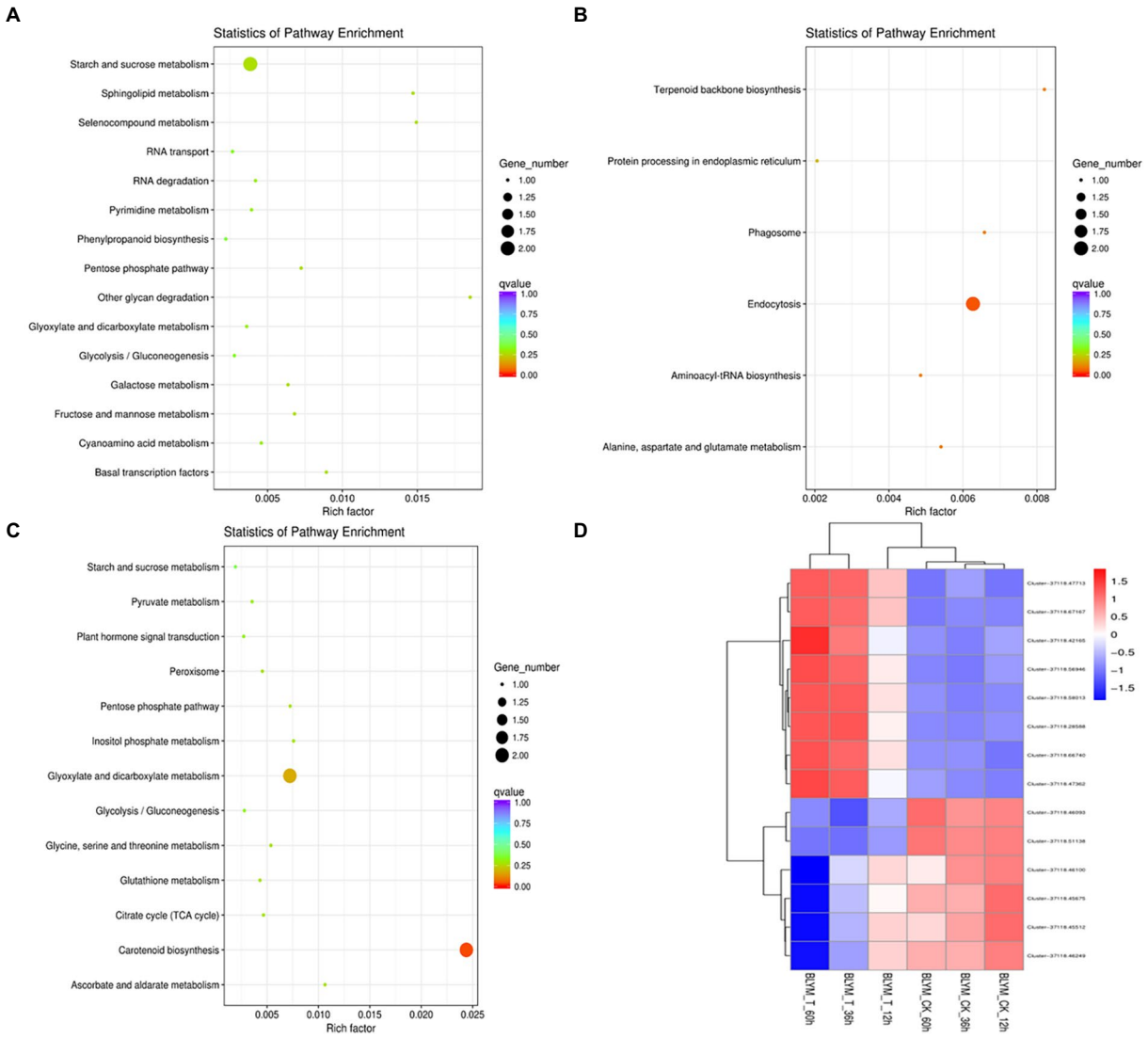


FIGURE 6
KEGG enrichment analysis and clustering analysis of *Helictotrichon virescens* hub genes from the brown **(A)**, blue **(B)** represents, and turquoise modules **(C)**. Cluster analysis representing the various hub genes **(D)**.

H_2O. Freezing treatment for 3 h significantly increased the activities of CAT, POD, and SOD. These results are consistent with the findings of Bermudagrass, which showed that treatment at 4°C over6 days increased the activities of CAT, POD, and SOD (Cheng et al., 2016). However, 4°C over 72 h low-temperature treatment increased SOD activity but decreased CAT and POD activities (Wei et al., 2015). These differences may be attributed to detailed changes in temperature, treatment time, and grass species. All data suggest that low-temperature stress modulates ROS by modulating antioxidant enzyme activities. Cheng et al. (2022) showed that with the prolongation of stress time, the ROS content in *H. virescens* leaves gradually increased, and the difference between the treatment and control groups reached significant or highly significant levels; at the same time, POD, SOD, and CAT were significantly enhanced, which is conducive to the scavenging of ROS in the *H. virescens* plants and other plants. In addition, due to the diversity of plant cold tolerance indicators and physiological adaptation mechanisms, the analysis of a single index often cannot truly reflect the cold tolerance of plants. Therefore, a comprehensive analysis of the related indicators of its cold tolerance can comprehensively evaluate and compare the cold resistance of plants. However, there is no report on the comprehensive evaluation of the physiological response and cold tolerance of forage varieties to low-temperature stress. He et al. (2021) used the membership function method to calculate the fresh weight, dry weight, chlorophyll content, SOD activity, MDA content, free proline content, and soluble sugar content of six kinds of grasses (*Elymus sibiricus'Qingmu No.1'*, *E.sibiricus'Tongde'*, *E.breviaristatus'Tongde'*, *Poa crymophila'Qinghai'*, *P.pratensis* var.*anceps Gaud'Qinghai'*, and *Festuca sinensis'Qinghai'*) under

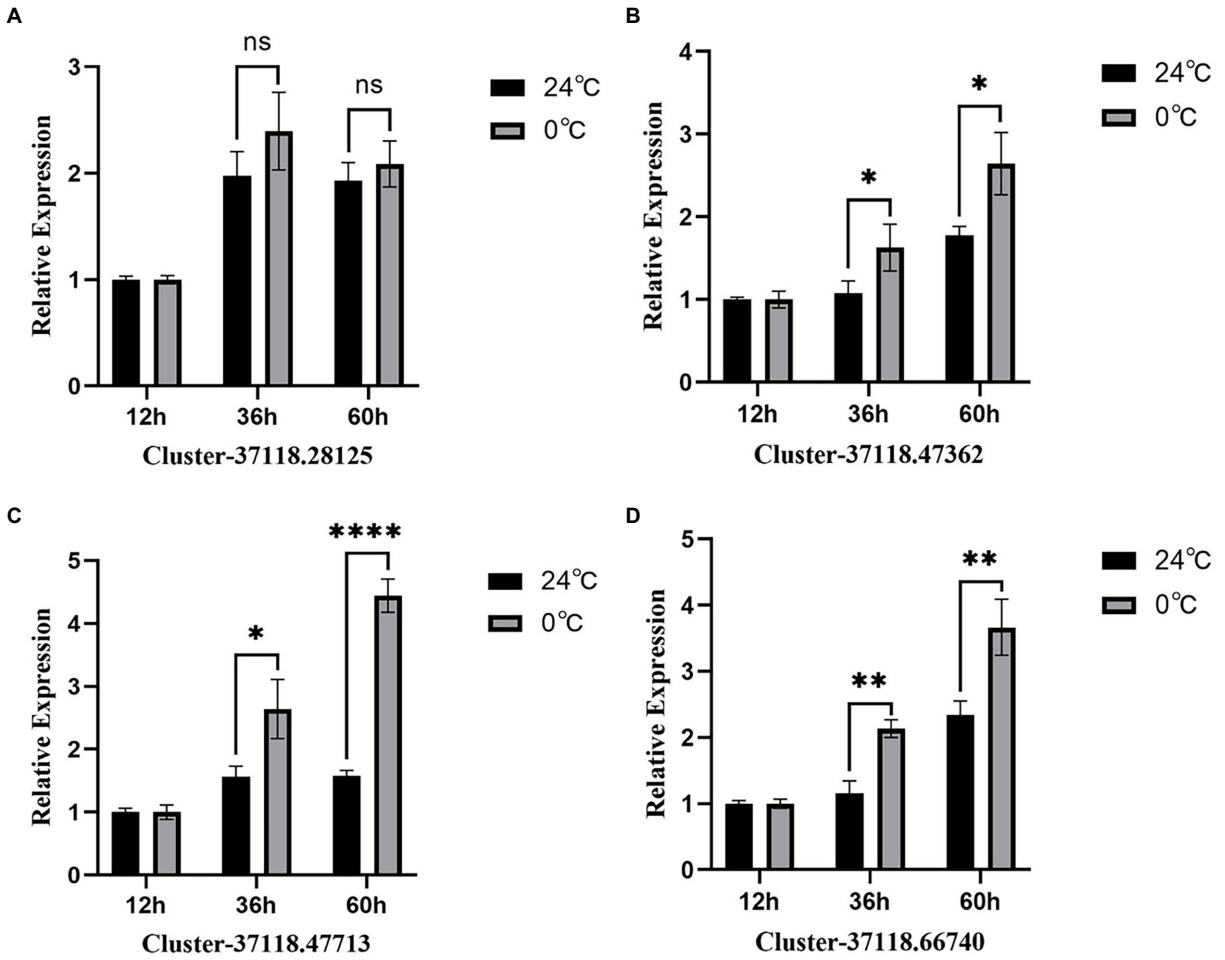


FIGURE 7
RT-PCR analysis of the four hub genes from the brown module. **(A–D)** are the expression results of four hub genes, respectively. * represents $p<0.05$,** represents $p<0.01$,*** represents p<0.001, **** represents $p<0.0001$.

low-temperature stress. According to the membership function value of the six species of grasses, the cold resistance of the six types of grass can be comprehensively evaluated so that the evaluation results can more comprehensively reflect the actual cold resistance of the six grass species. Cheng et al. (2022) showed that the chlorophyll content in *H. virescens* leaves showed a decreasing trend with the extension of the stress time. After 12 h of low temperature stress, the content of chlorophyll a was not significantly different between the treated group and the control group, and the content of chlorophyll b was significantly reduced; after 36 and 60 h of low temperature stress, the content of chlorophyll a and b in the treated group were significantly lower than those in the control group; In addition, the relative conductivity of leaves increased after low temperature stress. After 12 h of low temperature stress, the relative conductivity of leaves in the treated group was 1.15 times that of the control group. After prolonged low temperature stress to 60 h, the relative conductivity of leaves in the treated group increased. to 2.90 times that of the control group, indicating that the stability and tolerance of leaf cell membranes were damaged during low temperature stress; Under low temperature stress, the accumulation of proline (Pro) can reduce the cytosolic freezing point and increase the osmotic potential, thereby stabilizing the cell membrane system to prevent cell freezing and dehydration and reduce the exudation of solutes. The Pro content in the treated group and the control group showed a very significant difference, and the Pro content in the treatment group was 2.16 times that of the control group. In this study, as shown in Table 3, the method analysis of the phenotypic indicators of *H. virescens* under low-temperature stress and control treatments was carried out, and the results showed that Pro, Chia, Chib, POD, SOD, CAT, and ROS were significantly different between the treated and the control groups ($p<0.01$), the results of the above studies indicated that *H. virescens* had less low temperature damage and stronger resistance. As shown in Figure 1B, except for Pro and Rec, there was no correlation between Pro and POD, and Pro and ROS, and there were significant ($p<0.05$) andextremely significant ($p<0.01$) between other indicators. The above results indicate that these indicators can be used as phenotypic indicators for the identification of cold stress in *H. virescens*.

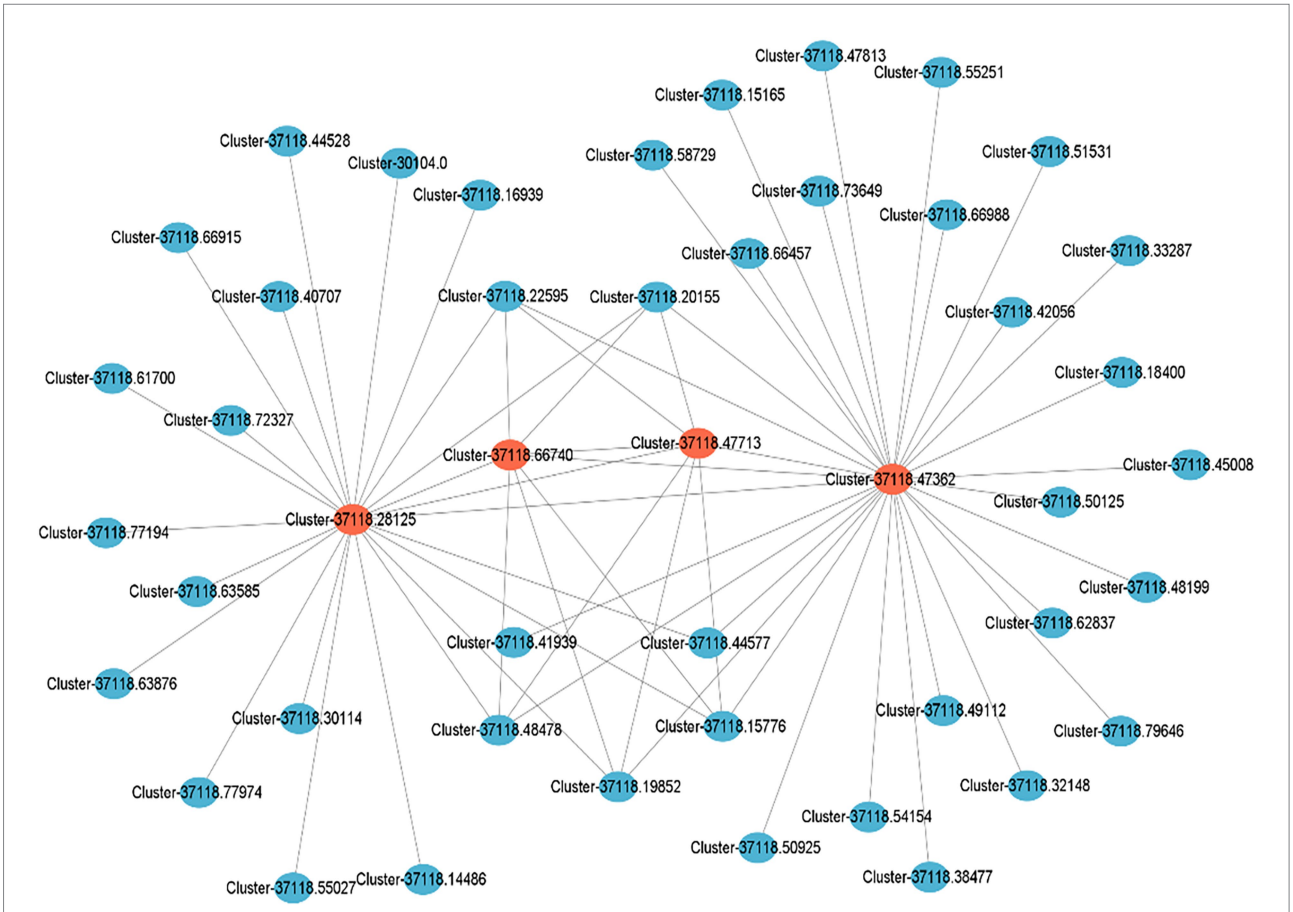


FIGURE 8
Local regulatory network of cold tolerance-related genes obtained from the brown module. Red circles represent the key cold tolerance-related genes, while blue ones represent the differentially expressed genes interacting with the key genes.

TABLE 4 Physical and chemical properties and subcellular localization prediction of hub gene protein related to cold tolerance in *H. virescens*.

Transcript ID	Amino acid amounts(aa)	Molecular weight (KD)	pl	Hydrophilicity coefficient	Transmembrane domain amounts	Subcellulat localization
Cluster-37118.28125	372	42.06	6.4	−0.381	3	Plasma membrane
Cluster-37118.47362	80	39	9.89	−0.035	1	Chloroplast
Cluster-37118.66740	744	36.65	5.93	−0.243	2	Plasma membrane
Cluster-37118.47713	306	27.61	9.06	−0.412	1	Mitochondrion

WGCNA is a key strategy to excavate cold tolerance genes in *Helictotrichon virescens*

Weighted Gene Co-expression Network Analysis is a research method in systems biology with many applications, including mining gene modules related to target traits in multi-sample transcriptome data. Unlike other network analysis methods, WGCNA can specifically identify genes associated with the target traits, screen core genes, and perform modular classification to obtain co-expression modules with high biological significance. Therefore, choosing a quantifiable target trait is a primary consideration for WGCNA analysis.

Qin et al. (2020) used the seedlings of cultivars C16 (CIP 397077.16) and C119 (CIP 398098.119) from the International Potato Center as test materials for the mannitol-induced drought stress experiment. The study obtained 15 gene co-expression modules closely related to root drought resistance, and several core genes with the highest correlation with the target trait were identified from four modules. The functional annotation showed that most of these genes were involved in the drought stress regulation pathway. Moreover, Li et al. (2019) used 47 transcriptome data sets of normal rice tissue for cold, drought, and

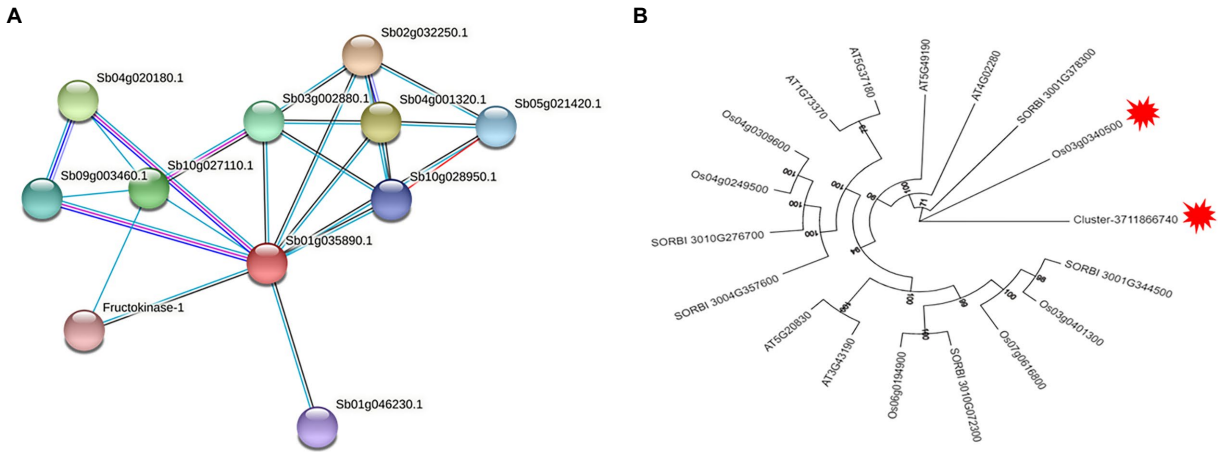


FIGURE 9
(A) Sb01g035890.1 protein interaction network. The red node in the figure is the sorghum protein Sb01g035890.1 aligned by Cluster-37118.66740, the red line represents gene fusion, the cyan line is from the database, the pink line is the experimental proof result, the black line is the co-expressed protein, and the blue line represents the gene symbiotic relationship; **(B)** Phylogenetic Analysis of sucrose synthase gene family.

salt stress treatment. The study identified 15 modules using WGCNA and found that the three known rice-related genes were present in each module. Two modules related to the three stress treatments were selected to construct the gene regulatory network in which 25 key genes related to stress resistance were predicted. Wang et al. (2021) used different peanut varieties as materials to analyze the similarities and differences in the expression of the stem-growth-related genes. The results showed that the dwarf type Df216 had 5,872 differential genes while tall type Huayu had 33. Df216 and the intermediate type Shanhua 108 had 6,662 differential genes, which were implicated in the biological origin and regulatory processes of the primary and secondary cell walls biosynthesis, phenylpropane biosynthesis and metabolism, lignin biosynthesis, cellulose synthase activity, and other molecular functions. Additionally, WGCNA identified five co-expression modules significantly correlated with the main stem height, whose core genes encoded caffeoyl-CoA-O-methyltransferase, transcription factors ATAF2, WAT1, and GDSL lipase. In the present study, three modules (turquoise, blue and brown modules) had the highest correlation with cold tolerance phenotype and were selected for further analysis. The brown module had 108 hub genes, while the blue and turquoise models had 96 and 21 hub genes, respectively. Only 11 hub genes in the brown module had significantly higher expression levels under low-temperature stress, while those in the other two modules were significantly lower than those under control conditions. Through GO enrichment analysis, the results showed that hub genes were significantly enriched in biological pathways, such as sucrose synthase activity, translation initiation factor activity, UDP-glucosyltransferase activity, and UDP-glycosyltransferase activity, among which, the above biological pathways were all related to sucrose synthesis Enzymes and glycosyltransferases are related, thus indicating that the enzyme activity is involved in the low temperature response process and plays a crucial role. Through KEGG enrichment analysis, the results show that hub genes are mainly enriched in biological pathways such as Sphingolipid metabolism, Selenocompound metabolism, Sphingolipid metabolism, Starch and sucrose metabolism, Pentose phosphate pathway, Fructose and mannose metabolism, Galactose metabolism, Glycolysis, and Gluconeogenesis. Among them, the above biological pathways are all related to sugar metabolism, which indicates that monosaccharide or polysaccharide metabolism is involved in the process of *H. virescens* in response to low temperature stress. Therefore, the 11 hub genes screened in the brown module could be the key cold-tolerance genes of *H. virescens*, and might need further studies.

The antioxidant defense system plays an important role in response to low-temperature stress in *Helictotrichon virescens*

Plants have complex mechanisms to respond to abiotic stresses (such as drought, salinity, and extreme temperatures). The stress signals are sensed by receptors on the cell membrane and then transduced to second messengers, a phenomenon that induces the activation of downstream physiological responses. The expression of different stress-responsive genes ultimately leads to a protective response in the whole plant (Varshney et al., 2011; Lei et al., 2014). Transcriptome profiling is an important strategy for elucidating the molecular components of cells and tissues for explaining functional elements of the genomes in response to different stimuli (Qiu et al., 2013). High-throughput RNA sequencing (RNA-seq) has been used for gene discovery and regulatory network studies, including stress response studies in higher plants (Deyholos, 2010; Kakumanu et al., 2012; Wang et al., 2013). In recent years, some studies have revealed the molecular

mechanism of abiotic stress by RNA-seq technology in turfgrass. For example, comparative transcriptome analysis of tall fescue and ryegrass exposed to heat or cold stress for 10 h showed that approximately 30 and 25% of genes showed significant changes under heat and cold stress, respectively, and of them, *HSFs* genes exhibited strong responses to heat and cold stresses (Wang et al., 2015). Transcriptome data of cold-acclimated and non-acclimated bermudagrass were studied, and the results showed that 5,867 genes were differentially expressed in cold-acclimated and unaccustomed bermudagrass, of which 2,181 were downregulated and 710 were up-regulated; among up-regulated genes, *AP2*, *NAC,* and *WRKY* family members were associated with cold stress (Zhu et al., 2015). However, in comparison to other grasses, little research has been done on the molecular mechanism of *H. virescens* response to low-temperature stress. It is well known that low-temperature stress can change the permeability of mesophyll cells in crop seedlings, and the change in permeability is inversely proportional to the external temperature and proportional to the action time; therefore, the low-temperature stress seriously causes cell membrane damage and greatly changes cell membrane permeability (Fridovich, 1978) collectively referred to the three enzymes SOD, CAT, and POD as cytoprotective enzyme systems, and their activity trends could reflect the cold tolerance of crops. In this study, *H. virescens* seedlings were subjected to low temperature stress at 0°C, the results showed that after 12, 36, and 60 h of stress, the leaves of *H. virescens* seedlings suffered less low temperature damage and the plants could grow normally. Subsequently, this study conducted a method comparison and analysis of the phenotypic indicators of *H. virescens* under treated group and control group, the results showed that POD, SOD, CAT, and ROS were significantly different between the treatment and the control groups ($p < 0.01$; Table 3). The above results show that under low-temperature stress, the high activity of antioxidant enzymes is beneficial to the removal of ROS in *H. virescens* plants and reduces the damage of ROS in the plants.

In order to elucidate the mechanism of response to cold stress of *H. virescens*, previous research results show that: at low -temperature, ROS content accumulation was the same as POD, SOD and CAT activity, the gene that regulates POD activity (Cluster37118.16911) and the gene that regulates ROS (Cluster37118.62042) may be involved in ROS elimination during low-temperature treatment (Cheng et al., 2022). However, in this study, we used weighted gene co-expression network analysis (WGCNA) to identify the hub genes associated with cold tolerance in *H. virescens*. Specifically, the correlation coefficients between the brow module and superoxide dismutase, peroxisome, catalase, and ROS were 0.94 ($p < 0.001$), 0.89 ($p < 0.001$), 0.93 ($p < 0.001$), and 0.9 (<0.001; Figure 2). Therefore, GO annotation results show that, as shown in Figure 4A, the hub genes of cluster-37118.47362, cluster-37118.47713, and cluster-37118.66740 were mainly enriched in sucrose synthase, translation initiation factor, and biological pathways activities, such as UDP-glucosyltransferase, nitric-oxide synthase, glucosyltransferase, translation factor, structural molecule activities, and structural constituent of ribosome. KEGG annotation results show that, as shown in Figure 6A and Supplementary Table 4, the cluster-37118.47362, cluster-37118.47713, and cluster-37118.66740 were mainly enriched in sphingolipid, selenocompound, sphingolipid, starch, sucrose, pentose phosphate, fructose, mannose, and galactose metabolism pathways, and other biological pathways. Notably, these biological pathways are all related to glucose metabolism, showing that monosaccharide or polysaccharide metabolism is involved in *H. virescens* response to low-temperature stress. Subsequently, RT-PCR analysis of the hub genes that were highly expressed in the brown module showed that the expression levels of cluster-37118.47362, cluster-37118.47713, and cluster-37118.66740 were significantly higher than those in the control conditions at 36 h and 60 h of low-temperature treatments (Figure 7). Compared with the results of Cheng et al. (2022), in this study, the phenotypic data of the analysis results of POD, SOD, CAT, and ROS phenotype indicators were combined with the FPKM value of differential genes to conduct WGCNA analysis, and the key hub genes were mined. This analysis method is more reasonable and more purposeful, and at the same time, according to the analysis results of hub gene expression, GO annotation results, KEGG annotation results, it can showed that cluster-37118.47362, cluster-37118.47713, and cluster-37118.66740 play an important role in the response to low-temperature stress in the *H. virescens*. On the basis of the above research, bioinformatics analysis of hub gene showed that, the protein isoelectric points of the hubs were hydrophobic proteins, and the hub genes were all localized to the plasma membrane. By constructing a protein interaction network, it was found that the hub gene Cluster-37118.66740 interacted with Sb09g003460.1 and Sb04g020180.1 proteins in *Sorghum bicolor.* Homologous gene identification results showed that, a total of five sucrose synthase activity (SUS) family members were identified in *Sorghum bicolor*, six in *Oryza sativa Japonica*, and six *in Arabidopsis thaliana*. In addition, KEGG enrichment analysis were performed on the above homologous genes (Supplementary Tables 9–11), the results showed that, all homologous genes were significantly enriched in metabolic pathways and starch and sucrose metabolism pathways, and studies have shown that the above biological pathways are related to abiotic stress. It is worth noting that, by constructing phylogenetic trees of the four species of *H. virescens*, *Sorghum bicolo*, *Oryza sativa Japonica*, and *Arabidopsis thaliana*, the results showed that the gene Cluster 37118.66740 and Os03g0340500 belonged to the same ancestral branch and were in the same subfamily (Figure 9B), subsequently, we performed functional annotation of Os03g0340500, the functional annotation as Sucrose-UDP glucosyltransferase 4. In plants, UDP-glycosyltransferase catalyzes the glycosyl transfer reaction, which transfers the sugar group from the activated donor molecule to the acceptor molecule, thereby regulating the activity of the acceptor

molecule in cells and organisms, such as biological activity, dissolution sex, and transport. In conclusion, the hub gene Cluster 37118.66740, as a key candidate gene, should be further analyzed in the follow-up studies.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://www.ncbi.nlm.nih.gov/search/all/?term=PRJNA810780.

Author contributions

MC, HW, RZ, XLei, MI, TY, XLi, YZ, YL, and RH designed the experiment. ZP and KC collated the transcriptome data. JZ and YC conducted the RT-PCR experiment. MC, XLei, TY, XLi, YZ, and YL drafted the manuscript. RH revised the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the Fundamental Research Funds for the Central Universities (ZYN2022053), Science and Technology Project of Sichuan Province (2020YJ0466), the Forage Innovation Team of Sichuan Province, and the Public Relations Project of Forage Breeding of Sichuan Province.

Acknowledgments

The authors are grateful to the laboratory members for their assistance and advice. We would also like to thank the editors and referees for their valuable comments on improving the manuscript. Special thanks to Fuyuan Liu, Bob Yang, Guangdi Li, and Neil Griffiths for their guidance.

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.2022.938859/full#supplementary-material

References

Abeynayake, S. W., Byrne, S., Nagy, I., Jonavičienė, K., Etzerodt, T. P., Boelt, B., et al. (2015). Changes in *Lolium perenne* transcriptome during cold acclimation in two genotypes adapted to different climatic conditions. *BMC Plant Biol* 15:250. doi: 10.1186/s12870-015-0643-x

Altschul, S. F. (2012). Basic local alignment search tool (BLAST). *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Augustyniak, A., Perlikowski, D., Rapacz, M., Kościelniak, J., and Kosmala, A. (2018). Insight into cellular proteome of *Lolium multiflorum/Festuca arundinacea* introgression forms to decipher crucial mechanisms of cold acclimation in forage grasses. *Plant Sci.* 272, 22–31. doi: 10.1016/j.plantsci.2018.04.002

Bocian, A., Kosmala, A., Rapacz, M., Jurczyk, B., and Zwierzykowski, Z. (2011). Differences in leaf proteome response to cold acclimation between *Lolium perenne* plants with distinct levels of frost tolerance. *J. Plant Physiol.* 168, 1271–1279. doi: 10.1016/j.jplph.2011.01.029

Bocian, A., Zwierzykowski, Z., Rapacz, M., Koczyk, G., and Kosmala, A. (2015). Metabolite profiling during cold acclimation of *Lolium perenne* genotypes distinct in the level of frost tolerance. *J. Appl. Genet.* 56, 439–449. doi: 10.1007/s13353-015-0293-6

Cheng, M. J., Cui, K. S., Zheng, M., Yang, T., Zheng, J., Li, X., et al. (2022). Physiological attributes and transcriptomics analyses reveal the mechanism response of *Helictotrichon virescens* to low temperature stress. *BMC Genomics* 23:280. doi: 10.1186/s12864-022-08526-4

Cheng, Z., Jin, R., Cao, M., Liu, X., and Chan, Z. (2016). Exogenous application of ABA mimic 1 (AM1) improves cold stress tolerance in bermudagrass (*Cynodon dactylon*). *Plant Cell Tissue Organ Cult.* 125, 231–240. doi: 10.1007/s11240-016-0941-5

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

Deyholos, M. K. (2010). Making the most of drought and salinity transcriptomics. *Plant Cell Environ.* 33, 648–654. doi: 10.1111/j.1365-3040.2009.02092.x

Fridovich, I. (1978). The biology of oxygen radicals. *Science*. 201, 875–880. doi: 10.1126/science.210504

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883

He, Z. H., Yang, C. Y., Wang, P., Bao, A. K., and Ma, Q. (2021). Physiological response and cold tolerance evaluation of six *Gramineae herbage* species to low temperature stress in alpine region. *Pratacult. Sci.* 38, 2019–2028. doi: 10.11829/j.issn.1001-0629.2021-0083

Hoffman, L., Dacosta, M., Ebdon, J. S., and Watkins, E. (2010). Physiological changes during cold acclimation of perennial ryegrass accessions differing in freeze tolerance. *Crop Sci.* 50, 1037–1047. doi: 10.2135/cropsci2009.06.0293

Hulke, B. S., Bushman, B. S., Watkins, E., and Ehlke, N. J. (2012). Association of freezing tolerance to *LpCBFIIIb* and *LpCBFIIIc* gene polymorphism in perennial ryegrass accessions. *Crop Sci.* 52, 2023–2029. doi: 10.2135/cropsci2011.09.0527

Kakumanu, A., Ambavaram, M. M. R., Klumas, C., Krishnan, A., Batlang, U., Myers, E., et al. (2012). Effects of drought on gene expression in maize reproductive and leaf meristem tissue revealed by RNA-Seq. *Plant Physiol.* 160, 846–867. doi: 10.1104/pp.112.200444

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, D480–D484. doi: 10.1093/nar/gkm882

Lei, X. T., Xiao, Y., Xia, W., Mason, A. S., Yang, Y. D., Ma, Z., et al. (2014). RNA-Seq analysis of oil palm under cold stress reveals a different C-repeat binding factor

(CBF) mediated gene expression pattern in *Elaeis guineensis* compared to other species. *PLoS One* 9:e114482. doi: 10.1371/journal.pone.0114482

Levitt, J. (1980). Responses of plants to environmental stresses. *Chill. Freez. High Temp. Stress* 38, 480–3645. doi: 10.2307/3899731

Li, X., Cheng, X., Liu, J., Zeng, H., Han, L., and Tang, W. (2011). Heterologous expression of the *Arabidopsis DREB1A/CBF3* gene enhances drought and freezing tolerance in transgenic *Lolium perenne* plants. *Plant Biotechnol. Rep.* 5, 61–69. doi: 10.1007/s11816-010-0157-9

Li, X. K., Li, R. J., and Zhang, B. J. (2019). Identification of abiotic stress-related gene co-expression networks using WGCNA. *Acta Agron. Sin.* 45, 1349–1364. doi: 10.3724/SP.J.1006.2019.82061

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

Paina, C., Byrne, S. L., Domnisoru, C., and Asp, T. (2014). Vernalization mediated changes in the *Lolium perenne* transcriptome. *PLoS One* 9:e107365. doi: 10.1371/journal.pone.0107365

Qin, T. Y., Sun, C., Bi, Z. Z., Liang, W. J., Li, P. C., Zhang, J.-L., et al. (2020). Identification of coexpression modules and core genes related to drought resistance in potato root based on WGCNA. *Acta Agron. Sin.* 46, 1033–1051. doi: 10.3724/SP.J.1006.2020.94130

Qiu, Z. B., Wan, L. C., Chen, T., Wan, Y. L., He, X. Q., Lu, S., et al. (2013). The regulation of cambial activity in Chinese fir (*Cunninghamia lanceolata*) involves extensive transcriptome remodeling. *New Phytol.* 199, 708–719. doi: 10.1111/nph.12301

Sigaud-Kutner, T., Pinto, E., Okamoto, O. K., Latorre, L. R., and Colepicolo, P. (2010). Changes in superoxide dismutase activity and photosynthetic pigment content during growth of marine phytoplankters in batch-cultures. *Physiol. Plant.* 114, 566–571. doi: 10.1034/j.1399-3054.2002.1140409.x

Tao, X., Wang, M. X., Dai, Y., Wang, Y., Fan, Y. F., Mao, P., et al. (2017). Identification and expression profile of CYPome in perennial ryegrass and tall fescue in response to temperature stress. *Front. Plant Sci* 8:1519. doi: 10.3389/fpls.2017.01519

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., et al. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28, 511–515. doi: 10.1038/nbt.1621

Varshney, R. K., Bansal, K. C., Aggarwal, P. K., Datta, S. K., and Craufurd, P. Q. (2011). Agricultural biotechnology for crop improvement in a variable climate: hope or hype? *Trends Plant Sci.* 16, 363–371. doi: 10.1016/j.tplants.2011.03.004

Wang, Y., Dai, Y., Tao, X., Wang, J. Z., Cheng, H. Y., Yang, H., et al. (2015). Heat shock factor genes of tall fescue and perennial ryegrass in response to temperature stress by RNA-Seq analysis. *Front. Plant Sci* 6:1226. doi: 10.3389/fpls.2015.01226

Wang, Y., Gao, F., Liu, Z. X., Zhao, J. H., Lai, H. J., Pan, X.-Y., et al. (2021). Using WGCNA module identified main stem growth of peanut gene expression. *J. Crop.* 47, 1639–1653. doi: 10.3724/SP.J.1006.2021.04223

Wang, X. C., Zhao, Q. Y., Ma, C. L., Zhang, Z. H., and Yang, Y. J. (2013). Global transcriptome profiles of *Camellia sinensis* during cold acclimation. *BMC Genomics* 14:415. doi: 10.1186/1471-2164-14-415

Wei, S. J., Du, Z. L., Gao, F., Ke, X., Li, J., Liu, J., et al. (2015). Global Transcriptome profiles of 'Meyer' zoysiagrass in response to cold stress. *PLoS One* 10:e0131153. doi: 10.1371/journal.pone.0131153

Young, M. D., Wakefield, M. J., Smyth, G. K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 11:R14. doi: 10.1186/gb-2010-11-2-r14

Yu, X. Q., Pijut, P. M., Byrne, S., Asp, T., Bai, G., Jiang, Y., et al. (2015). Candidate gene association mapping for winter survival and spring regrowth in perennial ryegrass. *Plant Sci.* 235, 37–45. doi: 10.1016/j.plantsci.2015.03.003

Zhao, C. X., Ma, X., Dong, W. K., Ma, K. J., Zhang, R., Huiling, M., et al. (2020). Transcriptome analysis of Poa Pratensis under low temperature stress. *J. Grassland* 28, 305–318. doi: 10.11733/j.issn.1007-0435.2020.02.003

Zhu, H., Yu, X., Xu, T., Wang, T., Du, L., Ren, G., et al. (2015). Transcriptome profiling of cold acclimation in bermudagrass (*Cynodon dactylon*). *Sci. Hortic.* 194, 230–236. doi: 10.1016/j.scienta.2015.08.024

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 08 September 2022
DOI 10.3389/fpls.2022.992024

OPEN ACCESS

EDITED BY
Linkai Huang,
Sichuan Agricultural University, China

REVIEWED BY
Guohui Yu,
Nanjing Agricultural University, China
Zhiqiang Zhang,
Inner Mongolia Agricultural University, China

*CORRESPONDENCE
Pan Zhang
zhangpan@neau.edu.cn
Guowen Cui
cgw603@163.com

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

RECEIVED 12 July 2022
ACCEPTED 08 August 2022
PUBLISHED 08 September 2022

CITATION
Yin H, Wang Z, Li H, Zhang Y, Yang M, Cui G and Zhang P (2022) *MsTHI1* overexpression improves drought tolerance in transgenic alfalfa (*Medicago sativa* L.). *Front. Plant Sci.* 13:992024. doi: 10.3389/fpls.2022.992024

MsTHI1 overexpression improves drought tolerance in transgenic alfalfa (*Medicago sativa* L.)

Hang Yin, Zhaoyu Wang, Han Li, Yu Zhang, Mei Yang, Guowen Cui* and Pan Zhang*

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

In recent years, drought stress caused by global warming has become a major constraint on agriculture. The thiamine thiazole synthase (THI1) is responsible for controlling thiamine production in plants displaying a response to various abiotic stresses. Nonetheless, most of the THI1 activities in plants remain largely unknown. In this study, we extracted *MsTHI1* from alfalfa and demonstrated its beneficial impact on improving the resistance of plants to stress conditions. The highest levels of *MsTHI1* expression were identified in alfalfa leaves, triggered by exposure to cold, drought, salt, or alkaline conditions. The upregulation of *MsTHI1* in drought-stressed transgenic plants resulted in enhanced accumulation of vitamin B1 (VB1), chlorophyll *a* (Chl *a*), chlorophyll *b* (Chl *b*), soluble protein, higher soil and plant analyzer development (SPAD) value, and the activity of peroxidase (POD), maintained Fv/Fm, and decreased lipid peroxidation. Moreover, overexpression of *MsTHI1* upregulated the transcription of *THI4*, *TPK1*, *RbcX2*, *Cu/Zn-SOD*, *CPK13*, and *CPK32* and downregulated the transcription of *TH1* and *CPK17* in transgenic alfalfa under drought stress. These results suggested that *MsTHI1* enhances drought tolerance by strengthening photosynthesis, regulating the antioxidant defense system, maintaining osmotic homeostasis, and mediating plant signal transduction.

KEYWORDS

alfalfa, *MsTHI1*, photosynthesis, antioxidant defense, osmotic homeostasis, drought tolerance

Introduction

Plants must deal with abiotic and biotic stresses in the wild, where they are constantly subjected to unfavorable environmental stresses (Zhu, 2016; Rao et al., 2020). The increased frequency and intensity of extreme weather events due to climate changes make drought stress a significant constraint on agriculture (Rivero et al., 2022). Climate change has made drought an even more substantial problem, dramatically reducing crop yields and quality and limiting plant development (Li et al., 2020). Plants respond to drought conditions by altering photosynthetic, physiological, metabolic, and molecular

activities, which in turn causes membrane rearrangement, osmotic pressure aggravation, the buildup of reactive oxygen species (ROS), in addition to functional abnormalities in cellular components (Roo et al., 2020; Yan et al., 2020). As a result, plants have developed sophisticated defense mechanisms to mitigate the detriment of drought conditions, from molecular and physiological to cellular and ecological levels (Xu and Zhou, 2008; Hussain et al., 2018). Rapid genetic enhancements in drought-tolerant crops may be possible if proteins or genes involved in such alterations are identified.

As an enzyme cofactor, thiamine [vitamin B1 (VB1)] is an integral part of the diets of all living species, since it is required for the appropriate metabolic activity of all living things (Strobbe et al., 2022). In plants, the biosynthesis of thiamine is promoted during adaptation responses to persistent abiotic stresses, such as drought, cold, heat, salinity, and oxidative stress (Rapala-Kozik et al., 2008; Yu et al., 2014; Meng et al., 2017; Amjad et al., 2021; Qian et al., 2022). Thiamine can improve the oxidation state of mitochondria, enhance the activities of pyruvate dehydrogenase (PDH), release ROS quickly when stimulated by stress, and activate downstream genes to induce stress resistance in plants. Thiamine thiazole synthase (THI1), also known as 4-methyl-5-hydroxyethyl thiazole phosphate (HET-P) synthase, is a major enzyme in thiamine biosynthesis, and it performs a crucial function in controlling thiamine production in plants (Chen et al., 2022). Nodule expansion and seed development in *Lotus japonicus* depend on the thiamine synthesis controlled by *THI1* (Nagae et al., 2016). Concurrently, *THI1* reacted to various abiotic stimuli, which may improve mitochondrial DNA damage tolerance (Rapala-Kozik et al., 2012). The expression of *THI* was upregulated in *Elaeis guineensis* (oil palm) under osmotic, salinity, and oxidative stress, and was also found to be predominantly upregulated in the early phase (2–6 h) of the NaCl or sorbitol treatment in *Arabidopsis* (Rapala-Kozik et al., 2012; Abidin et al., 2016). When temperatures increased, *THI1* was more prevalent than usual, while thiamine concentrations were lower than usual in *Oryza meridionalis* (Scafaro et al., 2010). *THI1* plays a vital role in the drought responsiveness in *Arabidopsis* and the abscisic acid (ABA) signaling in guard cells (Livak and Schmittgen, 2000). To regulate ABA activation of slow-type anion channels and ABA-induced stomatal closure in *Arabidopsis*, *THI1* can inhibit Ca^{2+}-dependent protein kinase (CPK33) activity in a plasma membrane-restricted way (Li et al., 2016).

A significant amount of *Medicago sativa*, also known as alfalfa, is grown globally as a leguminous feed crop. Alfalfa is essential for the growth of herbivorous animal husbandry and agricultural sustainability due to its high protein content and palatability and symbiotic nitrogen fixation (Yang et al., 2021). The alfalfa acreage has significantly risen in China since the Chinese government has started a strategy to strengthen the milk and alfalfa sectors in 2012 (Fan et al., 2018). Although the deep root structure of alfalfa protects it from drought on the dry and semi-arid ground, the scanty and erratic precipitation of the region has severely hampered crop yields and quality, stifling the growth in the alfalfa business (Zhang and Shi, 2018; Zhao et al., 2019). Enhancing drought resistance of alfalfa requires first learning about its molecular process by which the plant reacts to drought stress. Even though numerous *THI1* genes have important activities in other plants, little is known about *THI1* in alfalfa. In this investigation, the *MsTHI1* was extracted from alfalfa and was characterized. The transcription levels of *MsTHI1* in reaction to cold, drought, salt, and alkalinity were tested to identify how *MsTHI1* is expressed in response to abiotic stresses. The physiological functions of *MsTHI1* under drought stress have also been studied by overexpressing *MsTHI1* in *Nicotiana benthamiana* (tobacco) and alfalfa. In addition, the regulation mechanism was illustrated by quantifying the expression of associated genes involved in the thiamine pathway and stress response. Our findings demonstrated that *MsTHI1* upregulation increased drought resistance by boosting photosynthetic activities, decreasing ROS formation, and preserving osmotic stability. This study will improve our knowledge on how alfalfa reacts to droughts, which could have implications for stress-resistance breeding efforts.

Materials and methods

Plant materials and stress treatments

Alfalfa (*Medicago sativa* L. cv. Longmu 801) seeds were treated with 75% ethanol solution for 30 s, followed by treatment with 10% NaClO solution for 10 min. The seeds were then germinated on moist filter paper in Petri dishes after being cleaned 4–5 times with ultrapure water. When the seedlings were 5 days old, they were transplanted into a pot with nutrient soil, perlite, and vermiculite (1:1:1). They were placed in a growth chamber with 60% humidity and a 16 h photoperiod at 24°C and irrigated with 1/2 Hoagland nutrient solution every 2 days. The seedlings were exposed to stress treatments for 4 weeks at 0 (control), 3, 6, 12, 24, and 48 h. To simulate the effects of drought, salt, and alkaline stress, the alfalfa plants were transplanted into nutritional solutions containing 15% PEG-6000, 150 mM NaCl, and 150 mM $NaHCO_3$, respectively. To carry out the cold treatment, the plants were placed in a growth chamber maintained at 4°C. Leaves, stems, and roots were promptly removed, frozen in liquid nitrogen, and stored at −80°C for subsequent examination. The experiment was performed in triplicate.

Isolation and sequence analysis of *MsTHI1* gene

The total RNA from alfalfa leaves was isolated using an Ultrapure RNA kit (CWBIO, Beijing, China). Following

the manufacturer's instructions, first-strand complementary DNA (cDNA) was generated using the HiScript II 1st Strand cDNA Synthesis Kit (+gDNA wiper) (Vazyme, Nanjing, China). The coding DNA sequence (CDS) of *MsTHI1* was obtained *via* PCR with degenerate primers (*MsTHI1*-F/*MsTHI1*-R, Supplementary Table S1) that were designed based on the *MtTHI1* sequence of *Medicago truncatula* retrieved from the National Center for Biotechnology Information (NCBI). The coding uses a 2×Unique™ Taq Master Mix (with Dye) (Novogene, Beijing, China) following the manufacturer's instructions. The PCR product was inserted into the pEASY-Blunt Simple Cloning vector using ClonExpress®II One Step Cloning Kit (Vazyme, Nanjing, China) and sequenced by SangonBiotech Co. (Shanghai, China).

DNAMAN software performed multiple comparison determinations (version 9, LynnonBiosoft). The phylogenetic tree was generated using MEGA X software and the neighboring connection method. MEME was employed to examine the conserved motif (http://memesuite.org/). Protparam (http://expasy.org/Proteomics) was used to explore the physical and chemical characteristics of polygenes. The secondary structure of the protein was anticipated by the web application, SPOMA (http://pbil.ibcp.fr). The 3D structure of the *MsTHI1* amino acid sequence was predicted using SWISS-MODEL (https://swissmodel.expasy.org/). Using CELLOv.2.5 (http://cello.life.nctu.edu.tw/), the subcellular localization of amino acid sequences in alfalfa was predicted.

Subcellular localization of *MsTHI1* protein

The coding region of *MsTHI1* without a stop codon was amplified using specific primers (*MsTHI1*-BamHI-F/*MsTHI1*-SacI-R, Supplementary Table S1) with BamHI and SacI restriction sites. The amplicons were double-digested with SacI and BamHI, followed by joining with pCAMBIA-1300 to create a fusion plasmid, pCAMBIA1300-MsTHI1-green fluorescent protein (GFP). The *MsTHI1* fusion protein was transformed into the epidermal cells of *N. benthamiana* through the infiltration of *Agrobacterium tumefaciens*. A confocal laser scanning microscope was used to identify the fluorescence signal after 48 h incubation at 25°C in the dark (Leica TCS SP2 AOBS, Germany).

Quantitative real-time PCR analysis

Quantitative real-time PCR (qPCR) was used to determine how *MsTHI1* is expressed in different tissues and how it responds to environmental stress in transgenic alfalfa. Total RNA was removed from each piece of harvested tissue and reverse transcribed into cDNA. The CDS of *MsTHI1* was used to make the qPCR primers for *MsTHI1*. The GAPDH gene was utilized in alfalfa as an internal control (He et al., 2020). In transgenic tobacco plants, the tobacco *actin* gene (*NtActin*) was used as a reference gene. Using the online NCBI Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/), gene-specific primers for thiamine pathway genes were made and are listed in Supplementary Table S1. Three separate biological replicates and three identical reactions were done with each sample. The $2^{-\Delta\Delta CT}$ comparative method was used to determine how each gene was expressed (Livak and Schmittgen, 2000).

Plant transformation and generation of transgenic plants

For generating transgenic tobacco plants, the gene-specific primers (*MsTHI1*-PstI-F/*MsTHI1*-BamHI-R, Supplementary Table S1) of *MsTHI1* were inserted into the pCAMBIA1300 vector to make pCAMBIA1300-MsTHI1 expression vector. Afterward, the recombinant vector was put into the *A. tumefaciens* strain, EHA105. Then, the recombinant vector was introduced into EHA105, and transformed into wild-type (WT) tobacco by the agroinfiltration method (Kaur et al., 2021). A 20 $mg{\cdot}L^{-1}$ hygromycin (Hyg) was put into the medium for seed germination to find transgenic lines (T1) that are resistant to Hyg. Utilizing specific primers (*Hyg*-F/*Hyg*-R, Supplementary Table S1), a PCR test was used to detect the *Hyg* gene in homozygous transformants (T2), which was validated by qPCR using specific primers for *MsTHI1*. *NtActin* was used as a control within the lab. For more drought stress tests, lines 2 (OV#Nt2), 3 (OV#Nt3), and 7 (OV#Nt7) from T2-generation homozygous lines and WT lines were utilized.

To create transgenic alfalfa plants, the gene-specific primers (*MsTHI1*-SpeI-F/*MsTHI1*-XbaI-R, Supplementary Table S1) of *MsTHI1* were inserted into the pMDC123 vector to make pMDC123-MsTHI1 expression vector. After that, the recombinant vector was put into *A. tumefaciens* strain, LB4404, and transformed into the cotyledon nodes of alfalfa using the agroinfiltration method (Sun et al., 2020). Since the pMDC123 vector has a *Bar* resistance gene that makes it resistant to the herbicide glufosinate, 1.0 $mg{\cdot}L^{-1}$ of glufosinate-ammonium was used to pick out the transformants. The shoots grown back were planted in 1/2 Murashige and Skoog (MS) medium. The herbicide-resistant plants were then put through a PCR test to find the *Bar* gene by employing specific primers (*Bar*-F/*Bar*-R, Supplementary Table S1), which was affirmed by qPCR using specific primers for *MsTHI1*. The *GAPDH* gene from alfalfa was used as a control for the study. The homozygous lines were then passed on and used in more tests to find how drought affects plants.

Drought tolerance assay

On 1/2-strength MS medium plates, seeds of WT tobacco and transgenic tobacco lines were planted and kept at 4°C for 2 days before being moved to a growth chamber at 25°C for 7 days. After the seeds were germinated, they were moved to a greenhouse and put in plastic containers with vermiculite. Drought resistance tests were done on 3-week-old plants. The WT alfalfa and transgenic alfalfa were spread by cutting. Then, the same WT and transgenic plants were moved into plastic pots with vermiculite and put in a greenhouse. Drought tolerance tests were done on 6-week-old plants. For drought tolerance tests, the plants were irrigated for a week with a nutrient solution with 15% PEG-6000. Afterward, the treated leaves of the plants were picked right away, frozen in liquid nitrogen, and kept at −80°C for subsequent use.

Measurement of physiological and biochemical changes

Stress-treated WT, transgenic plants, and three categories of biological repeats, in addition to the three technical repeats, were set up at each time point. A kit was used to quantify the amount of VB1 based on the manufacturer's recommendations (Comin, Suzhou, China). The accumulation of malondialdehyde (MDA) was determined by a modified thiobarbituric acid (TBA) method as described by Puckette et al. (2007). The relative chlorophyll content [soil and plant analyzer development (SPAD) value] was measured by the SPAD-502 Chlorophyll Meter, and three leaves from each group were used to determine the value. Three measurements were made to find the average SPAD value. Three leaves kept in the dark for 20 min were measured thrice to get an average Fv/Fm value. The amount of superoxide anion radical ($O_2^{\cdot-}$) was measured by making some changes to the procedure reported by Jiang and Zhang (2001). The nitroblue tetrazolium (NBT) method (Giannopolitis and Ries, 1972) with a few minor changes were used to test the activity of superoxide dismutase (SOD, EC1.15.1.1). The guaiacol method was used to measure the peroxidase activity (POD, EC 1.11.1.7) (Polle et al., 1994). The amount of proline (Pro) was found using spectrophotometry and the ninhydrin method (Zhao et al., 2009). Wellburn and Lichtenthaler (1984) found out the quantity of chlorophyll *a* (Chl *a*) and chlorophyll *b* (Chl *b*) present in plants. Tietze (1969) used fluorometry to identify reduced glutathione (GSH) in the body. The Bradford method was employed to determine the soluble protein (Bradford, 1976).

Expression analysis of related genes

In WT and transgenic alfalfa, the expression of eight genes associated with the thiamine pathway and stress response was analyzed by employing qPCR; this was done to show how *MsTHI1* helps plants deal with drought. Some of these genes, like thiazole biosynthetic enzyme (THIAMIN4, *THI4*), 4-amino-2-methyl-5-hydroxymethyl pyrimidine phosphate (HMP-P) kinase/thiamin monophosphate (TMP) pyrophosphorylase (*TH1*), thiamine pyrophosphokinase (*TPK1*), chaperonin-like RbcX protein 2 (*RbcX2*), SOD [Cu-Zn] (*Cu/Zn-SOD*), and calcium-dependent protein kinases (*CPK13*, *CPK17*, and *CPK32*) play critical roles in thiamine synthesis and abiotic stress response. The *GAPDH* gene was utilized as an internal reference gene, and Supplementary Table S1 lists the primers employed in this study.

Data analysis

All experiments presented in this study were done three times in the same way. For statistical analysis with SPSS 22.0, the Student's *t*-test and one-way analysis of variance were used to identify significant differences.

Results

Isolation and characteristics of *MsTHI1*

The *MsTHI1* CDS (GenBank accession: MH206189) encoded a protein of 350 amino acids with three putative thiazole synthesis domains, 20 essential amino acids, 28 positively charged amino acid residues (Arg, Lys), and 36 negatively charged amino acid residues (Asp, Glu) (Supplementary Figure S1A). The *MsTHI1* has the chemical formula of $C_{1623}H_{2605}N_{441}O_{500}S_{19}$, the molecular weight (MW) of 36.91 kDa, and a predicted isoelectric point (PI) of 5.68. *MsTHI1* protein aqueous solution at 280 nm has an extinction coefficient of 18,575, fat coefficient of 90.3, an average hydrophilic coefficient of 0.112, and an instability coefficient of 33.24, suggesting the stability of the protein. The neighbor-joining approach used to generate the phylogenetic tree showed that *MsTHI1* and MtTHI1 were on the same branch and that *MsTHI1* was the most similar to *MtTHI1* (Supplementary Figure S1B). *MsTHI1* protein, as accurately predicted by its three-dimensional structure, consists of 32.86% α-helices, 18.57% β-extension chain, 8.29% β-rotation angle, and 40.29% irregular curls (Supplementary Figure S1C).

Subcellular localization of *MsTHI1*

To investigate the subcellular distribution of *MsTHI1*, a GFP-MsTHI1 fusion protein was transiently coexpressed in *N. benthamiana* leaves and then visualized using laser scanning confocal microscopy. MsTHI1-GFP fluorescence was found in the cell membrane and chloroplasts, as evidenced by the brilliant field under infected leaves, as well as the fusion picture

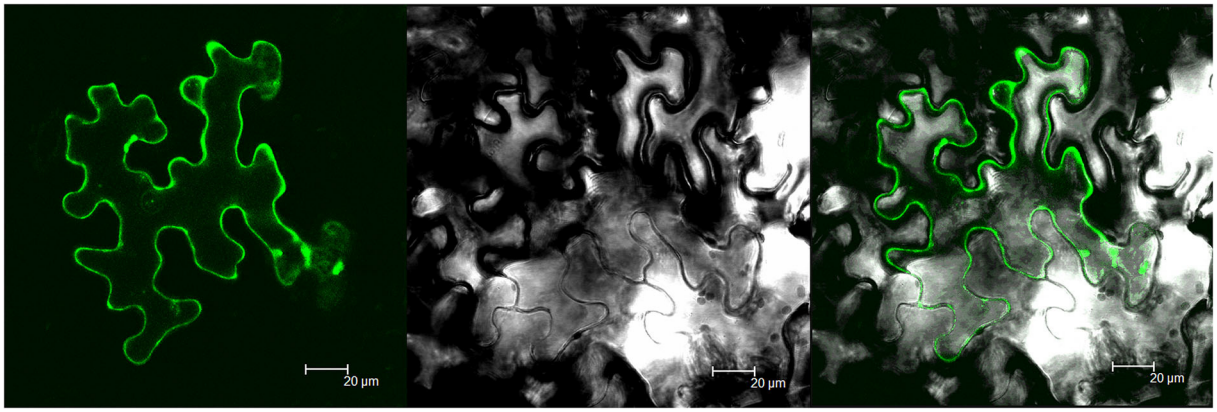


FIGURE 1
Subcellular localization of *MsTHI1*. **(A)** fusion GFP fluorescence. **(B)** bright field under infected leaves, **(C)** fusion picture. Scale bar, 20 μm.

(Figure 1). However, predictions of CELLO 2.5 to find where *MsTHI1* is subcellularly localized suggested that it is most likely to be found in chloroplasts (4.275). Therefore, the *MsTHI1* protein was primarily found in chloroplasts.

MsTHI1 expression patterns

The expression profiles of *MsTHI1* in distinct alfalfa tissues and its sensitivities to drought, cold, salt, and alkaline stress were studied through qPCR. The findings demonstrated that the expression pattern of *MsTHI1* in alfalfa leaves was more significant than in the roots and stems (Figure 2A). The expression level of *MsTHI1* was significantly suppressed during cold stress (Figure 2B). The expression level of *MsTHI1* changed comparably with 15% of PEG and 150 mM of NaCl treatment. They peaked at 3 h, dropped at 6 and 12 h, and climbed once more at 24 and 48 h. Contrary to the *MsTHI1* upregulation under salt stress, the *MsTHI1* expression was significantly suppressed at 6 and 12 h under drought stress. The expression level of *MsTHI1* at 150 mM of $NaHCO_3$ peaked at 3 h, followed by a drop from 6 to 48 h. These findings indicated that *MsTHI1* expression was tissue-specific in alfalfa and that *MsTHI1* might implicate the self-protection mechanisms of alfalfa in response to aggressive ecological scenarios.

The ability of transgenic plants to tolerate drought is enhanced by *MsTHI1* upregulation

Transgenic tobacco (T2) and alfalfa plants were created and validated with Hyg and *Bar*, respectively, through PCR to examine the involvement of *MsTHI1* in the plant drought tolerance (Supplementary Figure S2). *MsTHI1* transcript levels in OV#Nt2, OV#Nt3, as well as OV#Nt7 transgenic lines were 413.6, 3593.0, and 3014.8 times greater than those in WT tobacco plants, accordingly (Figure 3A). The expression of *MsTHI1* was 29.2 and 6.9 times greater in OV#Ms7 and OV#Ms9 transgenic lines, respectively, compared to that in WT alfalfa plants (Figure 3B). No phenotype was observed between *MsTHI1*-transformed and WT plants under normal control conditions in either transgenic tobacco or transgenic alfalfa (Figures 3C,D). When exposed to dryness for a week, the leaves of both WT tobacco and alfalfa plants wilted, and the growth was significantly suppressed. Despite this, OV#Nt2, OV#Nt3, OV#Nt7, OV#Ms7, and OV#Ms9 continued to develop properly with only a few yellow leaves. The results demonstrated that the drought resistance of transgenic plants was enhanced by overexpressing *MsTHI1*.

Modifications in the physiology of transgenic plants that overexpress *MsTHI1* during conditions of drought stress

To better demonstrate the role of *MsTHI1* in stress conditions tolerance, the physiological modifications in the transgenic tobacco or even alfalfa plants experiencing environmental stresses were studied. Transgenic tobacco with an increased VB1 concentration was found both in normal and drought conditions due to *MsTHI1* upregulation (Figure 4). During normal conditions, MDA levels in WT plants were higher than in OV#Nt2 and OV#Nt3 species. Even though the SPAD value of the transgenic tobacco was lesser than that of the WT plants in the control environment, all plants increased their SPAD value after being subjected to

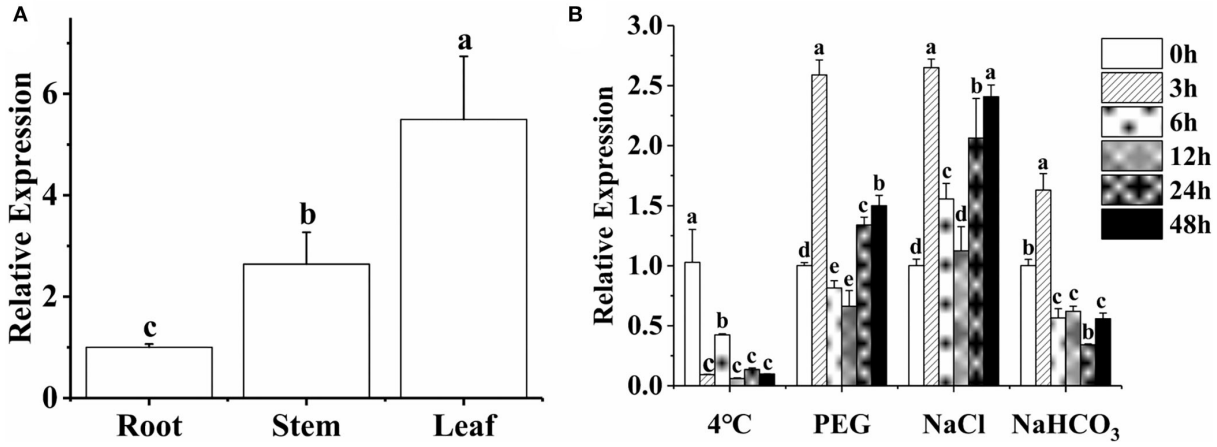


FIGURE 2
Expression pattern of *MsTHI1*. **(A)** Tissue-specific expression of *MsTHI* in the alfalfa root, stem, and leaf. **(B)** Expression analysis of *MsTHI1* in leaves of alfalfa subjected to cold (4°C), drought (15% PEG), salt (150 mM of NaCl), and alkaline (150 mM of $NaHCO_3$) stresses. The expression levels were normalized in the alfalfa *GAPDH* gene. Mean ± standard deviations from three biological replicates are shown with error bars. These letters denote variations that are statistically significant ($\alpha = 0.05$).

drought treatment, with OV#Nt7 plants showing the most significant increase. During drought conditions, the Fv/Fm ratio decreased more quickly in WT plants than in transgenic tobacco, with OV#Nt2 plants having the highest Fv/Fm ratio (0.79). Following drought stress, the WT and transgenic tobacco showed an increase in $O_2^{\cdot-}$ concentration and SOD and POD activities. While WT plants increased their $O_2^{\cdot-}$ level quicker than the transgenic tobacco, the latter eventually caught up. Considering drought stress, there was no significant difference between the WT and transgenic tobacco in free Pro concentration.

During drought conditions, VB1 levels in transgenic and WT alfalfa decreased, but the transgenic plants maintained a greater VB1 concentration than the WT ones. During drought conditions, MDA levels of transgenic and WT alfalfa were elevated, whereas the MDA level of transgenic alfalfa was lower under control or drought conditions. Chl *a* and Chl *b* concentrations were reduced during drought and were increased in *MsTHI*-transformed plants, mirroring the pattern of VB1. Transgenic and WT alfalfa both elevated their $O_2^{\cdot-}$ levels in response to drought conditions, whereas the $O_2^{\cdot-}$ level of transgenic alfalfa remained lower in control and drought conditions. However, the functions of SOD and the concentrations of GSH in transgenic alfalfa plants were significantly decreased compared to WT plants during control and droughts scenarios, confirming the role of VB1 in alfalfa drought tolerance. Transgenic alfalfa and WT plants exhibited statistically significant elevations in POD activity and Pro content in response to drought conditions.

Expression analysis of associated genes in the overexpression of *MsTHI1* in transgenic alfalfa under drought stress

Eight associated genes involving *THI4*, *TH1*, *TPK1*, *RbcX2*, *Cu/Zn-SOD*, *CPK13*, *CPK17*, and *CPK32* were assessed for their expression patterns in transgenic and WT alfalfa plants under drought conditions using qPCR. As depicted in Figure 6, drought stress dramatically upregulated the expression of *THI4*, *TPK1*, *RbcX2*, and *CPK32* in transgenic and WT alfalfa plants. Intriguingly, during drought stress, *TH1* and *CPK17* expression levels were reduced in transgenic alfalfa while they were increased in WT alfalfa. During drought conditions, *Cu/Zn-SOD* expression did not alter much in the transgenic alfalfa, whereas it was significantly upregulated in WT alfalfa. During drought stress, the *CPK13* expression was elevated in the transgenic alfalfa but suppressed in WT plants.

Discussion

To support proper metabolic activity, people depend on an adequate thiamin intake from food. Metabolic engineering to boost thiamin production in plants is a promising approach to reducing VB1 deficiency and, by extension, improving worldwide human health (Strobbe et al., 2021). In addition to its established roles in thiamine synthesis and mitochondrial DNA damage resistance, *THI1* contributes significantly to plant abiotic response to stress (Rapala-Kozik et al., 2012; Li et al., 2016). Several plant *THI1* genes that respond to

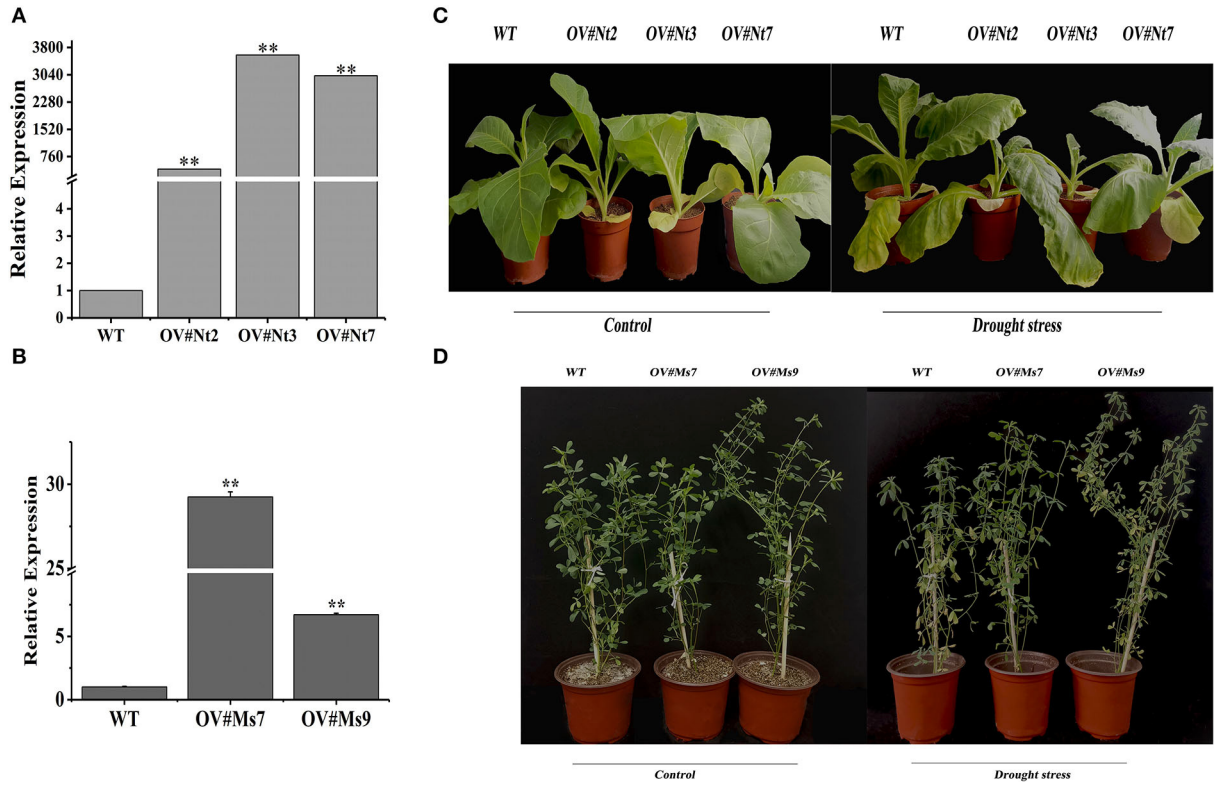


FIGURE 3
Identification of transformation with *MsTHI1 via* qPCR and phenotype of transformation with *MsTHI1* under control conditions and drought stress. **(A)** Identifying transgenic tobacco *via* quantitative real-time PCR (qPCR). **(B)** Identifying transgenic alfalfa *via* qPCR. **(C)** Phenotype of three-week-old transgenic tobacco plants with overexpressed *MsTHI1* under control and drought stress for seven days. **(D)** Phenotype of six-week-old transgenic alfalfa plants with overexpressed *MsTHI1* under control and drought stress for seven days. *NtActin* and alfalfa *GAPDH* gene were used as the internal reference control for identifying transgenic tobacco and alfalfa, respectively, *via* qPCR. Mean ± standard deviations from three biological replicates are shown with error bars. "**": extremely significant difference between WT and transformation line ($P < 0.01$). WT: wild type; OV#Nt2, OV#Nt3, and OV#Nt7: transgenic tobacco lines 2, 3, and 7, respectively. OV#Ms7 and OV#Ms9: transgenic alfalfa lines 7 and 9, respectively.

drought, salinity, heat, and oxidative stress have been identified (Livak and Schmittgen, 2000; Scafaro et al., 2010; Abidin et al., 2016; Mangel et al., 2017; Chen et al., 2022), but no research has been conducted on alfalfa. Here, we report the isolation and functional characterization of a *THI1* gene (called *MsTHI1*) from alfalfa. The protein encoded by *MsTHI1* is 350 amino acids long. It has three putative thiazole synthesis domains (Supplementary Figure S1A), all of which bind to 2-carboxylate-4-methyl-5-beta-(ethyladenosine 5'-diphosphate) thiazole, a possible intermediate in the thiazole biosynthesis of Eukaryotes (Godoi et al., 2006). The MsTHI1 protein was localized mainly in the chloroplast of plant cells (Figure 1), as is the case for the translation initiation at the first AUG (start codons) directs translocation of *THI1* to chloroplasts, indicating a high requirement for *THI1* in chloroplasts (Chabregas et al., 2003). Similar to *THI1* expression in *Arabidopsis* (Teixeira et al., 2005; Dong et al., 2015), the expression level of *MsTHI1* in alfalfa leaves was higher than that in the roots and stems (Figure 2A), confirming the tissue-specific expression of *THI1*. The *MsTHI1* was quickly upregulated in alfalfa leaves within 3 h during drought, salt, and alkaline stress, and downregulated under cold stress (Figure 2B), which might contribute to the self-protection mechanisms of alfalfa in response to unfavorable ecological scenarios. Although the expression of *MsTHI1* was similar in response to drought and salt, *MsTHI1* responses were more sensitive to salt, which may be associated with the characters of the alfalfa cultivar used in the study.

Overexpressing *MsTHI1* can increase drought resistance in transgenic plants. Drought stress significantly stunted the development of WT tobacco and alfalfa plants (Figure 3). To better understand the role of *MsTHI1* in plants during drought conditions, we studied the physiological and biochemical responses of transgenic plants. VB1 is a vital member of the vitamin B group, and the elevation in the level of VB1 is important in alleviating stressful conditions (Amjad et al., 2021). Our results demonstrated that VB1 levels in transgenic tobacco

10.3389/fpls.2022.992024

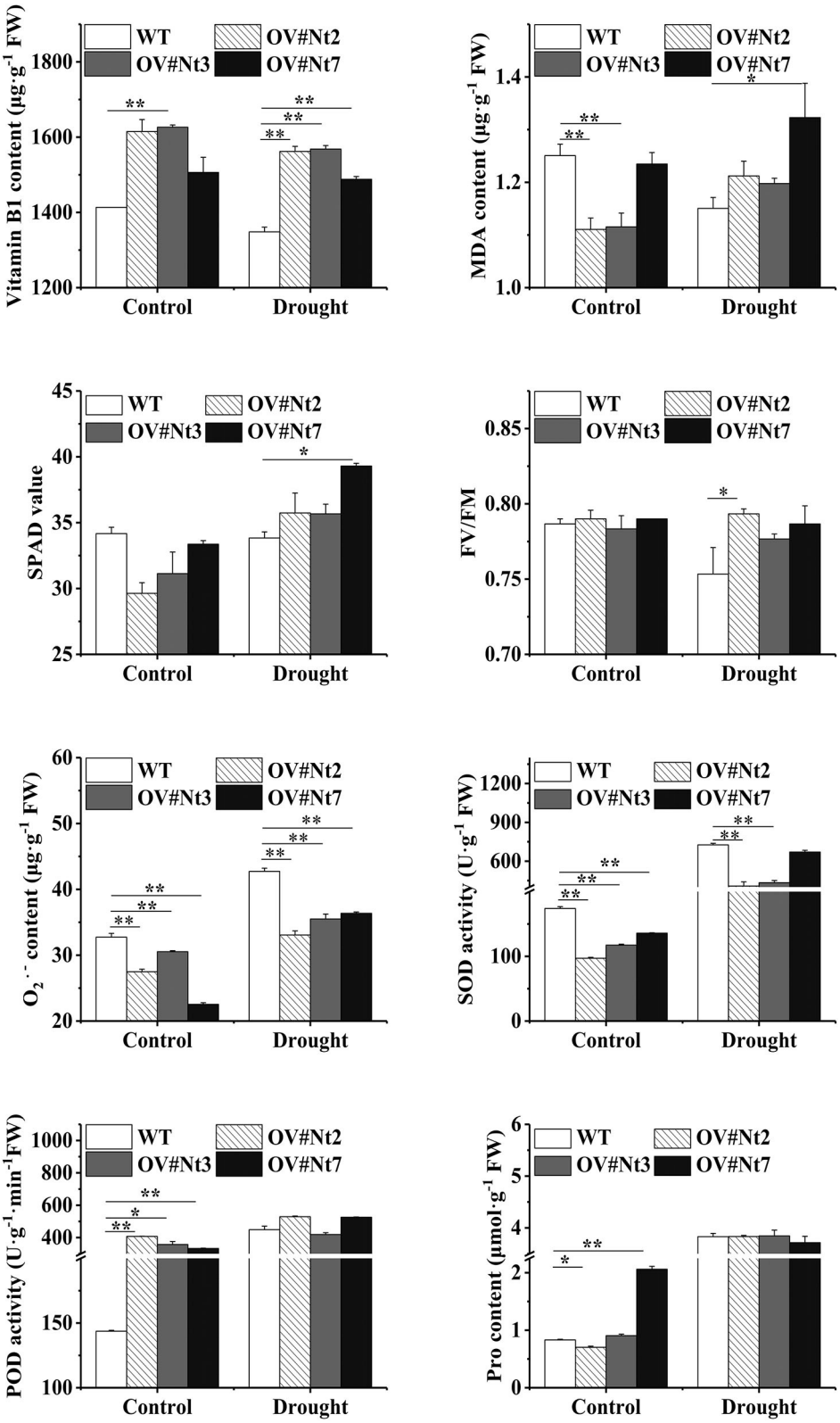


FIGURE 4
Modifications in the physiology of transgenic tobacco plants that overexpress *MsTHI1* during control or under drought stress. Mean ± standard deviations from three biological replicates are displayed with error bars. MDA, malondialdehyde. SPAD, soil and plant analyzer development.
(Continued)

FIGURE 4 (Continued)
Fv/Fm, the PS II primary light energy conversion efficiency. $O_2^{\cdot-}$, superoxide anion radical. SOD, superoxide dismutase. POD, peroxidase. Pro, proline. WT, wild type. OV#Nt2, OV#Nt3, and OV#Nt7, transgenic tobacco lines 2, 3, and 7, respectively. "*", significant difference between WT and transgenic tobacco lines ($P < 0.05$). "**", extremely significant difference between WT and transgenic tobacco lines ($P < 0.01$).

and alfalfa were always significantly higher than in WT plants. However, VB1 contents diminished in WT and transgenic plants under drought conditions suggesting that *MsTHI1* overexpression can counteract drought stress by consuming VB1. The MDA amount was also analyzed as a vital sign of lipid peroxidation in the membrane, reflecting stress-induced molecular damage. Following exposure to drought, transgenic *SbWRKY30 Arabidopsis*, as well as rice plants, had lower levels of MDA than WT plants (Yang et al., 2020). In this case, MDA levels in *MsTHI1*-overexpressed alfalfa were considerably lower than in WT alfalfa during control and drought stress, demonstrating a robust ability to reduce membrane damage. However, the MDA content in the transgenic tobacco was higher than that in the WT tobacco after drought stress. The results may reflect the difference in ROS scavenging capacity and plasma membrane damage in alfalfa and tobacco. Balanced photosynthesis under drought is necessary for better survivability and agricultural benefits in terms of biomass and output (Daszkowska-Golec et al., 2019). Leaf SPAD readings were directly correlated with grain output, and most of the high SPAD genotypes sustained hot canopies under drought (Raina et al., 2019). The chlorophyll contents and Fv/Fm, utilized as reliable benchmark indicators in identifying drought-adapted genotypes, were reduced under drought-treated *Zea mays* (maize) seedlings (Chen et al., 2016). In this study, the SPAD value in OV#Nt7 was substantially higher than that in the WT tobacco under drought stress. Simultaneously, the decline rate of Fv/Fm in the WT tobacco was faster than in transgenic plants during drought conditions. The levels of Chl *a* and Chl *b* reduced during drought conditions, but the alfalfa transformed with *MsTHI* had more significant Chl *a* and Chl *b* levels. These results revealed that *MsTHI1* upregulation might have a function of controlling chlorophyll content to balance photosynthesis, thus conferring a degree of drought conditions.

Drought stress can produce an immediate generation of ROS. Plants might elicit an antioxidant protective mechanism to scavenge superfluous ROS and protect the cells against oxidative stress during stressful circumstances (Sharma et al., 2012). The antioxidant enzymes, including SOD and POD, and non-enzymatic antioxidants, including GSH, are critical in detoxifying stress-induced ROS. Herein, we noticed that the transgenic tobacco and alfalfa revealed a lower $O_2^{\cdot-}$ level than WT plants during drought conditions (Figures 4, 5). Nevertheless, SOD activities and GSH concentration in the transgenic alfalfa were similarly lesser than in WT plants. The variations in POD activities depend on the expression pattern of *MsTHI1* in transformants. There was no difference in the POD activity between transgenic and WT tobacco under drought stress. Conversely, the concentration of soluble protein fell quickly during drought conditions. Previous findings have indicated that the rise of VB1 by stimulating thiamine metabolism protected cell organelles and decreased the ROS levels in the cells of *Taxus Chinensis*, and chlorophyll could be employed as an antioxidant (Meng et al., 2017; Agathokleous et al., 2020). Consequently, ROS scavenging in transgenic plants might correlate with *MsTHI1* upregulation, which triggered the thiamine metabolism and the aggregation of chlorophyll to decrease the amount of ROS. The soluble protein and free Pro are key osmotic regulators, which can raise the osmotic potential of cells and boost plant tolerance to stressful conditions (Hosseinifard et al., 2022). In the present research, the soluble protein level in transgenic alfalfa was mainly greater than that in WT alfalfa during control or drought settings. Although the level of Pro in transgenic alfalfa was consistently lower than that in WT alfalfa throughout control and drought situations, there was no difference in the level of Pro between transgenic and WT tobacco under drought conditions. These findings demonstrated that enhancing drought tolerance in transgenic plants might link with the buildup of soluble protein during drought environments.

Upregulated *MsTHI1* alfalfa responds to drought conditions by upregulating genes involved in the thiamine system and stress response. The suicide enzyme, THI4 is responsible for mediating the production of the thiazole precursor of thiamin (cThz-P), and hence, it must be destroyed and resynthesized (Sun et al., 2019). *Verticillium dahliae*, a disease of vascular plants, is more resistant to stresses like UV damage and oxidative stress due to *VdTHI4* (Hoppenau et al., 2014). *TH1*, a thiamine phosphate synthase, was upregulated in response to oxidative stress due to its involvement in the phosphorylation of HMP-P from plant thiamine thiazole to HMP-pyrophosphate (HMP-PP) (Strobbe et al., 2021). TPK is a crucial enzyme in the biosynthetic pathways of thiamine, and during environmental stresses, TPK activities were significantly increased in regulating thiamine metabolism in maize seedlings (Rapala-Kozik et al., 2008). During drought conditions, this research found that *THI4* and *TPK1* transcription levels were elevated in transgenic alfalfa, whereas *TH1* levels were declined (Figure 6). *RbcX2* can carboxylate Rubisco, which boosts photosynthesis and can also fold proteins to control the accumulation of soluble protein and drought resistance in plants (Liu et al., 2010; Doron et al., 2020). One form of SOD, Cu/Zn-SOD, is fundamental due to

10.3389/fpls.2022.992024

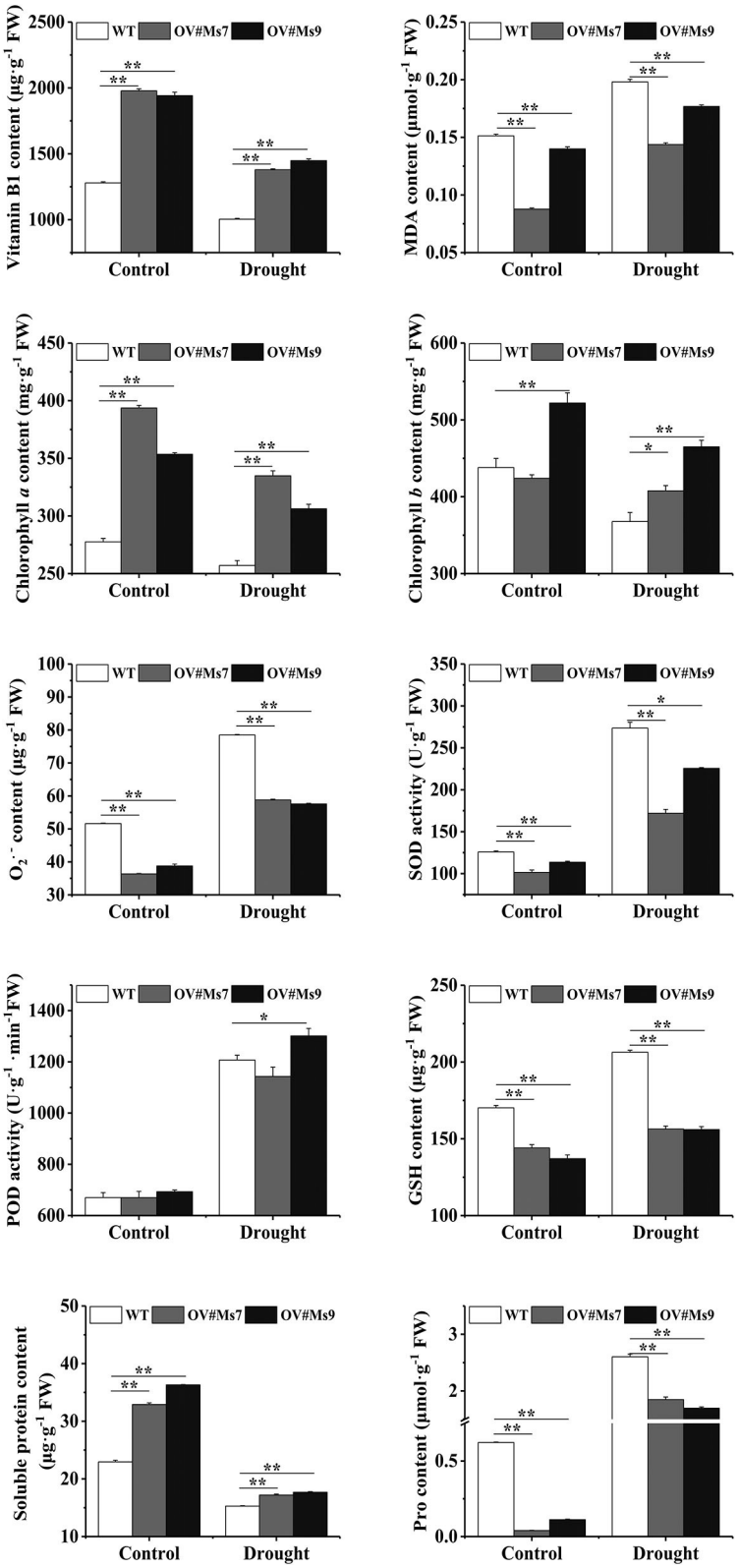


FIGURE 5
Physiological changes of transgenic alfalfa plants with overexpressing *MsTHI1* under control and drought stress. Mean ± standard deviations from three biological replicates are depicted with error bars. MDA, malondialdehyde. O_2^- superoxide anion radical. SOD, superoxide dismutase.
(Continued)

FIGURE 5 (Continued)
POD, peroxidase. GSH, reduced glutathione. Pro, proline. WT, wild type. OV#Ms7 and OV#Ms9, transgenic alfalfa lines 7 and 9, respectively. "*", significant difference between WT and transgenic tobacco lines ($P < 0.05$). "**", extremely significant difference between WT and transgenic tobacco lines ($P < 0.01$).

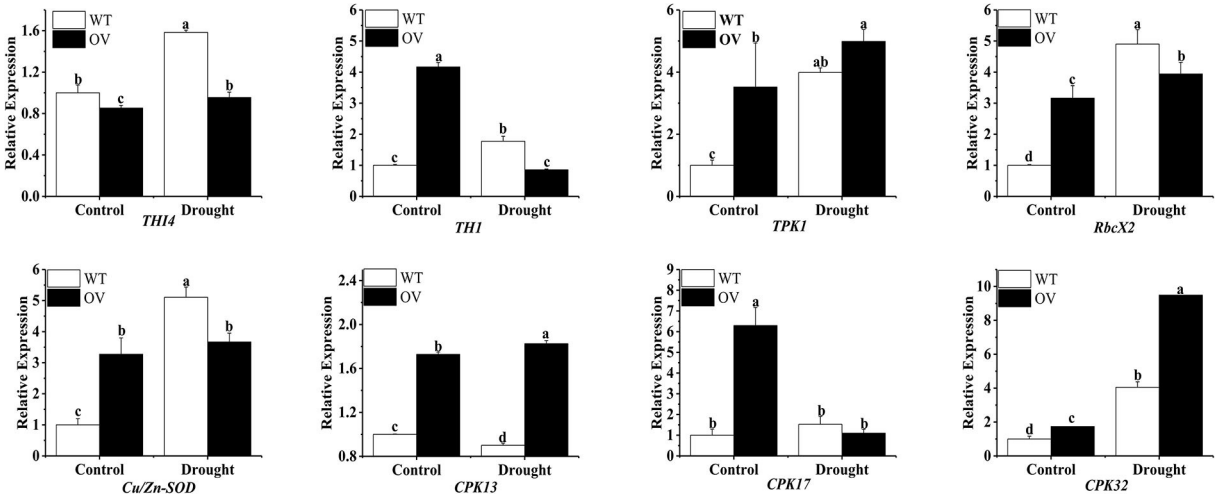


FIGURE 6
Expression of eight associated genes in transgenic alfalfa plants with overexpressing *MsTHI1* under drought stress. WT, wild-type alfalfa. OV, overexpressing *MsTHI1* alfalfa. *THI4*, thiazole biosynthetic enzyme THIAMIN4. *TH1*, HMP-P kinase/TMP pyrophosphorylase. *TPK1*, thiamine pyrophosphokinase. *RbcX2*, chaperonin-like RbcX protein 2. *Cu/Zn-SOD*, superoxide dismutase [Cu-Zn]. *CPK13*, *CPK17*, and *CPK32*: calcium-dependent protein kinases. Mean ± standard deviations from three biological replicates are shown with error bars. These letters denote statistically significant variations ($\alpha = 0.05$).

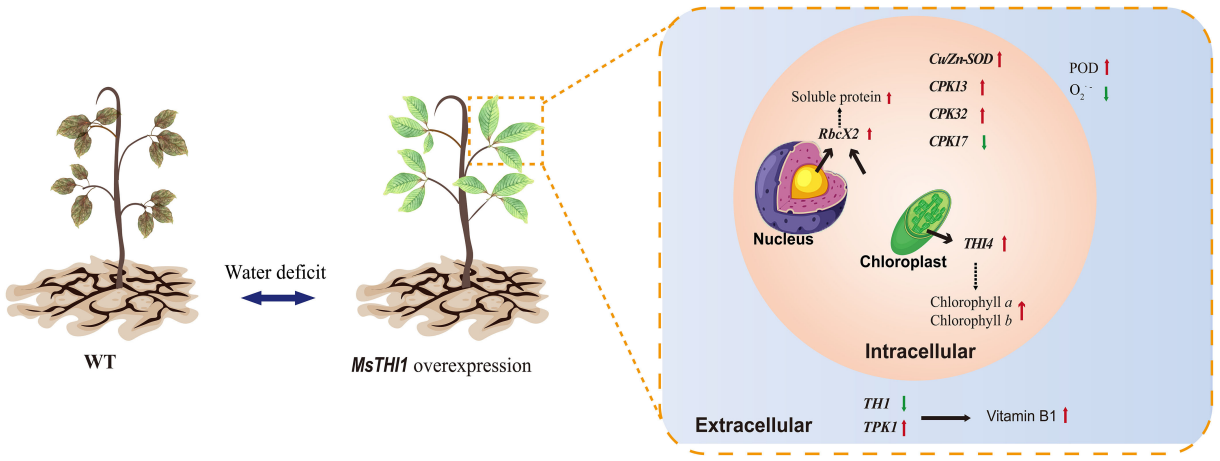


FIGURE 7
Possible function model of *MsTHI1* response to drought stress. The red arrows represent significant increases in physiological signs or genes being overexpressed. Green arrows represent decreased expressions of genes or declines in physiological markers. Dotted lines represent potential regulating pathways. *CPK13*, *CPK 17*, and *CPK32*: calcium-dependent protein kinases. *Cu/Zn-SOD*, superoxide dismutase [Cu-Zn]. O_2^-, superoxide anion radical. POD, peroxidase. *RbcX2*, chaperonin-like RbcX protein 2. *TH1*, HMP-P kinase/TMP pyrophosphorylase. *THI4*, thiazole biosynthetic enzyme THIAMIN4. *TPK1*, thiamine pyrophosphokinase. WT, wild type.

its role in ROS elimination and its correlation with drought tolerance in plants (Saed-Moucheshi et al., 2021; Zhang et al., 2021). Both transgenic and WT alfalfa showed an increase in *RbcX2* transcription in response to drought stress, while WT alfalfa showed a marked increase in *Cu/Zn-SOD* transcription. Changes in the content of soluble protein and SOD activity in the

drought tolerance experiment of transformants were in line with these findings. Serine/threonine protein kinases, of which CPKs are a subfamily, interact with *THI1* to confer resistance to abiotic stressors, such as dehydration, salinity, and low temperature (Atif et al., 2019; Zhang et al., 2020; Chen et al., 2021). Variations in *CPK13*, *CPK17*, and *CPK32* transcription were observed in the present study. These findings indicated that CPKs might have a role in regulating the response of *MsTHI1* to drought stress; however, the mechanisms by which CPKs regulated *MsTHI1* expression still require investigation.

Accordingly, we postulated that *MsTHI1* participation in the plant drought tolerance might include a complicated regulating network that includes the physiological and gene expression levels (Figure 7). The overexpression of *MsTHI1* can enhance the transcription of *THI4*, *TPK1*, *RbcX2*, *Cu/Zn-SOD*, *CPK13,* and *CPK32*, and decrease the transcription of *TH1* and *CPK17* to increase the accumulation of VB1, Chl *a*, Chl *b*, soluble protein, and POD activities, which helps preserve osmotic balance and strengthens photosynthesis and ROS scavenging.

Conclusion

Overall, we could identify the thiamine thiazole synthase, *MsTHI1*, from alfalfa and determine its beneficial effect on the enhancement of drought stress tolerance. *MsTHI1* localized to chloroplasts and alfalfa leaves exhibited excellent expression levels. Simultaneously, exposure to stressful conditions increased *MsTHI1*. The upregulation of *MsTHI1* enhanced drought tolerance in plants and led to an enhancement in photosynthesis by increasing SPAD, maintaining Fv/Fm, and aggregating Chl *a* and Chl *b*, reducing the ROS accumulation and lipid peroxidation by elevating VB1 synthesis and POD activity, reducing MDA, and maintaining the balance of osmotic by strengthening the accumulation of soluble protein. Moreover, *MsTHI1* mediated signal transduction by upregulating the transcription of *THI4*, *TPK1*, *RbcX2*, *Cu/Zn-SOD*, *CPK13,* and *CPK32*, and downregulating the transcription of *TH1* and *CPK17* in transgenic alfalfa under drought stress. Our findings shed light on the mechanism by which *MsTHI1* alleviates plant drought stress and suggest that it should be incorporated into successful breeding initiatives to enhance the tolerance of alfalfa to drought.

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

HY, GC, and PZ designed the experiment. HY experimented and wrote the manuscript. ZW, HL, YZ, and MY performed part of the experiment and analyzed and discussed data analysis. GC and PZ supervised the project and revised the manuscript. All authors agree to be accountable for the content of the work, contributed to the article and approved the submitted version.

Funding

This research was funded by the Heilongjiang Provincial Natural Science Foundation of China (YQ2021C019).

Acknowledgments

We would like to thank Prof. Hua Cai from the Northeast Agricultural University for kindly providing the pMDC123 vector and all reviewers for their helpful comments on 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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.992024/full#supplementary-material

SUPPLEMENTARY FIGURE S1
Alignment of homologous protein sequences. **(A)**. phylogenetic tree generation **(B)**, and three-dimensional modeling of *MsTHI1* **(C)**.

SUPPLEMENTARY FIGURE S2
Identification of transformation with *MsTHI via* PCR. **(A)**. PCR identification of transgenic tobacco with *Hyg*. **(B)**. PCR identification of transgenic alfalfa with *Bar*. M, Trans2K DNA Marker; "-", negative control; "+", positive control; WT, wild type; number, transformation line.

SUPPLEMENTARY TABLE S1
List of primers used in this study.

References

Abidin, A. A. Z., Yee, W. S., Rahman, N. S. A., Idris, Z. H. C., and Yusof, Z. N. B. (2016). Osmotic, oxidative and salinity stresses upregulate the expressions of Thiamine (vitamin B1) biosynthesis genes (*THIC* and *THI1*/*THI4*) in oil palm (*Elaeis guineensis*). *J. Oil Palm Res.* 28, 308–319. doi: 10.21894/jopr.2016.2803.07

Agathokleous, E., Feng, Z., and Penuelas, J. (2020). Chlorophyll hormesis: are chlorophylls major components of stress biology in higher plants? *Sci. Total Environ.* 726, 138637. doi: 10.1016/j.scitotenv.2020.138637

Amjad, S. F., Mansoora, N., Yaseen, S., Kamal, A., Butt, B., Matloob, H., et al. (2021). Combined use of endophytic bacteria and pre-sowing treatment of thiamine mitigates the adverse effects of drought stress in wheat (*Triticum aestivum* L.) Cultivars. *Sustainability* 13, 6582. doi: 10.3390/su13126582

Atif, R. M., Shahid, L., Waqas, M., Ali, B., Rashid, M. A. R., Azeem, F., et al. (2019). Insights on Calcium-Dependent Protein Kinases (CPKs) signaling for abiotic stress tolerance in plants. *Int. J. Mol. Sci.* 20, 5298. doi: 10.3390/ijms20215298

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248–254. doi: 10.1016/0003-2697(76)90527-3

Chabregas, M. S., Luche, D. D., Van Sluys M-A., Menck C. F. M., Silva-Filho M. C. (2003). Differential usage of two in-frame translational start codons regulates subcellular localization of *Arabidopsis thaliana THI1*. *J. Cell Sci.* 116, 285–291. doi: 10.1242/jcs.00228

Chen, D. Q., Wang, S. W., Cao, B. B., Cao, D., Leng, G. H., Li, H. B., et al. (2016). Genotypic variation in growth and physiological response to drought stress and re-watering reveals the critical role of recovery in drought adaptation in maize seedlings. *Front. Plant Sci.* 6, 1241. doi: 10.3389/fpls.2015.01241

Chen, P. F., Liu, P., Zhang, Q. F., Zhao, L., Hao, X. R., Liu, L., et al. (2022). Dynamic physiological and transcriptome changes reveal a potential relationship between the circadian clock and salt stress response in *Ulmus pumila*. *Mol. Genet. Genom.* 297, 303–317. doi: 10.1007/s00438-021-01838-2

Chen, X., Ding, Y., Yang, Y., Song, C., Wang, B., Yang, S., et al. (2021). Protein kinases in plant responses to drought, salt, and cold stress. *J. Integr. Plant Biol.* 63, 53–78. doi: 10.1111/jipb.13061

Daszkowska-Golec, A., Collin, A., Sitko, K., Janiak, A., Kalaji, H. M., and Szarejko, I. (2019). Genetic and physiological dissection of photosynthesis in barley exposed to drought stress. *Int. J. Mol. Sci.* 20, 6341. doi: 10.3390/ijms20246341

Dong W., Stockwell V. O., and Goyer A. (2015). Enhancement of Thiamin content in arabidopsis thaliana by metabolic engineering. *Plant Cell Physiol.* 56, 2285–2296. doi: 10.1093/pcp/pcv148

Doron, L., Xu, L., Rachmilevitch, S., and Stern, D. B. (2020). Transgenic overexpression of rubisco subunits and the assembly factor RAF1 are beneficial to recovery from drought stress in maize. *Environ. Exp. Botany* 177, 104126. doi: 10.1016/j.envexpbot.2020.104126

Fan, Q., Creamer, R., and Li, Y. (2018). Time-course metabolic profiling in alfalfa leaves under *Phoma medicaginis* infection. *PloS ONE* 13, e0206641. doi: 10.1371/journal.pone.0206641

Giannopolitis, C. N., and Ries, S. K. (1972). Superoxide dismutase I. Occurrence in higher plants. *Plant Physiol.* 59, 309–314. doi: 10.1104/pp.59.2.309

Godoi, P. H. C., Galhardo, R. S., Luche, D. D., Van Sluys, M.-A., Menck, C. F. M., and Oliva, G. (2006). Structure of the thiazole biosynthetic enzyme THI1 from *Arabidopsis thaliana*. *J. Biol. Chem.* 281, 30957–30966. doi: 10.1074/jbc.M604469200

He, L., Liu, Y., He, H., Liu, Y., Qi, J., Zhang, X., et al. (2020). A molecular framework underlying the compound leaf pattern of *Medicago truncatula*. *Nature Plants* 6, 511–521. doi: 10.1038/s41477-020-0642-2

Hoppenau, C. E., Tran, V. T., Kusch, H., Asshauer, K. P., Landesfeind, M., Meinicke, P., et al. (2014). *Verticillium dahliae VdTHI4*, involved in thiazole biosynthesis, stress response and DNA repair functions, is required for vascular disease induction in tomato. *Environ. Exp. Botany* 108, 14–22. doi: 10.1016/j.envexpbot.2013.12.015

Hosseinifard, M., Stefaniak, S., Ghorbani Javid, M., Soltani, E., Wojtyla, Ł., and Garnczarska, M. (2022). Contribution of exogenous proline to abiotic stresses tolerance in plants: a review. *Int. J. Mol. Sci.* 23, 5186. doi: 10.3390/ijms23095186

Hussain, H. A., Hussain, S., Khaliq, A., Ashraf, U., Anjum, S. A., Men, S., et al. (2018). Chilling and drought stresses in crop plants: implications, cross talk, and potential management opportunities. *Front. Plant Sci.* 9, 393. doi: 10.3389/fpls.2018.00393

Jiang, M., and Zhang, J. (2001). Effect of abscisic acid on active oxygen species, antioxidative defence system and oxidative damage in leaves of maize seedlings. *Plant Cell Physiol.* 42, 1265–1273. doi: 10.1093/pcp/pce162

Kaur, M., Manchanda, P., Kalia, A., Ahmed, F. K., Nepovimova, E., Kuca, K., et al. (2021). Agroinfiltration Mediated Scalable Transient Gene Expression in Genome Edited Crop Plants. *Int. J. Mol. Sci.* 22, 10882. doi: 10.3390/ijms221910882

Li, C. L., Wang, M., Wu, X. M., Chen, D. H., Lv, H. J., Shen, J. L., et al. (2016). THI1, a Thiamine Thiazole Synthase, Interacts with Ca^{2+}-Dependent Protein Kinase CPK33 and modulates the S-type anion channels and stomatal closure in *Arabidopsis*. *Plant Physiol.* 170, 1090–1104. doi: 10.1104/pp.15.01649

Li, P., Liu, H., Yang, H., Pu, X., Li, C., Huo, H., et al. (2020). Translocation of drought-responsive proteins from the chloroplasts. *Cells* 9, 259. doi: 10.3390/cells9010259

Liu, C., Young, A. L., Starling-Windhof, A., Bracher, A., Saschenbrecker, S., Rao, B. V., et al. (2010). Coupled chaperone action in folding and assembly of hexadecameric Rubisco. *Nature* 463, 197–U181. doi: 10.1038/nature08651

Livak, K., and Schmittgen, T. (2000). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Mangel, N., Fudge, J. B., Fitzpatrick, T. B., Gruissem, W., and Vanderschuren, H. (2017). Vitamin B-1 diversity and characterization of biosynthesis genes in cassava. *J. Exp. Botany* 68, 3351–3363. doi: 10.1093/jxb/erx196

Meng, D. L., Yu, X. H., Ma, L. Y., Hu, J., Liang, Y. L., Liu, X. D., et al. (2017). Transcriptomic response of Chinese Yew (*Taxus chinensis*) to cold stress. *Front. Plant Sci.* 8, 468. doi: 10.3389/fpls.2017.00468

Nagae, M., Parniske, M., Kawaguchi, M., and Takeda, N. (2016). The thiamine biosynthesis gene THI1 promotes nodule growth and seed maturation. *Plant Physiol.* 172, 2033–2043. doi: 10.1104/pp.16.01254

Polle, A., Otter, T., and Seifert, F. (1994). Apoplastic Peroxidases and Lignification in Needles of Norway Spruce (*Picea abies* L.). *Plant Physiol.* 106, 53–60. doi: 10.1104/pp.106.1.53

Puckette, M. C., Hua, W., and Mahalingam, R. (2007). Physiological and biochemical responses to acute ozone-induced oxidative stress in Medicago truncatula. *Plant Physiol. Biochem.* 45, 70–79. doi: 10.1016/j.plaphy.2006.12.004

Qian, R. J., Hu, Q. D., Ma, X. H., Zhang, X. L., Ye, Y. J., Liu, H. J., et al. (2022). Comparative transcriptome analysis of heat stress responses of *Clematis lanuginosa* and *Clematis crassifolia*. *BMC Plant Biol.* 22, 138. doi: 10.1186/s12870-022-03497-w

Raina, S. K., Rane, J., Raskar, N., Singh, A. K., Govindasamy, V., Kumar, M., et al. (2019). Physiological traits reveal potential for identification of drought tolerant mungbean *Vigna radiata* (L.) Wilczek genotypes under moderate soil-moisture deficit. *Indian J. Genet. Plant Breed.* 79, 427–437. doi: 10.31742/IJGPB.79.2.6

Rao, S., Tian, Y., Xia, X., Li, Y., and Chen, J. (2020). Chromosome doubling mediates superior drought tolerance in *Lycium ruthenicum* via abscisic acid signaling. *Hortic. Res.* 7, 40. doi: 10.1038/s41438-020-0260-1

Rapala-Kozik, M., Kowalska, E., and Ostrowska, K. (2008). Modulation of thiamine metabolism in *Zea mays* seedlings under conditions of abiotic stress. *J. Exp. Botany* 59, 4133–4143. doi: 10.1093/jxb/ern253

Rapala-Kozik, M., Wolak, N., Kujda, M., and Banas, A. K. (2012). The upregulation of thiamine (vitamin B-1) biosynthesis in *Arabidopsis thaliana* seedlings under salt and osmotic stress conditions is mediated by abscisic acid at the early stages of this stress response. *BMC Plant Biology* 12, 2. doi: 10.1186/1471-2229-12-2

Rivero, R. M., Mittler, R., Blumwald, E., and Zandalinas, S. I. (2022). Developing climate-resilient crops: improving plant tolerance to stress combination. *Plant J.* 109, 373–389. doi: 10.1111/tpj.15483

Roo, L. D., Salomón, R. L., and Steppe, K. (2020). Woody tissue photosynthesis reduces stem CO2 efflux by half and remains unaffected by drought stress in young Populus tremula trees. *Plant Cell Environ.* 43, 981–991. doi: 10.1111/pce.13711

Saed-Moucheshi, A., Sohrabi, F., Fasihfar, E., Baniasadi, F., Riasat, M., and Mozafari, A. A. (2021). Superoxide dismutase (SOD) as a selection criterion for triticale grain yield under drought stress: a comprehensive study on genomics and expression profiling, bioinformatics, heritability, and phenotypic variability. *BMC Plant Biol.* 21, 148. doi: 10.1186/s12870-021-02919-5

Scafaro, A. P., Haynes, P. A., and Atwell, B. J. (2010). Physiological and molecular changes in *Oryza meridionalis* Ng., a heat-tolerant species of wild rice. *J. Exp. Botany* 61, 191–202. doi: 10.1093/jxb/erp294

Sharma, P., Jha, A. B., Dubey, R. S., and Pessarakli, M. (2012). Reactive oxygen species, oxidative damage, and antioxidative defense mechanism in plants under stressful conditions. *J. Botany* 2012, 217037. doi: 10.1155/2012/217037

Strobbe, S., Verstraete, J., Fitzpatrick, T. B., Faustino, M., Lourenco, T. F., Oliveira, M. M., et al. (2022). A novel panel of yeast assays for the assessment of thiamin and its biosynthetic intermediates in plant tissues. *New Phytologist* 234, 748–763. doi: 10.1111/nph.17974

Strobbe, S., Verstraete, J., Stove, C., and Van Der Straeten, D. (2021). Metabolic engineering provides insight into the regulation of thiamin biosynthesis in plants. *Plant Physiol.* 186, 1832–1847. doi: 10.1093/plphys/kiab198

Sun, J., Sigler, C. L., Beaudoin, G. A. W., Joshi, J., Patterson, J. A., Cho, K. H., et al. (2019). Parts-prospecting for a high-efficiency thiamin thiazole biosynthesis pathway. *Plant Physiol.* 179, 958–968. doi: 10.1104/pp.18.01085

Sun, N., Song, T., Ma, Z., Dong, L., Zhan, L., Xing, Y., et al. (2020). Overexpression of *MsSiR* enhances alkali tolerance in alfalfa (*Medicago sativa* L.) by increasing the glutathione content. *Plant Physiol. Biochem.* 154, 538–546. doi: 10.1016/j.plaphy.2020.07.001

Teixeira, R. D., Paiva, F. L., Dantas, d,.A. J., Mayumi, K. P., Ribeiro, A.F.C., C., et al. (2005). Functional characterization of the thi1 promoter region from *Arabidopsis thaliana*. *J. Exp. Botany* 56, 1797–1804. doi: 10.1093/jxb/eri168

Tietze, F. (1969). Enzymic method for quantitative determination of nanogram amounts of total and oxidized glutathione: applications to mammalian blood and other tissues. *Anal. Biochem.* 27, 502–522. doi: 10.1016/0003-2697(69)90064-5

Wellburn, A. R., and Lichtenthaler, H. K. (1984). *Formulae and Program to Determine Total Carotenoids and Chlorophylls A and B of Leaf Extracts in Different Solvents*. Springer Netherlands.

Xu, Z., and Zhou, G. (2008). Responses of leaf stomatal density to water status and its relationship with photosynthesis in a grass. *J. Exp. Botany* 59, 3317–3325. doi: 10.1093/jxb/ern185

Yan, C., Song, S., Wang, W., Wang, C., Li, H., Wang, F., et al. (2020). Screening diverse soybean genotypes for drought tolerance by membership function value based on multiple traits and drought-tolerant coefficient of yield. *BMC Plant Biol.* 20, 321–321. doi: 10.1186/s12870-020-02519-9

Yang, M., Duan, X., Wang, Z., Yin, H., Zang, J., Zhu, K., et al. (2021). Overexpression of a Voltage-Dependent Anion-Selective Channel (VDAC) protein-encoding gene, *MsVDAC*, from *Medicago sativa* confers cold and drought tolerance to transgenic tobacco. *Genes* 12, 1706. doi: 10.3390/genes12111706

Yang, Z., Chi, X., Guo, F., Jin, X., Luo, H., Hawar, A., et al. (2020). *SbWRKY30* enhances the drought tolerance of plants and regulates a drought stress-responsive gene, *SbRD19*, in sorghum. *J. Plant Physiol.* 246, 153142. doi: 10.1016/j.jplph.2020.153142

Yu, J. G., Lee, G. H., Lee, S. C., and Park, Y. D. (2014). Gene expression and phenotypic analyses of transgenic Chinese cabbage over-expressing the cold tolerance gene, *BrCSR*. *Hortic. Environ. Biotechnol.* 55, 415–422. doi: 10.1007/s13580-014-0054-1

Zhang, C., and Shi, S. (2018). Physiological and proteomic responses of contrasting Alfalfa (*Medicago sativa* L.) varieties to PEG-induced osmotic stress. *Front. Plant Sci.* 9, 242. doi: 10.3389/fpls.2018.00242

Zhang, H., Liu, D., Yang, B., Liu, W.-Z., Mu, B., Song, H., et al. (2020). Arabidopsis CPK6 positively regulates ABA signaling and drought tolerance through phosphorylating ABA-responsive element-binding factors. *J. Exp. Botany* 71, 188–203. doi: 10.1093/jxb/erz432

Zhang, X., Zhang, L., Chen, Y., Wang, S., Fang, Y., Zhang, X., et al. (2021). Genome-wide identification of the SOD gene family and expression analysis under drought and salt stress in barley. *Plant Growth Regul.* 94, 49–60. doi: 10.1007/s10725-021-00695-8

Zhao, L., Liu, F., Xu, W., Di, C., Zhou, S., Xue, Y., et al. (2009). Increased expression of OsSPX1 enhances cold/subfreezing tolerance in tobacco and Arabidopsis thaliana. *Plant Biotechnol. J.* 7, 550–561. doi: 10.1111/j.1467-7652.2009.00423.x

Zhao, Y., Ma, W., Wei, X., Long, Y., Zhao, Y., Su, M., et al. (2019). Identification of exogenous nitric oxide-responsive miRNAs from Alfalfa (*Medicago sativa* L.) under drought stress by high-throughput sequencing. *Genes* 11, 30. doi: 10.3390/genes11010030

Zhu, J. K. (2016). Abiotic stress signaling and responses in plants. *Cell* 167, 313–324. doi: 10.1016/j.cell.2016.08.029

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 20 September 2022
DOI 10.3389/fpls.2022.985572

OPEN ACCESS

EDITED BY
Wengang Xie,
Lanzhou University, China

REVIEWED BY
Zhihua Wu,
Zhejiang Normal University, China
Dong Luo,
Lanzhou University, China

*CORRESPONDENCE
Xiao-Fei Ma
maxiaofei@lzb.ac.cn

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

RECEIVED 04 July 2022
ACCEPTED 22 August 2022
PUBLISHED 20 September 2022

CITATION
Fang T, Zhou S, Qian C, Yan X, Yin X, Fan X, Zhao P, Liao Y, Shi L, Chang Y and Ma X-F (2022) Integrated metabolomics and transcriptomics insights on flavonoid biosynthesis of a medicinal functional forage, *Agriophyllum squarrosum* (L.), based on a common garden trial covering six ecotypes.
Front. Plant Sci. 13:985572.
doi: 10.3389/fpls.2022.985572

Integrated metabolomics and transcriptomics insights on flavonoid biosynthesis of a medicinal functional forage, *Agriophyllum squarrosum* (L.), based on a common garden trial covering six ecotypes

Tingzhou Fang[1,2], Shanshan Zhou[3], Chaoju Qian[1], Xia Yan[4,5], Xiaoyue Yin[1], Xingke Fan[1], Pengshu Zhao[1,2], Yuqiu Liao[1,2], Liang Shi[1,2], Yuxiao Chang[6] and Xiao-Fei Ma[1,5]*

[1]Key Laboratory of Stress Physiology and Ecology in Cold and Arid Regions, Department of Ecology and Agriculture Research, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou, China, [2]College of Resources and Environment, University of Chinese Academy of Sciences, Beijing, China, [3]Faculty of Environmental Science and Engineering, Shanxi Institute of Science and Technology, Jincheng, China, [4]Key Laboratory of Eco-Hydrology of Inland River Basin, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou, China, [5]Marsgreen Biotech Jiangsu Co., Ltd., Haian, China, [6]Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, China

Agriophyllum squarrosum (L.) Moq., well known as sandrice, is an important wild forage in sandy areas and a promising edible and medicinal resource plant with great domestication potential. Previous studies showed flavonoids are one of the most abundant medicinal ingredients in sandrice, whereby isorhamnetin and isorhamnetin-3-glycoside were the top two flavonols with multiple health benefits. However, the molecular regulatory mechanisms of flavonoids in sandrice remain largely unclear. Based on a common garden trial, in this study, an integrated transcriptomic and flavonoids-targeted metabolomic analysis was performed on the vegetative and reproductive periods of six sandrice ecotypes, whose original habitats covered a variety of environmental factor gradients. Multiple linear stepwise regression analysis unveiled that flavonoid accumulation in sandrice was positively correlated with temperature and UVB and negatively affected by precipitation and sunshine duration, respectively. Weighted co-expression network analysis (WGCNA) indicated the *bHLH* and *MYB* transcription factor (TF) families might play key roles in sandrice flavonoid biosynthesis regulation. A total of 22,778 differentially expressed genes (DEGs) were identified between ecotype DL and ecotype AEX, the two extremes in most environmental factors, whereby 85 DEGs could be related to known flavonoid biosynthesis pathway. A sandrice flavonoid biosynthesis network embracing the detected 23 flavonoids in this research was constructed. Gene families *Plant flavonoid O-methyltransferase*

(*AsPFOMT*) and *UDP-glucuronosyltransferase* (*AsUGT78D2*) were identified and characterized on the transcriptional level and believed to be synthases of isorhamnetin and isorhamnetin-3-glycoside in sandrice, respectively. A trade-off between biosynthesis of rutin and isorhamnetin was found in the DL ecotype, which might be due to the metabolic flux redirection when facing environmental changes. This research provides valuable information for understanding flavonoid biosynthesis in sandrice at the molecular level and laid the foundation for precise development and utilization of this functional resource forage.

KEYWORDS

***Agriophyllum squarrosum*, sandrice, flavonoids, transcriptome, metabolome, common garden, isorhamnetin-3-glycoside, isorhamnetin**

Introduction

Plant secondary metabolites play important roles in mediating plant responses to environmental factors (Yang et al., 2012). As the largest group of secondary metabolites in plants, flavonoids are regarded as one of the key adaptations to additional abiotic stresses of early land plants during the transition toward a non-aquatic lifestyle in land colonization (Albert et al., 2018; Wen et al., 2020). Though its mechanism of anti-oxidative activity remains unclear to some extent, flavonoids were reported as defense compounds in plants (Hernández et al., 2009). Excessive accumulation of flavonoids provides extensive stress tolerance *via* eliminating reactive oxygen species (ROS), and responses to biotic stresses (like pathogen infection and insect predation) as well as nearly all kinds of abiotic stresses, including saline-alkali stress, dehydration, heat, cold, high light intensity, UV radiation, etc. (Nakabayashi et al., 2014; Nakabayashi and Saito, 2015; Ma et al., 2020; Böttner et al., 2021; Dong and Lin, 2021; Xu et al., 2022). On the other hand, the past few decades witnessed growing evidence for the multiple benefits of flavonoids on human wellbeing as plant-derived nutraceuticals. Flavonoids were proved to be anti-inflammatory agents and anti-oxidants, and therefore a therapeutic intervention for a variety of chronic diseases and metabolic syndromes (Galleano et al., 2012; Yi, 2018), including cancer (Le Marchand, 2002; Burkard et al., 2017), cardiovascular diseases (Hodgson and Croft, 2010; Fusi et al., 2020), diabetes (Hussain et al., 2020; Proença et al., 2022), obesity (Rufino et al., 2021), respiratory system diseases (Somerville et al., 2016), and neurodegenerative diseases, e.g., Alzheimer's and Parkinson's diseases (Solanki et al., 2015; Spagnuolo et al., 2018; Tahir et al., 2021).

Flavonoids have been conservatively estimated to comprise over 8,000 metabolites and could be categorized into six subclasses according to different substituting positions on the common diphenylpropane (C6-C3-C6) backbone, e.g., flavones, flavonols, flavanones, flavanols, anthocyanidins, and isoflavones (Dong and Lin, 2021). Flavonoid biosynthesis in plants begins with the general phenylpropanoid pathway (GPP), a three-step enzymatic reaction catalyzed by phenylalanine ammonia lyase (PAL), cinnamate 4-hydroxylase (C4H), and p-coumaroyl coenzyme A ligase (4CL), generating p-coumaroyl-CoA (Vogt, 2010). The two main outlets of the phenylpropanoid pathway, lignin, and flavonoid, branch off from here and are catalyzed by Hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase (HCT) and chalcone synthase (CHS), respectively. CHS catalyzes the first step in sandrice flavonoid biosynthesis to synthesize chalcone and directs the metabolic flux to flavonoid metabolism. Next, chalcone isomerase (CHI) acts on chalcones and generates flavanones, for example, naringenin. Then, a series of successive enzymatic reactions synthesize different classes of flavonoid compounds, whereby methylation and glycosylation are two prevalent processes (Shelton et al., 2012; Wang S. et al., 2019).

Plants are the only natural resource of flavonoids (Hernández et al., 2009). Due to their resilience to local marginal environments, remarkable biodiversity, and multiple dietary benefits, pseudocereals have greater potential as sources of natural flavonoid exploitation and utilization. Sandrice [*Agriophyllum squarrosum* (L.)] is an edible and medicinal resource psammophyte of Chenopodiaceae, which is widely distributed in the vast arid and semi-arid sandy lands in Central Asia, the Caucasus, Mongolia, Siberia, and Northern China (Qian et al., 2016). Sandrice is a pioneer species on mobile sands, being able to adapt to the cruel environment in deserts like extreme temperature, dehydration, and sand burial, and was proved to have a wide range of genetic and phenotypic variation among different populations, which was derived from local adaptation to their original marginal habitats (Yin et al., 2016; Zhao et al., 2017, 2022; Qian et al., 2021). Although this underutilized feed crop has not been domesticated, the seeds of sandrice have a long consumption history and are rich in

essential amino acids, crude fiber, polyunsaturated fatty acids, etc., and tender stems and leaves of sandrice are wild feedings for livestock in the sandy areas, making it an ideal functional food and natural feed crop (Li et al., 1992; Chen et al., 2014; Wang Q. et al., 2019; Han et al., 2021; Zhao et al., 2021). Additionally, the aerial part of sandrice was documented as a medicinal plant for kidney inflammation, dyspepsia, antipyretic, and analgesic treatment in Traditional Chinese Medicine (TCM) as well as in Mongolian medicine. Recent research on sandrice not only elucidated its potential as a feed additive for improving the growth performance of livestock (Liang et al., 2021), but also clarified the efficacy of human diabetes treatment and liver protection (Saqier et al., 2019; Bao et al., 2020). Moreover, our previous studies provided evidence of the abundance and diversity of multiple active pharmaceutical ingredients in sandrice, such as terpenoids, phenolic acids, alkaloids, and especially flavonoids (Yin et al., 2018; Zhou et al., 2021a,b). Thus, the plant is of great domestication potential and capacity of offering dietary and medical benefits to human health, in the context of global climate change and continuous consumption expansion.

Our latest research showed that isorhamnetin is the most abundant flavonoid among a variety of ecotypes of sandrice (Zhou et al., 2021a). Isorhamnetin, and its derivate, isorhamnetin-3-glycoside, were proved to have multiple health benefits, including anti-virus (Urzainqui and Carrasco, 1989), cardiovascular and cerebrovascular protection (Gong et al., 2020), anticancer (Saud et al., 2013; Hu et al., 2019; Park et al., 2019), anti-inflammation (Yang et al., 2013; Xu et al., 2022), detoxification (Kim et al., 2021), and could be a prescription of diabetes (Matboli et al., 2021; Kalai et al., 2022). Together with a few other flavonoids, isorhamnetin was recommended as a dietary supplement to enhance immunity and relieve inflammation of the respiratory system, in the context of the COVID-19 global pandemic (Li et al., 2021). Isorhamnetin, or 3-methylquercetin, is formed *via* the methylation of quercitrin at the 3′-OH and catalyzed by S-adenosyl-L-methionine (SAM) dependent O-methyltransferase family (OMTs, Xie et al., 2022). Plant OMTs can be divided into families PlOMTI (caffeoyl-CoA O-methyltransferase, CCoAOMT) and PlOMTII (caffeic acid 3-O-methyltransferase, COMT) according to their molecular weights and Mg^{2+}-dependency, or categorized into five subfamilies based on substrate preference (Liu et al., 2019). Moreover, three novel OMTs that had high similarity with CCoAOMT, while presented substrate preference in flavonoids was identified in *Mesembryanthemum crystallinum*, and formed a subclade in the CCoAOMT family and designated as PFOMT or CCoAOMT-like (Ibdah et al., 2003). A few studies provided experimental evidence of members in the PFOMT subclade being capable of specifically catalyzing the methylation of quercetin to form isorhamnetin in *Mesembryanthemum crystallinum* (Ibdah et al., 2003), *Serratula tinctoria* (Huang et al., 2004), *Glycine max* (Kim et al., 2006), *Citrus reticulata* (Liu et al., 2020), *Prunus persica* (Xie et al., 2022), etc. However, there is currently no report available on systematic analyses of PFOMT in sandrice or its closely related species, which hinders the precise development of this valuable resource plant.

In our previous studies, differentially accumulation of multiple flavonoids was detected between high- and middle-altitudinal sandrice ecotypes *in situ* (data unpublished). After transplanting to a common garden trial in a middle-altitudinal location, flavonoids were diversified and enriched among 14 ecotypes of sandrice along altitudinal gradients, leading us to the deduction that the enrichment of flavonoids in the high-altitude populations of sandrice was a consequence of physiological response to environmental stresses other than of genetic differentiation involved in local adaptation to the high altitude (Zhou et al., 2021a). To provide further evidence to this hypothesis on the transcriptional level, and to elucidate the underlying molecular basis and regulatory modules of characteristic flavonoid biosynthesis, especially that of isorhamnetin and isorhamnetin-3-glycoside in sandrice, here, based on a common garden trial, we expanded our previous research by integrated analysis combined metabolomics and transcriptomics of six sandrice ecotypes, whose original habitats covering multiple environmental factors gradients, across two developmental stages. Multivariance regression analysis provides solid evidence on the question of whether and how sandrice flavonoid accumulation involves in local adaptation to their original habitats, as a complement and revise to univariate linear regression. Weighted co-expression network analysis (WGCNA) and subsequent hub gene digging were conducted to discover possible regulatory molecules, e.g., transcription factors (TF), regarding flavonoid biosynthesis. Besides, one of our aims was to identify synthases of isorhamnetin and isorhamnetin-3-glycoside in sandrice, for their abundance and potential health benefits. In brief, this research provided preliminary but valuable support for flavonoid biosynthesis in sandrice, and pave the way for the precise development of this resource plant in the health industry.

Materials and methods

Plant growth, common garden trial, and sample collection

Seeds of six sandrice ecotypes, AEX, DL, JST, NM, TGX, and XSW, were collected originally from their habitats, which cover multiple environmental factors gradients in Northern China deserts and sandy lands (**Figure 1A** and **Supplementary Table 1**). Seeds were planted in a common garden trial conducted in natural environmental conditions nearby the City of Wuwei, Gansu Province (Northwest China, Tengger Desert, 37°54′10.98"N, 102°54′4.2"E, 1,530 m) in

April 2019. Mature leaf samples from all the six ecotypes were collected in June (vegetative growth stage of sandrice) and August (reproductive growth stage of sandrice), and were assigned as V and R, respectively. Pooled leaves from the same developmental stage of at least two or more individuals of the six ecotypes were immediately frozen in liquid nitrogen, preparing for ribonucleic acid (RNA) extraction.

Ribonucleic acid extraction and RNA-seq library construction

RNA sequencing was performed to understand the molecular mechanism of flavonoid biosynthesis in sandrice. Total RNA was extracted with the RNAprep Pure Plant Plus kit (TIANGEN, Beijing, China) and treated with RNase-free DNase I (TIANGEN, Beijing, China), following the manufacturer's guidebooks. RNA concentration and purity were measured using NanoDrop 2000 (Thermo Fisher Scientific, Wilmington, DE). RNA integrity was assessed using the RNA Nano 6000 Assay kit of the Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, United States). Twelve RNA sequencing samples were generated using NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina® (NEB, Ipswich, MA, United States) following the manufacturer's recommendations. Library quality was assessed by the Agilent Bioanalyzer 2100 system. After cluster generation, the library preparations were sequenced on an Illumina Hiseq Xten platform (Illumina Inc., San Diego, CA, United States) and paired-end reads were generated. The raw sequencing files of transcriptomic data will be available in the NCBI SRA database with accession number PRJNA853545.

Quality control, assembly, and gene functional annotation

Clean data were obtained by removing reads containing adapter, poly-N, and low-quality paired-end raw reads. All clean data with high quality were subsequently mapped to sandrice reference genome (unpublished data) and formed a joint transcript using HISAT2-Stringtie pipeline with default parameter. To provide insight into the functions of newly identified sandrice transcripts, the annotation for all unigenes was performed by running BLAST against the following databases: Nt (NCBI non-redundant nucleotide sequences); Swiss-Prot (A manually annotated non-redundant protein sequence database); InterProt (The InterPro protein families and domains database); TrEMBL (Atlas of Protein Sequence and Structure); GO (Gene Ontology database); and TAIR (A resource for integrated Arabidopsis data) with an *E*-value threshold of 1e-5 as significant hits.

Multivariance stepwise regression analysis

Multivariance stepwise regression analysis was conducted by IBM SPSS Statistic 25 (SPSS Inc., Armonk, NY, United States) to determine the main environmental factors affecting flavonoid accumulation in sandrice. The 22 environmental climatic factors from the original habitats of the six ecotypes (**Supplementary Tables 1**, **2**) were extracted by ArcGIS 10.2 (ESRI Inc., California, United States) and DIVA-GIS 7.5.0.

Metabolic profiling and differentially accumulated flavonoids analysis

Details of flavonoid-targeted metabolic profiling were listed in our previous research (Zhou et al., 2021a). OPLS-DA analyses for differentially accumulated flavonoids (DAF) identification between different developmental stages within an ecotype and across ecotypes were performed using SIMCA 13.0 (Umetrics, Umeå, Sweden) and OmicShare tools.[1]

Identification and biological function analysis of differentially expressed genes

Differential expression analysis between developmental stages and among ecotypes was performed using the EdgeR package (3.32.1). Genes met an adjusted *P*-value < 0.05 and the absolute value of $\log_2$(fold change) ≥ 1 were assigned as DEGs. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of DEGs were implemented by the clusterProfiler R package and OmicShare tools, using hypergeometric testing to find GO and KEGG entries that are significantly enriched compared to the in-house genome background. Pathway enrichment analysis of DEGs was performed using the KEGG Automatic Annotation Server (KAAS).[2]

Weighted co-expression network analysis and hub genes digging

Read counts of the whole transcript from all the 12 samples were subjected to R package WGCNA (v3.2.2) to construct gene co-expression networks. All samples were initially clustered to analyze the sample height. Following the application of the scale-free topology criterion, a soft

1 https://www.omicshare.com/tools/
2 http://www.genome.jp/tools/kaas/

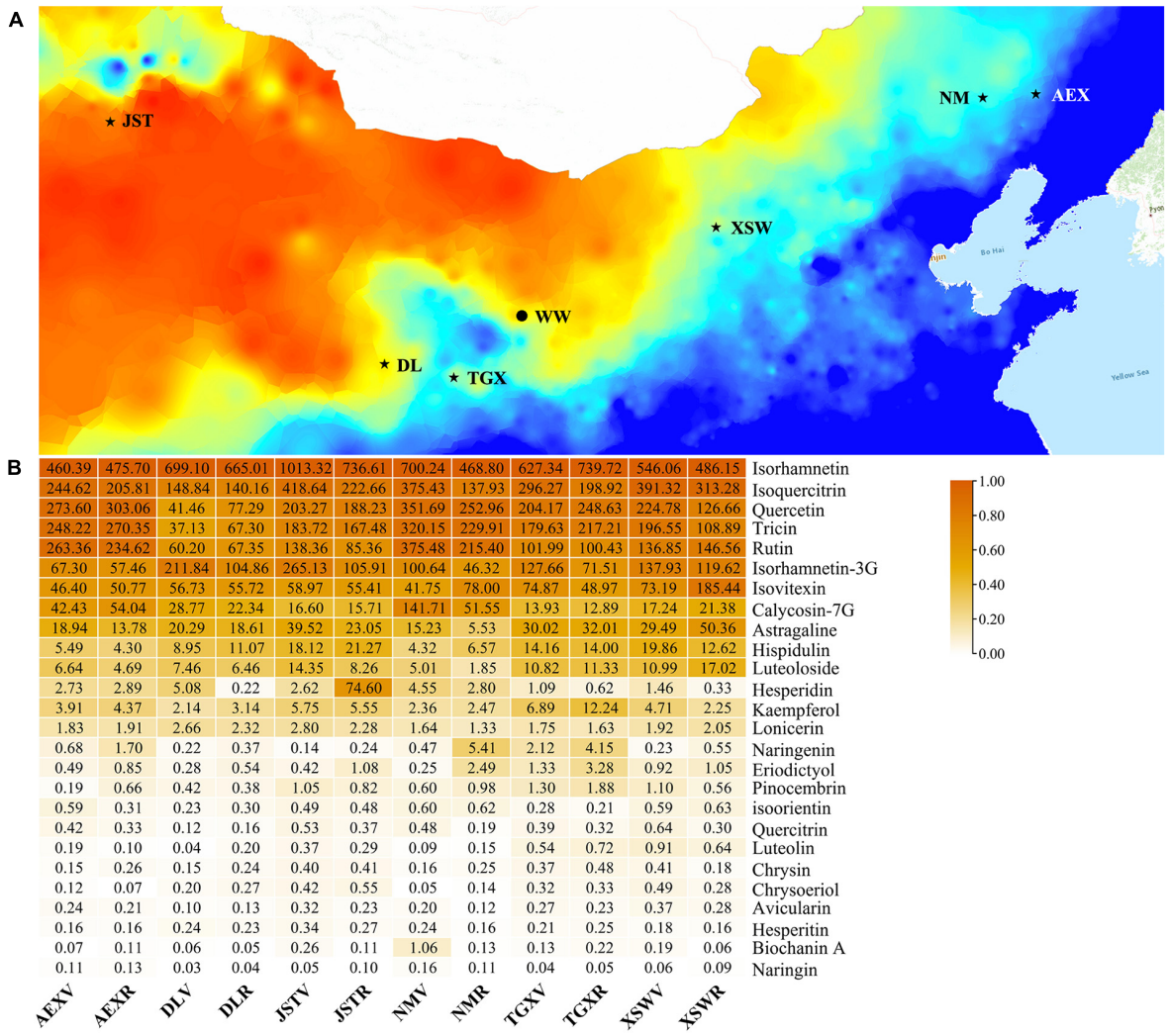


FIGURE 1
Map of samples collected and the contents of flavonoids in the 12 samples. **(A)** Original habitats of the six ecotypes and the location of the common garden trials are marked as stars and dots, respectively. The base map showed the gradient of annual precipitation, from high (blue) to low (red). **(B)** The number in each cell is the average content in μg/g. Modified from Zhou et al. (2021a). Stars are locations of original habitats of each sample.

threshold of 20 was chosen. Based on the topological overlap-based dissimilarity measure, unigenes were first hierarchically clustered, and the gene dendrogram was used for module detection by the dynamic tree cut method with mergeCutHeight = 0.25 and minModuleSize = 50. We calculated the correlation coefficients between the module eigengene (ME) and 26 flavonoid contents. In the WGCNA, gene connectivity was based on the edge weight (ranging from 0 to 1) determined by the topology overlap measure, which reflects the strength of the communication between genes. The weights across all edges of a node were summed and used to define the level of connectivity by Cytoscape plug-in CytoHubba, and flavonoids-related genes in a ME with high connectivity were considered as hub genes. Illustration of genes connectivity within a specific module was performed in Cytoscape 3.9.0.

Whole-transcriptional identification of *AsOMT* and *UDP-glucuronosyltransferase* gene family

First, known plant OMT (Xie et al., 2022) or UGT78D2 protein sequences were subjected to the BLASTP process against the whole transcripts to find possible homologs. Second, an HMMER search for conserved domain structure of PF00891, S-adenosyl-L-methionine (SAM)-dependent methyltransferase,

was conducted against the sandrice protein database on a local server. Union set of the two results above was then submitted to the NCBI web CD-search tool[3] and manually checked to exclude sequences with incomplete conserved domain, resulting in a protein set of candidates AsOMT. For phylogenetic analysis of inter-species OMT tree and intra-species AsOMT and AsUGT78D2 trees, multi-alignment of OMT (or UGT78D2) protein sequences was conducted by MEGA X built-in Muscle program, and a neighborhood-joining (NJ) tree was constructed by MEGA X. Visualization of the phylogenetic trees and gene structure were performed in the iTOL online server[4] and GSDS 2.0 online server,[5] respectively; conserved motifs were found by the help of MEME suite[6] while expression profile heatmaps, conserved domains, and conserved motifs were visualized by TBtools.

Results

Multiple environmental variables affect flavonoid accumulation in sandrice

A total of 26 types of flavonoids in all the 12 tested samples by flavonoids-targeted metabolomics were identified in our latest research, whereby isorhamnetin, isoquercitrin, quercetin, tricin, isorhamnetin-3G, and rutin were the top six flavonoids with the highest contents (average content ≥ 100 μg/g, **Figure 1B**; Zhou et al., 2021a). Various environmental factors co-shaped the unique original habitats. Thus, multivariance stepwise regression provides more robust evidence other than univariate linear regression and was conducted to take a closer aspect of the combined relationship between the 22 environmental climatic factors in the original habitats of the six ecotypes and 14 highly accumulated flavonoids in sandrice (**Table 1**). In general, the amounts of most flavonoids had a positive correlation with temperature and UVB, whereas negatively affected by precipitation and sunshine duration (SD). Among all the regression equations, resolution (adjusted R^2) of rutin (0.736), hispidulin (0.672), quercetin (0.602), and tricin (0.601) were of the highest value with a p-value of 0, indicating the four equations provided better explanations to the relationship between corresponding flavonoid contents and environmental factors. The content of rutin was negatively affected by ALT (altitude) and Tmin (the average minimum temperature in June/August); hispidulin accumulation had a positive correlation with aat10 (accumulated temperature, ≥ 10℃) and Tmin, while was negatively affected by PWM (Precipitation of Wettest Month), MTWEQ (Mean Temperature of Wettest Quarter), and PRE (average precipitation in June/August). Equations in quercetin and tricin were found quite similar. These two flavonoids were both positively affected by PS (precipitation seasonality) and MTWEQ, whereas had negative correlations with SD and TAR (Temperature Annual Range), respectively. Isorhamnetin was negatively affected by PWM, while isoquercitrin had a positive correlation with aat10 and UVB, and a negative correlation with Tmax (the average maximum temperature in June/August), respectively. Though with rather low resolution (0.297), UVB positively affected the accumulation of isorhamnetin-3-glycoside in sandrice, which was in accordance with previous research in peach fruit (Xie et al., 2022).

Biofunction prediction of transcripts and flavonoid biosynthesis-related differentially expressed genes identification

Paired-end RNA sequencing generated 8.67 Gb of raw data. After filtering, 8.63 Gb clean data with an average Q30 value of 91.06 were mapped to the sandrice reference genome (unpublished data), and the average mapping ratio of the 12 samples was about 94%. A total of 69,323 unigenes were generated by the HISAT2-Stringtie pipeline. The mean length of unigenes was 1,292 bp, with N50 of 9,388 bp and GC content of 38%, respectively. The statistics of length distribution showed that 22,917 unigenes ranged from 1,000 bp to 2,000 bp, accounting for 33.06% of the total sequences, while 16,531 unigenes were 2,000–3,000 bp, accounting for 23.85% of the total sequences (**Supplementary Figure 1A**). Annotation of all unigenes generated 38,624, 50,716, 44,967, 55,947, 57,655, 52,039 significant hits against GO, InterPro, SwissProt, Nr, TrEMBL, and TAIR databases, respectively (**Supplementary Figure 1B**). A total of 7,840 uncharacterized unigenes were designated as novel genes in sandrice.

To clarify possible changes during different growth periods and variant gene background of sandrice on the transcriptional level, DEGs were identified among all the six ecotypes across the developmental stage. In general, more upregulated DEGs were identified in comparison between vegetative stage (V) and reproductive stage (R) than downregulated DEGs among most ecotypes, except for TGX and XSW (**Figure 2A**). Venn diagram of DEG_{VvsR} showed that all ecotypes shared 1,834 common DEG, representing the core transcriptional changes between vegetative and reproductive stages (**Supplementary Figure 2A**). Comparative KEGG enrichment of DEG_{VvsR} (**Supplementary Figure 2C**) unveiled that the six ecotypes have experienced the same transcriptional changes in "Starch and sucrose metabolism," "Glycerophospholipid metabolism," "GPI-anchor biosynthesis," and "Terpenoid

3 https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi
4 https://itol.embl.de/
5 http://gsds.gao-lab.org/
6 https://meme-suite.org/meme/

TABLE 1 Multivariance stepwise regression equation of flavonoid contents and environmental factors.

Compounds	Adjusted R^2	Equation of regression	*P*-value	*F*
Total flavonoids	0.116	Y = 2146.964–49.847PWEQ	0.011	7.035
Isorhamnetin	0.205	Y = 896.279–3.057PWM	0.001	12.827
Isoquercitrin	0.399	Y = −1120.472 + 0.024aat10+869.483UVB–26.172Tmax	0.000	9.530
Quercetin	0.602	Y = 2719.683+7.298PS–0.702SD+33.444MTWEQ–42.609TAR	0.000	18.379
Tricin	0.601	Y = 2470.176+6.726PS–0.629SD+31.760MTWEQ–40.106TAR	0.000	18.356
Isorhamnetin3G	0.297	Y = −681.872+479.657UVB	0.000	20.471
Rutin	0.736	Y = 802.331–0.168ALT-31.129Tmin	0.000	65.070
Isovitexin	NA	NA	NA	NA
Calycosin7G	0.242	Y = −127.505+1.594PS	0.000	15.646
Astragalin	0.115	Y = 35.448–0.136PWM	0.011	6.961
Hesperidin	0.217	Y = −4.412+0.005Arid	0.001	13.749
Hispidulin	0.674	Y = 38.108–0.072PWM+0.001aat10–2.5MTWEQ-0.129PRE+1.336Tmin	0.000	20
Luteoloside	0.132	Y = 12.820–0.052PWM	0.007	8.008
Kaempferol	0.396	Y = 33.452+1.997MMTR–0.019SD	0.000	16.066
Lonicerin	0.186	Y = 4.669–0.026PS	0.001	11.481

Flavonoids with an average amount of more than 1 μg/g were taken into consideration. The four equations with resolution (R^2) > 0.5 were presented in bold. PWEQ, Precipitation of Wettest Quarter; PWM, Precipitation of Wettest Month; aat10, accumulated temperature, ≥ 10°C; UVB, UVB radiation; Tmax/min, the average maximum/minimum temperature of June or August; PS, Precipitation Seasonality (Coefficient of Variation); SD, sunshine duration; MTWEQ, Mean Temperature of Wettest Quarter; TAR, Temperature Annual Range; Alt, Altitude; Arid, aridity; PRE, average precipitation of June or August; MMTR, Mean Monthly Temperature Range.

backbone biosynthesis." However, flavonoid biosynthesis was not one of the characteristic differences between vegetative growth and reproductive growth in sandrice, which was in accordance with metabolomic analysis (**Figure 2B**).

Due to its lowest total flavonoid contents among all the ecotypes, we chose ecotype DL as CK to perform DEG identification among the other five ecotypes in the two developmental stages. In the vegetative stage, except for DLVvsJSTV, all the other comparisons had nearly twofold of upregulated DEGs than downregulated ones. However, in the reproductive stage, the comparison of DLRvsAEXR had the most numerous of DEG, indicating these two ecotypes were of great discrepancy in the reproductive stage (**Figure 2A**). As shown in **Supplementary Figures 2B,D**, the majority of DEG was shared among the five comparation in each of the developmental stages, and mainly distributed in KEGG terms "Glycerophospholipid metabolism," "Phosphatidylinositol signaling system," "GPI-anchor biosynthesis," "Terpenoid backbone biosynthesis," and "Ether lipid metabolism" (**Supplementary Figure 2C**). Flavonoid biosynthesis-related DEG was screened out based on their KEGG Orthology (ko00940, phenylpropanoid biosynthesis, and ko00941, flavonoid biosynthesis) in the enrichment result. A total of 24 flavonoids-related DEGs were found, of which five genes, *peroxidase* (*POD*), *4CL*, *COMT*, and *glycosyl hydrolase family protein* (*GHF*), were the most abundant flavonoids-related genes, the copy numbers of eight genes, *CHS*, *flavanone 3-dioxygenase* (*F3H*), *CHI*, *glycosyltransferase* (*GTF*), *Cytochrome P450 98A2* (*CYP98A2*), *feruloyl-CoA 6-hydroxylase* (*F6H*), *CCoAOMT*, and *flavanol synthase* (*FLS*), were of moderate quantity, while the rest 11 genes were scarcely found differentially expressed (**Supplementary Figure 3**). On the other hand, Members in one gene family have diverse expression patterns, for example, three copies of CCoAOMT-coding genes, *AsqAEX008163*, *AsqAEX019070*, and *AsqAEX013633*, were highly expressed in XSWR, NMV, and AEXV, respectively. In general, more highly expressed flavonoid-related genes were found in AEX and NM in the two developmental stages than in other samples, while only a few genes expressed in DL and TGX (**Supplementary Figure 4**), which could be due to the different stress intensity impacted on the variant genetic background of these ecotypes in the common garden trial.

Weighted co-expression network analysis unveiled hub genes in key flavonoids regulation in sandrice

WCGNA integrating transcriptomics and metabolomics was conducted to find the regulation pattern of flavonoids in sandrice. A total of 46,391 (66.9%) unigenes (out of 69,323; after filtering the ones with low expression) were subjected to WGCNA for the characterization of gene networks in flavonoid biosynthesis. Based on their expression profiles, the tested 12 samples were clearly divided in PCA, whereby AEX and NM, the two low-altitude ecotypes, were separated from the other samples along dim1; while the vegetative stage and reproductive stage of each ecotype were divided along dim2 (**Supplementary Figure 5A**). A scale-free topology model with a soft threshold of 20

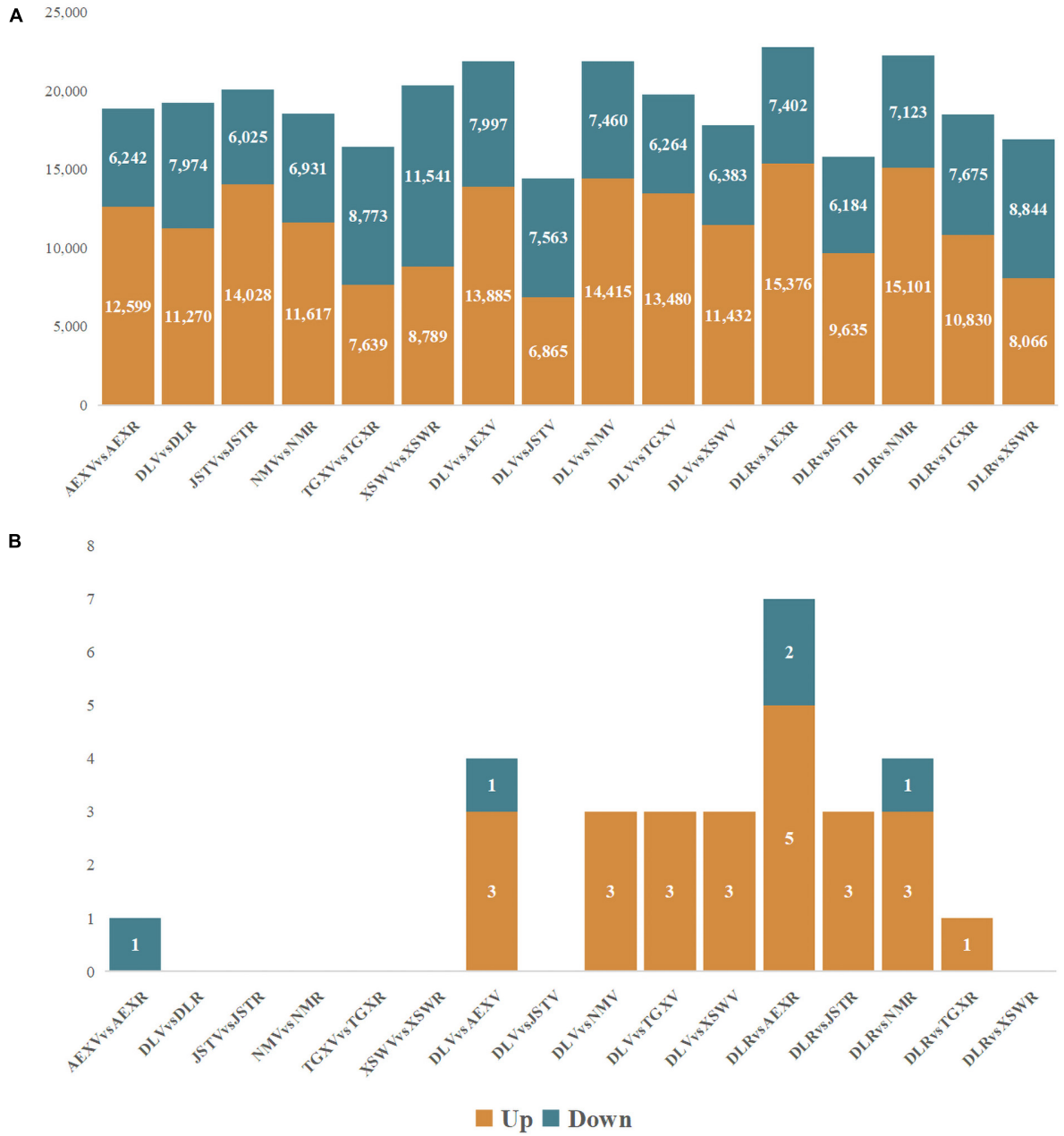


FIGURE 2
Pairwise comparison of DEG **(A)** and DAF **(B)**. Up- and downregulated genes or flavonoids are illustrated in orange and turquoise, respectively.

(Supplementary Figures 5B,C) was used, resulting in 50 module eigengenes (ME) according to gene co-expression patterns (Supplementary Figures 5D,E). Then, an investigation of correlations was conducted between the 50 co-expression modules and the top six flavonoids in sandrice (Figure 3). We identified four MEs that were significantly ($p < 0.05$) correlated with the top six flavonoids in sandrice, which were MEblue with highly positive coefficients associated with quercetin (0.84), tricin (0.85), and rutin (0.89); MEfloralwhite, with isorhamnetin (0.76); MElightgreen, with isorhamnetin-3-glycoside (0.74); and MElightyellow, with isoquercitrin (0.68), respectively. Additionally, quercetin and tricin were significantly negatively correlated with MEskyblue (−0.82 and −0.81, respectively). We also found some high correlations in the module–traits relationship, such as biochanin A (MEpink, 0.96), isovitexin (MEdarkred, 0.96), hesperidin (MEdarkgray, 0.99), naringenin, and eriodictyol (MElightsteelblue1, 0.94 and 0.93, respectively), providing insights into the utilization of

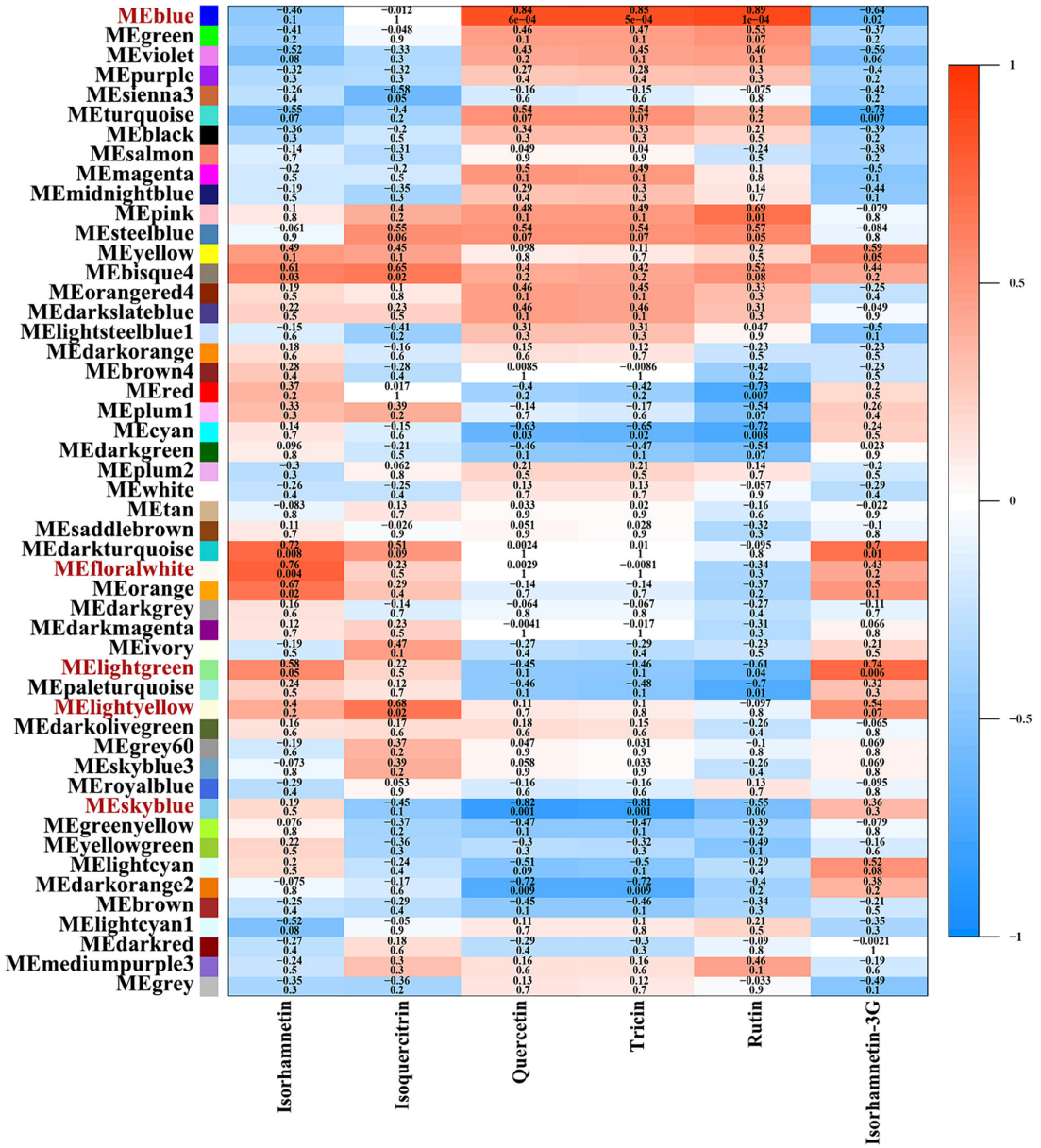


FIGURE 3
Module−trait relationships, partial. This figure shows relationships between 50 module eigengenes and the top six flavonoids. The complete module-trait relationships involving all the 26 detected flavonoids was provided in **Supplementary Figure 6**. The correlation and *p*-value of corresponding flavonoids and gene modules are shown as numbers up and down in each cell, respectively. The color scale on the right side represents the module−trait correlation from −1 (blue) to 1 (red).

the corresponding compounds in the future (**Supplementary Figure 6**).

KEGG enrichment of the five modules with high correlations with the top flavonoids in sandrice mentioned above, i.e., MEblue, MEskyblue, MElightgreen, MElightyellow, and MEfloralwhite, was performed to clarify their potent biofunction (**Supplementary Table 3**). Next, hub genes in MEblue, MEskyblue, and MEfloralwhite were screened out based on their annotation and connectivity in the corresponding module (**Figure 4**). MEblue had a highly positive correlation with quercetin, tricin, rutin, and naringin, and contained eight flavonoid synthases, including POD, CHI, CHS, C4H, cinnamyl alcohol dehydrogenase (CAD), cytochrome P450 81B2 (CYP81B2), and two COMTs. We found a lot of TFs in

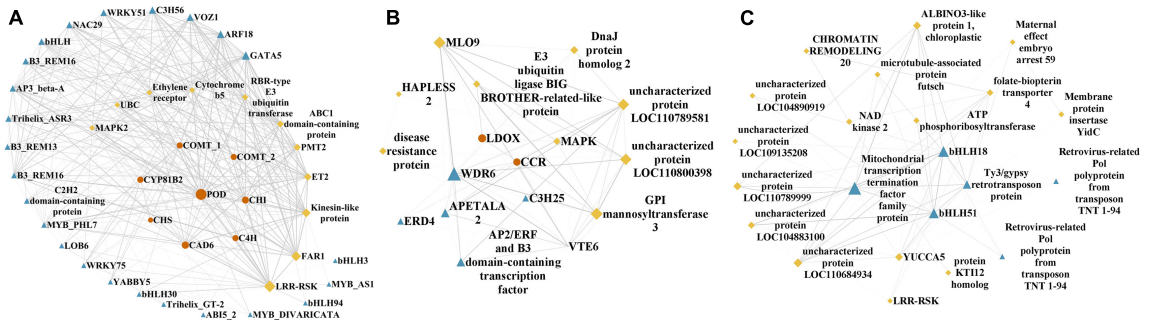


FIGURE 4
Hub genes identified in flavonoids-correlated modules. Hub genes in **(A)** MEblue, **(B)** MEskyblue, and **(C)** MEfloralwhite, respectively. Flavonoid biosynthesis genes, transcription factors or regulators, and other high connectivity genes are shown as orange dots, blue triangle, and yellow diamonds, respectively. Shape size indicated connectivity, while line width indicated the weight of edges.

this module, such as four bHLHs, three MYBs, three B3s, two WRKYs, etc. Interestingly, FAR-RED IMPAIRED RESPONSE 1 (FAR1) was also found in this module, indicating that quercetin, tricin, and rutin might be involved in plants' light signaling transduction and photomorphogenesis (**Figure 4A**; Ma and Li, 2018). MEskyblue was in negative correlation with quercetin and tricin. A member of the WD-domain repeat (WDR) TF family, WDR6, was screened out in MEskyblue (**Figure 4B**). WDR family was documented as a component of the MYB-bHLH-WDR (MBW) complex that regulates the biosynthesis of flavonoids, especially anthocyanin, and leucoanthocyanidin dioxygenase (LDOX) is one of its targets (Xu et al., 2015). Besides, other four TFs, including APETALA2, AP2/ERF, C3H25, and early responsive to dehydration stress 4 (ERD4), were found negatively regulated the accumulation of quercetin and tricin (**Figure 4B**). In MEfloralwhite, the module with a positive correlation with isorhamnetin, two members in the *bHLH* TF family, *bHLH 18* and *bHLH 51*, were identified as hub genes, indicating the two TFs might play key roles in the regulation of isorhamnetin biosynthesis (**Figure 4C**). A few uncharacterized genes and three transposons, two transposons TNT 1–94 and one member in Ty3/Gypsy family, were involved in the regulation of isorhamnetin, whose function requires further study (**Figure 4C**). In MElightgreen and MElightyellow, with a high correlation with isoquercitrin and isorhamnetin-3-glycoside, respectively, no flavonoids-related gene was detected. Details of hub genes in each module are listed in **Supplementary Table 4**.

Identification of putative isorhamnetin and isorhamnetin-3-glycoside biosynthesis in sandrice

Members of the *OMT* gene family were reported in previous studies to code the O-methyltransferases that catalyze

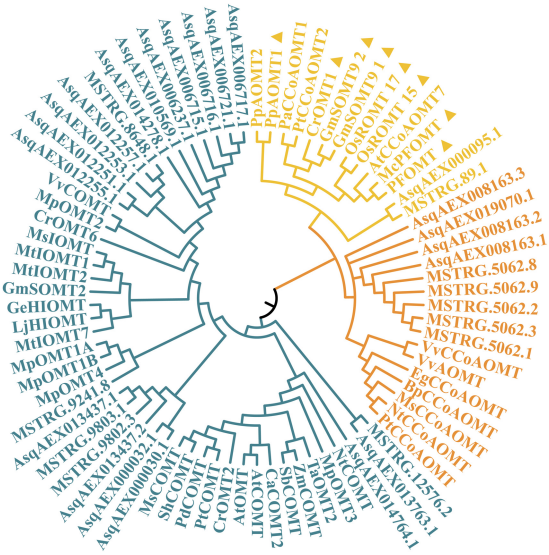


FIGURE 5
Phylogenetic relationship of 33 candidate AsOMT and 45 known OMTs in other plants. CCoAOMT subfamily, PFOMT subclade, and COMT subfamily are colored in orange, yellow, and blue, respectively. Sequences marked with triangles are the reported isorhamnetin biosynthesis genes.

methylation on quercetin at the 3′-OH to form isorhamnetin. As shown in **Figure 5**, a total of 33 sandrice candidate OMTs were identified based on the whole transcripts, including 11 members in the CCoAOMT subfamily containing two members in the PFOMT subclade, and 22 members in the COMT subfamily. OMTs specifically catalyze the biosynthesis of isorhamnetin in other plants, such as GmSOMT9, OsROMT, and McPFOMT, etc., were all located in a subclade PFOMT in the CCoAOMT subfamily (Xie et al., 2022), and clustered with sandrice transcripts *MSTRG89.1* and *AsqAEX000095.1*. Of note, AsOMT shared relatively low similarity with OMTs in other

species and were clustered into solely subclades, indicating OMT in sandrice had experienced early differentiation and possibly subsequent gene duplication events. The intra-species AsOMT phylogenetic tree was divided into two subfamilies, CCoAOMT and COMT, in accordance with the inter-species tree (**Figure 6**). Members in the CCoAOMT subfamily embraced the conserved domain AdoMet_MTases_superfamily, and the COMT subfamily contained a dimerized domain and an AdoMet_MTases domain. In the AsPFOMT subclade, *AsqAEX000095.1* had low expression in the 12 samples. Thus, *MSTRG89.1* might be the candidate PFOMT-coding gene in isorhamnetin biosynthesis in sandrice.

On the other hand, isorhamnetin-3-glycoside is formed by glycosylation on isorhamnetin at 3-OH. Gene *UGT78D2* was speculated to encode glycosyltransferases and act on isorhamnetin to generate isorhamnetin-3-glycoside. To figure out the isorhamnetin-3-glycoside biosynthesis route in sandrice, a total of 14 candidates sandrice *UGT78D2* genes were identified, among which 12 genes were expressed effectively in the 12 samples (**Supplementary Figure 7**). Some of them present ecotype- and/or developmental stage-specific expression modes, such as *AsqAEX021169.1* and *AsqAEX021169.2*. These two genes had almost identical gene structures and conserved motifs but expressed complementarily in the 12 samples. *MSTRG.19033.1* expressed specifically in low-altitude ecotypes, AEX and NM, indicating the accumulation of isorhamnetin-3-glycoside might be involved in stress resistance (**Supplementary Figure 7**).

The molecular basis of flavonoid biosynthesis in sandrice

OPLS-DA analysis between all the six ecotypes across the two developmental stages unveiled that there was no obvious change in flavonoid contents during sandrice growth; whereas, the reproductive stage of DL (DLR) and AEX (AEXR) had the most discrepancy in flavonoid accumulation (**Figure 2B**), of which contents of total flavonoids, quercetin, tricin, rutin, and isoquercitrin were higher in AEXR than DLR, while accumulation of isorhamnetin and isorhamnetin-3-glycoside were higher in DLR than in AEXR, significantly (**Table 2** and **Supplementary Figure 8**).

To clarify the molecular basis underlying metabolomic differences between AEXR and DLR, biofunction prediction of $DEG_{DLRvsAEXR}$ was conducted based on transcriptomics. Compared to DLR, there were 15,376 and 7,402 up- and downregulated DEGs identified in AEXR, respectively (**Figure 2A**). GO enrichment showed that the top five enriched terms were cytoplasm (GO:0005737), carbohydrate binding (GO:0005529), ABC-type transporter activity (GO:0140359), RNA processing (GO:0006396), and kinase activity (GO:0016301; **Supplementary Figure 9A**), while the KEGG enrichment analysis categorized these genes mainly into Biosynthesis of secondary metabolites (ko01110), carbon metabolism (ko01200), starch and sucrose metabolism (ko00500), biosynthesis of amino acids (ko01230), phenylpropanoid biosynthesis (ko00940) in class metabolism, spliceosome (ko03040) in class Information Processing, and plant hormone signal transduction (ko04075) in class Environmental Information Processing (**Supplementary Figure 9B**).

A flavonoid biosynthesis network in sandrice containing the detected 23 compounds was then constructed (**Figure 7**). The pathway of sandrice flavonoid biosynthesis starts with the transformation of phenylalanine to p-coumaroyl-CoA. The two main outlets of the GPP, lignin biosynthesis, and flavonoid biosynthesis, were shunted here, *via* generating naringenin and a series of CoA, respectively (Dong and Lin, 2021). The majority of the flavonoids-related $DEG_{DLRvsAEXR}$ (51 out of 85) was identified in the lignin biosynthesis pathway, including three copies of *4CL*, two *C4H*s, four *Cytochrome P450 98A2*s (*C3′H*), three *shikimate O-hydroxycinnamoyltransferases* (*HCT*s), seven *COMT*s, three *CCoAOMT*s, two *F6H*s, one *cytochrome P450 84A1* (*F5H*), five *CAD*s, and 21 *POD*s. Most of them had higher expression in AEXR than DLR, indicating lignin biosynthesis was activated in ecotype AEX, and possibly resulted from stresses (Meng et al., 2022). On the other hand, naringenin is the intermediate product of biosynthesis pathways of different classes of flavonoids (Zhao et al., 2020), from which 21 out of all the 23 detected flavonoids, except pinocembrin and chrysin, originated. However, only a few enzyme-coding genes were found in the flavonoid biosynthesis pathway in sandrice, i.e., two copies of *CHI*, one *CHS*, three *FLS*s, two *flavonol-3-O-glucoside L-rhamnosyltransferases* (*FG2*), one putative *OMT*, and five putative *UGT78D2*s. Lignin biosynthesis is one of the most well-studied pathways and is conserved in higher plants to a large extent, while clarifying flavonoid biosynthesis genes in small species required further endeavors, which might partially explain the reason why more genes were identified in lignin biosynthesis than flavonoids biosynthesis in sandrice. Quercetin is an important intermediate and serves as a tee-valve in sandrice flavonoid biosynthesis, from which three biosynthesis routes generate: quercetin-quercitrin; quercetin-isoquercitrin-rutin; and quercetin-isorhamnetin-isorhamnetin-3-glycoside (**Figure 7**). UGT78D1, UGT78D2, and PFOMT, respectively, catalyze the formation of quercitrin, isoquercitrin, and isorhamnetin. UGT78D2 also acts on isorhamnetin to form isorhamnetin-3-glycoside, while FG2 acts on isoquercitrin to form rutin. Interestingly, according to OPLS-DA, the contents of quercetin, isoquercitrin, and rutin were significantly lower in DLR, while isorhamnetin and isorhamnetin-3-glycoside accumulated higher in DLR (**Table 2**), indicating metabolic flux was somehow redirected to

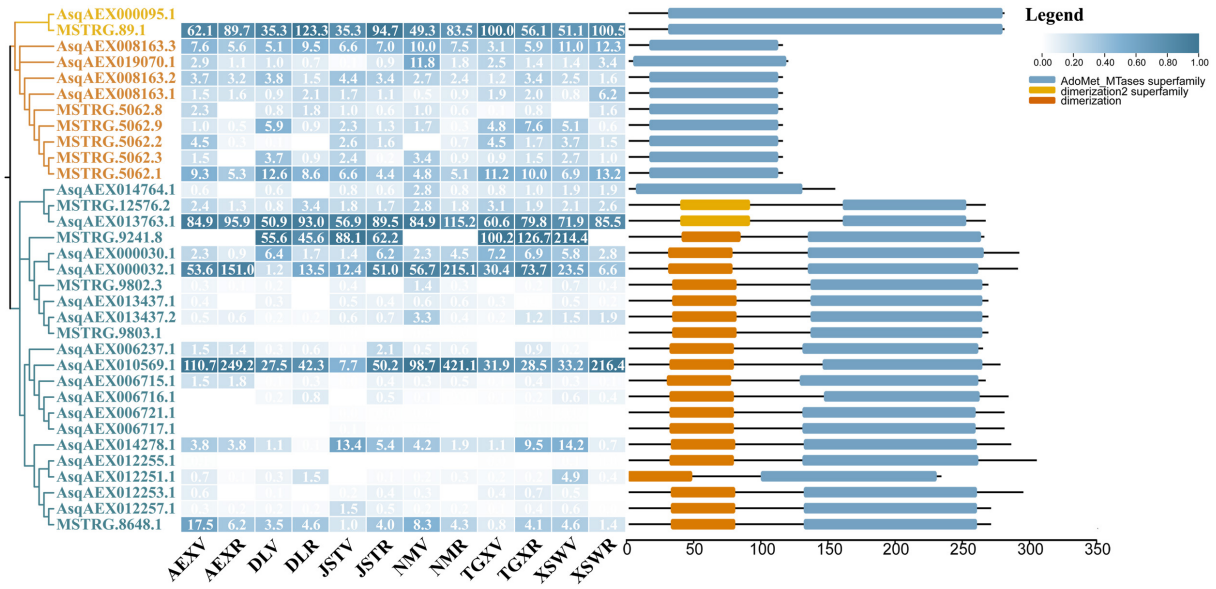


FIGURE 6
AsOMT gene family identification. From left to right, the three columns showed the phylogenetic tree, heatmap of expression profiles (transcripts per million, TPM) in each of the 12 samples, and conserved domain visualization of the 33 candidates AsOMT, respectively. In the phylogenetic tree, the *AsPFOMT* subclade, *AsCCoAOMT* subfamily, and *AsCOMT* subfamily are shaded in yellow, orange, and blue, respectively. Numbers in each cell of the expression profile heatmap are genes' original TPM values. Key conserved domains are shown as different colored bars, according to the actual length of peptides.

TABLE 2 DAF (DLRvsAEXR) identified by OPLS-DA.

Compounds	Log2FC	VIP	*P*-value
Total flavonoids	0.40	2.56	0.02
Isorhamnetin	− 0.68	2.22	0.02
Quercetin	2.39	2.15	0.00
Tricin	2.46	2.04	0.00
Rutin	2.37	1.79	0.00
Isoquercitrin	1.15	1.37	0.00
Isorhamnetin3G	− 1.09	1.06	0.00

VIP, Variable important in projection.

isorhamnetin generating rather than isoquercitrin generating in ecotype DL (**Figure 7**).

Discussion

Flavonoid accumulation in sandrice resulted more from stress-responding than other local adaptation

Flavonoid accumulation can be induced by multiple stresses and is regarded as the consequence of local adaptation in some species. For instance, highland *Zay maize* often accumulates flavonoids in its leaves for coping with the UV-B radiation (Casati and Walbot, 2005). Based on two common garden trials, three flavonoid biosynthesis-related genes, *anthocyanidin synthase* (*ANS*), *anthocyanidin reductase* (*ANR*), and *flavonoid 3'hydroxylase* (*F3′H*) were found outliers among seven natural population of *Pinus yunnanensis* (Sun et al., 2020). In Populus, multiple flavonoid-related genes were subjected to environment selection pressure and the region-specific allele played an important role in local adaptive evolution (Lu et al., 2021). Moreover, in our previous study, *CCoAOMT*, *F3′H*, and *HCT* were under diversifying selection along elevation gradients, indicating flavonoid accumulation was involved in sandrice high-altitude adaptation (data unpublished). Nevertheless, neither selective signal nor gene structure variance in flavonoid-related genes was detected in the present research (data not displayed). Furthermore, in multivariance stepwise regression analysis, most flavonoids had a low resolution in correlation with environmental factors in their original habitats (**Table 1**), indicating that local adaptation was not enough to explain the diversity of flavonoid accumulation among sandrice ecotypes. The common garden trial is a method of detecting local adaptation as well as studying phenotypic plasticity, which provides an equivalent field condition that eliminated environmental heterogeneity in ecotypes' original habitats, while simultaneously creating new stresses (Dorman et al., 2009). Hence, the phenotype observed in a common garden (flavonoid accumulation in this research) was a consequence of genetic background that was

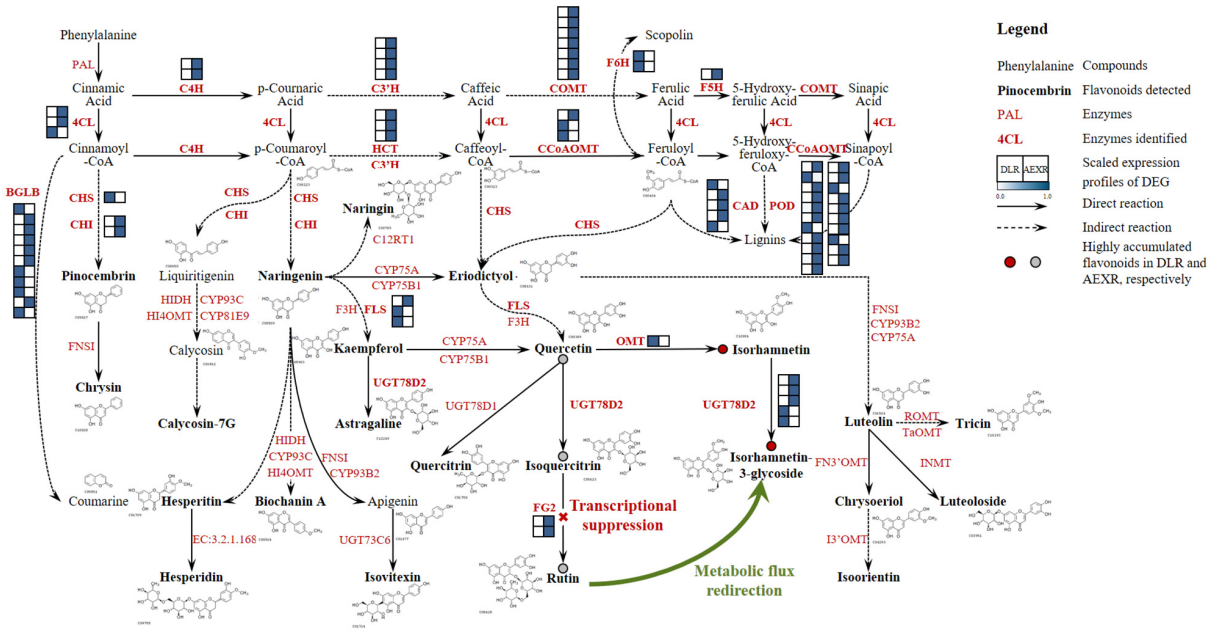


FIGURE 7
Flavonoid biosynthesis pathway in sandrice. Detected flavonoids and $DEG_{DLRvsAEXR}$ in this research are shown in bold fonts, black and red, respectively. The structural formula of the corresponding compound is also presented. DAF between AEXR and DLR are shown as dots, whereby red and gray dots indicate up- and downregulated flavonoids in DLR than in AEXR, respectively. The red cross and green curved arrow indicate the possible mechanism of trade-off between rutin and isorhamnetin biosynthesis. 4CL, 4-coumarate–CoA ligase; BGLB, Beta-glucosidase; C3′H, 5-O-(4-coumaroyl)-D-quinate 3′-monooxygenase; C4H, Trans-cinnamate 4-monooxygenase; CAD, Cinnamyl alcohol dehydrogenase; CCoAOMT, Caffeoyl-CoA O-methyltransferase; CHI, Chalcone-flavanone isomerase; CHS, Chalcone synthase; COMT, Caffeic acid 3-O-methyltransferase; CYP75A, Flavonoid 3′,5′-hydroxylase; CYP75B1, Flavonoid 3′-monooxygenase; CYP93B2, Flavone synthase II; CYP93C, 2-hydroxyisoflavanone synthase; CYP98A2, Cytochrome P450 98A2; F3H, Flavanone 3-dioxygenase; F3′H, Flavonoid 3′hydroxylase; F5H, Cytochrome P450 84A1; F6H, Feruloyl-CoA 6-hydroxylase; FG2, Flavonol-3-O-glucoside L-rhamnosyltransferase; FLS, Flavanol synthase; FN3′OMT, Flavone 3′-O-methyltransferase; FNSI, Flavone synthase I; GHF, Glycosyl hydrolase family protein; GTF, Glycosyltransferase; HCT, Shikimate O-hydroxycinnamoyltransferase; HI4OMT, 2,7,4′-trihydroxyisoflavanone 4′-O-methyltransferase/isoflavone 4′-O-methyltransferase; HIDH, 2-hydroxyisoflavanone dehydratase; I3′OMT, Isoorientin 3′-O-methyltransferase; INMT, Amine N-methyltransferase OMT, SAM-dependent O-methyltransferase; PAL, Phenylalanine ammonia-lyase; POD, Peroxidase; ROMT, Tricin synthase; TaOMT, Tricetin 3′,4′,5′-O-trimethyltransferase; UGT73C6, Flavonol-3-O-L-rhamnoside-7-O-glucosyltransferase; UGT78D1, Flavonol-3-O -rhamnosyltransferase; UGT78D2, Flavonol-3-O-glycosyltransferase; EC:3.2.1.168, Hesperidin 6-O-alpha-L-rhamnosyl-beta-D-glucosidase.

generated from local adaptation (G), response to environmental cues (E), and their combined effects (G × E; Del Valle et al., 2018; Zhou et al., 2021a). WW (common garden location) was of mid-value among a series of environmental factors in the six original habitats. The diversity in expression profiles of flavonoid-related DEGs among different ecotypes (**Supplementary Figure 4**) indicated that low-land ecotypes (AEX and NM) were subjected to intense stress, such as dehydration and intense UVB radiation, while highland ecotypes, e.g., DL and TGX, were in stress relieve or stress release. In a word, flavonoid accumulation in sandrice is more like a consequence of response to environmental factors other than local adaptation to *in situ* habitats. Unfortunately, we were unable to quantify the influences of genetic backgrounds, environmental factors, and the much trickier factor, phenotypic plasticity. In this regard, genetic–environment interactions (G × E) analysis with strict variable control will be our future research direction.

Flavonoid biosynthesis pathway in sandrice

Flavonoids obtain their structural biodiversity from a series of enzymatic substitution reactions, whereby methylation and glycosylation are two prevalent processes (Shelton et al., 2012; Wang S. et al., 2019). In the flavonoid biosynthesis pathway, quercetin serves as an important intermediate and generates isorhamnetin, quercitrin, and isoquercitrin. PFOMT catalyzed methyl transferring on quercetin 3′-OH to form isorhamnetin. It is reported that isorhamnetin rarely significantly accumulates in plants (Huang et al., 2004). However, our latest research showed isorhamnetin was highly enriched in sandrice multiple ecotypes despite different genetic backgrounds, indicating this compound is of great value to sandrice growth and development. In sandrice, homologs of *PFOMT* had either low expression or low correlation with isorhamnetin contents (**Figure 6**). This may be partially due

to the complexity and reticulum of the flavonoid biosynthesis pathway. As an intermediate in the biosynthetic cascade of methylated flavonoids, isorhamnetin accumulation is mainly affected by its direct precursor and downstream compounds (quercetin and isorhamnetin-3-glycoside, respectively), and more importantly, the activity and catalytic efficiency of PFOMT. Therein, more experimental evidence, *in vivo* or *in vitro*, will help clarify the functions and regulatory modules of *AsPFOMT*.

Since flavonoids glycosylation results in an increase in stability (through the protection of reactive nucleophilic groups) and water solubility, natural flavonoids in plants are mostly found in the form of glycosides (Gachon et al., 2005). Two homologous enzymes in the *UGT78* gene family catalyze the first 3-O-glycosylation steps on flavonols, whereby UGT78D1 and UGT78D2 catalyzed rhamnosyl and glucosyl transferring on quercetin 3-OH, respectively (Le Roy et al., 2016), and UGT78D2 also acts on isorhamnetin to form isorhamnetin-3-glycoside (**Figure 7**). In this study, the *AsUGT78D2* gene family embraced 14 genes. Most enzymes in the flavonoid biosynthesis pathway are encoded by multi-genes families rather than a major gene (Vogt, 2010). Though we had little clue about the biological implications of the redundant function of members in a multi-gene family, it can be inferred as strategies in facing variant environment cues. Even though none of the *AsUGT78D2* family members had a significant positive correlation with isorhamnetin-3-glycoside content in sandrice, the 14 homologs were of diverse expression profiles (**Supplementary Figure 7**) and were considered to function redundantly and complementarily in isorhamnetin-3-glycoside biosynthesis, under the circumstance of adverse condition and different developmental stages.

Trade-off between isorhamnetin and rutin might contribute to stress-responding in sandrice

Interestingly, the highland ecotype DL had relatively low contents of quercetin, isoquercitrin, and rutin, but was abundant in isorhamnetin and isorhamnetin-3-glycoside (**Figure 1B**), which was significant when compared to AEX (**Table 2**). Ecotype DL might redirect metabolic flux from route quercetin-isoquercitrin-rutin to route quercetin-isorhamnetin-isorhamnetin-3-glycoside by transcriptional suppression on *FG2* (**Figure 7**). DL locates at the Qinghai-Tibet Plateau, which is of the highest altitude (3,130 m) among the six tested ecotypes, with a combination of ecological factors of low precipitation, low temperature, long SD, and intense UVB radiation (**Supplementary Table 1**). Our previous studies depicted that quercetin was abundant in sandrice highland ecotype, and over-accumulated in lowland ecotype after transplanting into the middle-altitudinal common garden, indicating quercetin can respond to environmental factors involved in high-altitudinal habitats. The character of quercetin has been thoroughly studied. It is reported that the high efficiency in ROS elimination of quercetin rendered plants with prevalent protection against intense UVB radiation, heavy metal, osmotic stress, microbial infection, etc. Moreover, quercetin could induce phytohormone signaling pathways, including auxin (IAA) and abscisic acid (ABA), and regulate the activity of enzymatic and non-enzymatic antioxidants in a dose-dependent manner (Jańczak-Pieniążek et al., 2021; Singh et al., 2021). On the other hand, rutin was documented to be a plant defense compound against cold stress (Reimer et al., 2022), salinity (Lim et al., 2012), and accumulated after UVB radiation (Huang et al., 2016). Limited studies on isorhamnetin function in plants *in vivo* unveiled that isorhamnetin and isorhamnetin-3-glycoside could respond to UVB, while the latter was upregulated by fungus infection (Santin et al., 2018; Xie et al., 2022). According to research on the antioxidant property of quercetin and its derivates, quercetin is believed to have the highest potential, followed by isorhamnetin and isorhamnetin-3-glycoside, whereas in terms of lipid-peroxidation inhibition, isorhamnetin showed higher activities than quercetin (Singh et al., 2021). If isorhamnetin and its derivate can respond to stresses occurred in the common garden, then enrichment of isorhamnetin should have been observed in lowland ecotypes; If quercetin, isoquercitrin, and rutin are responsible for ROS elimination, then a similar decrease in biosynthesis of these three flavonoids should have been monitored in the other highland ecotype TGX. However, no such phenomenon was observed. The reason for the trade-off between rutin and isorhamnetin in sandrice ecotype DL remains unclear, but the unusual high accumulation of isorhamnetin and isorhamnetin-3-glycoside indicated that the ecotype might readjust the homeostasis between quercetin (isoquercitrin or rutin) and isorhamnetin (and isorhamnetin-3-glycoside) to cope with environmental changes.

Conclusion

In this research, we conducted multivariance stepwise regression analysis and elucidated the main environmental factors, i.e., precipitation, temperature, UVB, and sunshine duration, that affect flavonoid accumulation in sandrice. Based on a common garden trial, an integrated analysis of transcriptomics and flavonoids-targeted metabolomics from six sandrice ecotypes in vegetative and reproductive stages provided valuable data in outlining the flavonoid biosynthesis pathway in sandrice. WGCNA and hub gene digging illustrated the molecular regulatory network underlying the biosynthesis of the characteristic flavonoids in sandrice, for instance, isorhamnetin,

quercetin, rutin, isoquercitrin, tricin, and isorhamnetin-3-glycoside, whereby *bHLH* and *MYB* TF family might be the key transcriptional regulatory elements in flavonoid accumulation. Moreover, identification and characterization the of *AsPFOMT* gene family and *AsUGT78D2* gene family unveiled that members in the two families might serve as synthases of isorhamnetin and isorhamnetin-3-glycoside, respectively. Finally, the hypothesis that the trade-off between rutin and isorhamnetin in highland ecotype DL provided insights into sandrice stress-responding mechanisms. Although there are still many gaps to be filled, the present research provided preliminary but valuable data support in elucidating the molecular basis of sandrice flavonoid biosynthesis, paving the way to precise development of this valuable resource forage in the health industry.

Data availability statement

The data presented in this study are deposited in the NCBI SRA repository, accession number: PRJNA853545.

Author contributions

X-FM: conceptualization. TF, SZ, XYY, and YL: data curation. CQ and X-FM: funding acquisition and writing—review and editing. SZ: investigation. CQ, XY, and YC: methodology. TF and XF: software. XF, PZ, and LS: validation. TF and XYY: visualization. TF: writing—original draft. All authors listed have made a substantial, direct, and intellectual contribution to the work, and have read and agreed to the published version of the manuscript.

Funding

This research was funded by the Special Funds for Scientific and Technological Innovation and Development in Gansu Province (Research and Development Center of Desert Nutrition and Medicinal Plant Germplasm Innovation) and the Biological Resources Programme, Chinese Academy of Sciences (Grant no. KFJ-BRP-007-015).

Conflict of interest

Authors XY and X-FM were employed by the Marsgreen Biotech Jiangsu Co., Ltd.

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.985572/full#supplementary-material

References

Albert, N. W., Thrimawithana, A. H., McGhie, T. K., Clayton, W. A., Deroles, S. C., Schwinn, K. E., et al. (2018). Genetic analysis of the liverwort Marchantia polymorpha reveals that R2R3MYB activation of flavonoid production in response to abiotic stress is an ancient character in land plants. *New Phytol.* 218, 554–566. doi: 10.1111/nph.15002

Bao, S., Wu, Y.-L., Wang, X., Han, S., Cho, S., Ao, W., et al. (2020). Agriophyllum oligosaccharides ameliorate hepatic injury in type 2 diabetic db/db mice targeting INS-R/IRS-2/PI3K/AKT/PPAR-γ/Glut4 signal pathway. *J. Ethnopharmacol.* 257:112863. doi: 10.1016/j.jep.2020.112863

Böttner, L., Grabe, V., Gablenz, S., Böhme, N., Appenroth, K. J., Gershenzon, J., et al. (2021). Differential localization of flavonoid glucosides in an aquatic plant implicates different functions under abiotic stress. *Plant Cell Environ.* 44, 900–914. doi: 10.1111/pce.13974

Burkard, M., Leischner, C., Lauer, U. M., Busch, C., Venturelli, S., and Frank, J. (2017). Dietary flavonoids and modulation of natural killer cells: Implications in malignant and viral diseases. *J. Nutr. Biochem.* 46, 1–12. doi: 10.1016/j.jnutbio.2017.01.006

Casati, P., and Walbot, V. (2005). Differential accumulation of maysin and rhamnosylisoorientin in leaves of high-altitude landraces of maize after UV-B exposure. *Plant Cell Environ.* 28, 788–799. doi: 10.1111/j.1365-3040.2005.01329.x

Chen, G., Zhao, J., Zhao, X., Zhao, P., Duan, R., Nevo, E., et al. (2014). A psammophyte Agriophyllum squarrosum (L.) Moq.: A potential food crop. *Genet. Resour. Crop Evol.* 61, 669–676. doi: 10.1007/s10722-014-0083-8

Del Valle, J. C., Buide, M. L., Whittall, J. B., and Narbona, E. (2018). Phenotypic plasticity in light-induced flavonoids varies among tissues in Silene littorea (*Caryophyllaceae*). *Environ. Exp. Bot.* 153, 100–107. doi: 10.1016/j.envexpbot.2018.05.014

Dong, N.-Q., and Lin, H.-X. (2021). Contribution of phenylpropanoid metabolism to plant development and plant–environment interactions. *J. Integr. Plant Biol.* 63, 180–209. doi: 10.1111/jipb.13054

Dorman, M., Sapir, Y., and Volis, S. (2009). Local adaptation in four Iris species tested in a common-garden experiment. *Biol. J. Linn. Soc.* 98, 267–277. doi: 10.1111/j.1095-8312.2009.01265.x

Fusi, F., Trezza, A., Tramaglino, M., Sgaragli, G., Saponara, S., and Spiga, O. (2020). The beneficial health effects of flavonoids on the cardiovascular system: Focus on K+ channels. *Pharmacol. Res.* 152:104625. doi: 10.1016/j.phrs.2019.104625

Gachon, C. M. M., Langlois-Meurinne, M., and Saindrenan, P. (2005). Plant secondary metabolism glycosyltransferases: The emerging functional analysis. *Trends Plant Sci.* 10, 542–549. doi: 10.1016/j.tplants.2005.09.007

Galleano, M., Calabro, V., Prince, P. D., Litterio, M. C., Piotrkowski, B., Vazquez-Prieto, M. A., et al. (2012). Flavonoids and metabolic syndrome. *Ann. N. Y. Acad. Sci.* 1259, 87–94. doi: 10.1111/j.1749-6632.2012.06511.x

Gong, G., Guan, Y.-Y., Zhang, Z.-L., Rahman, K., Wang, S.-J., Zhou, S., et al. (2020). Isorhamnetin: A review of pharmacological effects. *Biomed. Pharmacother.* 128:110301. doi: 10.1016/j.biopha.2020.110301

Han, L., Qiu, S., Cao, S., Yu, Y., Yu, S., and Liu, Y. (2021). Molecular characteristics and physicochemical properties of very small granule starch isolated from Agriophyllum squarrosum seeds. *Carbohydr. Polym.* 273:118583. doi: 10.1016/j.carbpol.2021.118583

Hernández, I., Alegre, L., Van Breusegem, F., and Munné-Bosch, S. (2009). How relevant are flavonoids as antioxidants in plants? *Trends Plant Sci.* 14, 125–132. doi: 10.1016/j.tplants.2008.12.003

Hodgson, J. M., and Croft, K. D. (2010). Tea flavonoids and cardiovascular health. *Mol. Asp. Med.* 31, 495–502. doi: 10.1016/j.mam.2010.09.004

Hu, J., Zhang, Y., Jiang, X., Zhang, H., Gao, Z., Li, Y., et al. (2019). ROS-mediated activation and mitochondrial translocation of CaMKII contributes to Drp1-dependent mitochondrial fission and apoptosis in triple-negative breast cancer cells by isorhamnetin and chloroquine. *J. Exp. Clin. Cancer Res.* 38:225. doi: 10.1186/s13046-019-1201-4

Huang, T.-S., Anzellotti, D., Dedaldechamp, F., and Ibrahim, R. K. (2004). Partial Purification, Kinetic Analysis, and Amino Acid Sequence Information of a Flavonol 3-O-Methyltransferase from Serratula tinctoria. *Plant Physiol.* 134, 1366–1376. doi: 10.1104/pp.103.036442

Huang, X., Yao, J., Zhao, Y., Xie, D., Jiang, X., and Xu, Z. (2016). Efficient Rutin and Quercetin Biosynthesis through Flavonoids-Related Gene Expression in Fagopyrum tataricum Gaertn. Hairy Root Cultures with UV-B Irradiation. *Front. Plant Sci.* 7:63. doi: 10.3389/fpls.2016.00063

Hussain, T., Tan, B., Murtaza, G., Liu, G., Rahu, N., Saleem Kalhoro, M., et al. (2020). Flavonoids and type 2 diabetes: Evidence of efficacy in clinical and animal studies and delivery strategies to enhance their therapeutic efficacy. *Pharmacol. Res.* 152:104629. doi: 10.1016/j.phrs.2020.104629

Ibdah, M., Zhang, X.-H., Schmidt, J., and Vogt, T. (2003). A Novel Mg2+-dependent O-Methyltransferase in the Phenylpropanoid Metabolism of Mesembryanthemum crystallinum. *J. Biol. Chem.* 278, 43961–43972. doi: 10.1074/jbc.M304932200

Jańczak-Pieniążek, M., Migut, D., Piechowiak, T., Buczek, J., and Balawejder, M. (2021). The Effect of Exogenous Application of Quercetin Derivative Solutions on the Course of Physiological and Biochemical Processes in Wheat Seedlings. *Int. J. Mol. Sci.* 22:6882. doi: 10.3390/ijms22136882

Kalai, F. Z., Boulaaba, M., Ferdousi, F., and Isoda, H. (2022). Effects of isorhamnetin on diabetes and its associated complications: A review of *in vitro* and *in vivo* studies and a post hoc transcriptome analysis of involved molecular pathways. *Int. J. Mol. Sci.* 23:704. doi: 10.3390/ijms23020704

Kim, B. G., Lee, H. J., Park, Y., Lim, Y., and Ahn, J.-H. (2006). Characterization of an O-methyltransferase from soybean. *Plant Physiol. Biochem.* 44, 236–241. doi: 10.1016/j.plaphy.2006.05.003

Kim, M., Jee, S.-C., Kim, K.-S., Kim, H.-S., Yu, K.-N., and Sung, J.-S. (2021). Quercetin and Isorhamnetin Attenuate Benzo[a]pyrene-Induced Toxicity by Modulating Detoxification Enzymes through the AhR and NRF2 Signaling Pathways. *Antioxidants* 10:787. doi: 10.3390/antiox10050787

Le Marchand, L. (2002). Cancer preventive effects of flavonoids—a review. *Biomed. Pharmacother.* 56, 296–301. doi: 10.1016/S0753-3322(02)00186-5

Le Roy, J., Huss, B., Creach, A., Hawkins, S., and Neutelings, G. (2016). Glycosylation Is a Major Regulator of Phenylpropanoid Availability and Biological Activity in Plants. *Front. Plant Sci.* 7:735. doi: 10.3389/fpls.2016.00735

Li, S., Chang, X., and Zhao, X. (1992). Study of Agriophyllum Squarrosum—pioneer plant on shifting sand. *J. Arid Land Environ. Resour.* 6, 63–70. doi: 10.13448/j.cnki.jalre.1992.04.009

Li, S., Cheng, C.-S., Zhang, C., Tang, G.-Y., Tan, H.-Y., Chen, H.-Y., et al. (2021). Edible and Herbal Plants for the Prevention and Management of COVID-19. *Front. Pharmacol.* 12:656103. doi: 10.3389/fphar.2021.656103

Liang, Y., Ren, Y., Ran, F., Jiao, D., Du, X., and Ma, X.-F. (2021). Effect of dietary different proportions of Agriophyllumsquarrosumon growth performance and blood indexes of Tan sheep. *Feed Res.* 44, 15–20. doi: 10.13557/j.cnki.issn1002-2813.2021.18.004

Lim, J.-H., Park, K.-J., Kim, B.-K., Jeong, J.-W., and Kim, H.-J. (2012). Effect of salinity stress on phenolic compounds and carotenoids in buckwheat (*Fagopyrum esculentum* M.) sprout. *Food Chem.* 135, 1065–1070. doi: 10.1016/j.foodchem.2012.05.068

Liu, X., Zhao, C., Gong, Q., Wang, Y., Cao, J., Li, X., et al. (2020). Characterization of a caffeoyl-CoA O-methyltransferase-like enzyme involved in biosynthesis of polymethoxylated flavones in Citrus reticulata. *J. Exp. Bot.* 71, 3066–3079. doi: 10.1093/jxb/eraa083

Liu, Z., Fan, M., Li, C., and Xu, J.-H. (2019). Dynamic gene amplification and function diversification of grass-specific O-methyltransferase gene family. *Genomics* 111, 687–695. doi: 10.1016/j.ygeno.2018.04.005

Lu, W., Du, Q., Xiao, L., Lv, C., Quan, M., Li, P., et al. (2021). Multi-omics analysis provides insights into genetic architecture of flavonoid metabolites in Populus. *Ind. Crops Prod.* 168:113612. doi: 10.1016/j.indcrop.2021.113612

Ma, L., and Li, G. (2018). FAR1-RELATED SEQUENCE (FRS) and FRS-RELATED FACTOR (FRF) family proteins in Arabidopsis growth and development. *Front. Plant Sci.* 9:692. doi: 10.3389/fpls.2018.00692

Ma, S., Lv, L., Meng, C., Zhang, C., and Li, Y. (2020). Integrative analysis of the metabolome and transcriptome of Sorghum bicolor reveals dynamic changes in flavonoids accumulation under saline–alkali stress. *J. Agric. Food Chem.* 68, 14781–14789. doi: 10.1021/acs.jafc.0c06249

Matboli, M., Saad, M., Hasanin, A. H. A., Saleh, L., Baher, W., Bekhet, M. M., et al. (2021). New insight into the role of isorhamnetin as a regulator of insulin signaling pathway in type 2 diabetes mellitus rat model: Molecular and computational approach. *Biomed. Pharmacother.* 135:111176. doi: 10.1016/j.biopha.2020.111176

Meng, J., Zhang, Y., Wang, G., Ji, M., Wang, B., He, G., et al. (2022). Conduction of a chemical structure-guided metabolic phenotype analysis method targeting phenylpropane path-way via LC-MS: Ginkgo biloba and soybean as examples. *Food Chem.* 390:133155. doi: 10.1016/j.foodchem.2022.133155

Nakabayashi, R., and Saito, K. (2015). Integrated metabolomics for abiotic stress responses in plants. *Curr. Opin. Plant Biol.* 24, 10–16. doi: 10.1016/j.pbi.2015.01.003

Nakabayashi, R., Yonekura-Sakakibara, K., Urano, K., Suzuki, M., Yamada, Y., Nishizawa, T., et al. (2014). Enhancement of oxidative and drought tolerance in Arabidopsis by overaccumulation of antioxidant flavonoids. *Plant J.* 77, 367–379. doi: 10.1111/tpj.12388

Park, C., Cha, H.-J., Choi, E. O., Lee, H., Hwang-Bo, H., Ji, S. Y., et al. (2019). Isorhamnetin Induces Cell Cycle Arrest and Apoptosis Via Reactive Oxygen Species-Mediated AMP-Activated Protein Kinase Signaling Pathway Activation in Human Bladder Cancer Cells. *Cancers* 11:1494. doi: 10.3390/cancers11101494

Proença, C., Ribeiro, D., Freitas, M., and Fernandes, E. (2022). Flavonoids as potential agents in the management of type 2 diabetes through the modulation of α-amylase and α-glucosidase activity: A review. *Crit. Rev. Food Sci. Nutr.* 62, 3137–3207. doi: 10.1080/10408398.2020.1862755

Qian, C., Yan, X., Fang, T., Yin, X., Zhou, S., Fan, X., et al. (2021). Genomic Adaptive Evolution of Sand Rice (*Agriophyllum squarrosum*) and Its Implications for Desert Ecosystem Restoration. *Front. Genet.* 12:656061. doi: 10.3389/fgene.2021.656061

Qian, C., Yin, H., Shi, Y., Zhao, J., Yin, C., Luo, W., et al. (2016). Population dynamics of Agriophyllum squarrosum, a pioneer annual plant endemic to mobile sand dunes, in response to global climate change. *Sci. Rep.* 6:26613. doi: 10.1038/srep26613

Reimer, J. J., Shaaban, B., Drummen, N., Sanjeev Ambady, S., Genzel, F., Poschet, G., et al. (2022). Capsicum Leaves under Stress: Using Multi-Omics Analysis to Detect Abiotic Stress Network of Secondary Metabolism in Two Species. *Antioxidants* 11:671. doi: 10.3390/antiox11040671

Rufino, A. T., Costa, V. M., Carvalho, F., and Fernandes, E. (2021). Flavonoids as antiobesity agents: A review. *Med. Res. Rev.* 41, 556–585. doi: 10.1002/med.21740

Santin, M., Neugart, S., Castagna, A., Barilari, M., Sarrocco, S., Vannacci, G., et al. (2018). UV-B Pre-treatment Alters Phenolics Response to Monilinia fructicola Infection in a Structure-Dependent Way in Peach Skin. *Front. Plant Sci.* 9:1598. doi: 10.3389/fpls.2018.01598

Saqier, Bao, S., Han, S., and Ao, W. (2019). Effects of Agriophyllum squarrosum extracts on glucose metabolism in KKAy mice and the associated underlying mechanisms. *J. Ethnopharmacol.* 241:112009. doi: 10.1016/j.jep.2019.112009

Saud, S. M., Young, M. R., Jones-Hall, Y. L., Ileva, L., Evbuomwan, M. O., Wise, J., et al. (2013). Chemopreventive activity of plant flavonoid isorhamnetin in colorectal cancer is mediated by oncogenic Src and β-catenin. *Cancer Res.* 73, 5473–5484. doi: 10.1158/0008-5472.CAN-13-0525

Shelton, D., Stranne, M., Mikkelsen, L., Pakseresht, N., Welham, T., Hiraka, H., et al. (2012). Transcription Factors of Lotus: Regulation of Isoflavonoid Biosynthesis Requires Coordinated Changes in Transcription Factor Activity. *Plant Physiol.* 159, 531–547. doi: 10.1104/pp.112.194753

Singh, P., Arif, Y., Bajguz, A., and Hayat, S. (2021). The role of quercetin in plants. *Plant Physiol. Biochem.* 166, 10–19. doi: 10.1016/j.plaphy.2021.05.023

Solanki, I., Parihar, P., Mansuri, M. L., and Parihar, M. S. (2015). Flavonoid-based therapies in the early management of neurodegenerative diseases. *Adv. Nutr.* 6, 64–72. doi: 10.3945/an.114.007500

Somerville, V. S., Braakhuis, A. J., and Hopkins, W. G. (2016). Effect of Flavonoids on Upper Respiratory Tract Infections and Immune Function: A Systematic Review and Meta-Analysis12. *Adv. Nutr.* 7, 488–497. doi: 10.3945/an.115.010538

Spagnuolo, C., Moccia, S., and Russo, G. L. (2018). Anti-inflammatory effects of flavonoids in neurodegenerative disorders. *Eur. J. Med. Chem.* 153, 105–115. doi: 10.1016/j.ejmech.2017.09.001

Sun, Y.-Q., Zhao, W., Xu, C.-Q., Xu, Y., El-Kassaby, Y. A., De La Torre, A. R., et al. (2020). Genetic Variation Related to High Elevation Adaptation Revealed by Common Garden Experiments in Pinus yunnanensis. *Front. Genet.* 10:1405. doi: 10.3389/fgene.2019.01405

Tahir, M. S., Almezgagi, M., Zhang, Y., Bashir, A., Abdullah, H. M., Gamah, M., et al. (2021). Mechanistic new insights of flavonols on neurodegenerative diseases. *Biomed. Pharmacother.* 137:111253. doi: 10.1016/j.biopha.2021.111253

Urzainqui, A., and Carrasco, L. (1989). Degradation of cellular proteins during poliovirus infection: Studies by two-dimensional gel electrophoresis. *J. Virol.* 63, 4729–4735. doi: 10.1128/JVI.63.11.4729-4735.1989

Vogt, T. (2010). Phenylpropanoid biosynthesis. *Mol. Plant* 3, 2–20. doi: 10.1093/mp/ssp106

Wang, Q., Shao, H., Zhang, Z., Yan, S., Huang, F., Zhang, H., et al. (2019). Phenolic profile and antioxidant properties of sand rice (Agriophyllum squarrosum) as affected by cooking and *in vitro* digestion. *J. Sci. Food Agric.* 99, 3871–3878. doi: 10.1002/jsfa.9609

Wang, S., Alseekh, S., Fernie, A. R., and Luo, J. (2019). The Structure and Function of Major Plant Metabolite Modifications. *Mol. Plant* 12, 899–919. doi: 10.1016/j.molp.2019.06.001

Wen, W., Alseekh, S., and Fernie, A. R. (2020). Conservation and diversification of flavonoid metabolism in the plant kingdom. *Curr. Opin. Plant Biol.* 55, 100–108. doi: 10.1016/j.pbi.2020.04.004

Xie, L., Guo, Y., Ren, C., Cao, Y., Li, J., Lin, J., et al. (2022). Unravelling the consecutive glycosylation and methylation of flavonols in peach in response to UV-B irradiation. *Plant Cell Environ.* 45, 2158–2175. doi: 10.1111/pce.14323

Xu, W., Dubos, C., and Lepiniec, L. (2015). Transcriptional control of flavonoid biosynthesis by MYB-bHLH-WDR complexes. *Trends Plant Sci.* 20, 176–185. doi: 10.1016/j.tplants.2014.12.001

Xu, Y., Li, J., Lin, Z., Liang, W., Qin, L., Ding, J., et al. (2022). Isorhamnetin alleviates airway inflammation by regulating the Nrf2/Keap1 pathway in a mouse model of COPD. *Front. Pharmacol.* 24:860362. doi: 10.3389/fphar.2022.860362

Yang, C.-Q., Fang, X., Wu, X.-M., Mao, Y.-B., Wang, L.-J., and Chen, X.-Y. (2012). Transcriptional Regulation of Plant Secondary MetabolismF. *J. Integr. Plant Biol.* 54, 703–712. doi: 10.1111/j.1744-7909.2012.01161.x

Yang, J. H., Kim, S. C., Shin, B. Y., Jin, S. H., Jo, M. J., Jegal, K. H., et al. (2013). O-Methylated flavonol isorhamnetin prevents acute inflammation through blocking of NF-κB activation. *Food Chem. Toxicol.* 59, 362–372. doi: 10.1016/j.fct.2013.05.049

Yi, Y.-S. (2018). Regulatory roles of flavonoids on inflammasome activation during inflammatory responses. *Mol. Nutr. Food Res.* 62:1800147. doi: 10.1002/mnfr.201800147

Yin, C., Qian, C., Chen, G., Yan, X., and Ma, X.-F. (2016). The Influence of Selection of Ecological Differentiation to the Phenotype Polymorphism of Agriophyllum Squarrosum. *J. Desert Res.* 36, 364–373. doi: 10.7522/j.issn.1000-694X.2016.00051

Yin, X., Wang, W., Qian, C., Fan, X., Yan, X., Yan, G., et al. (2018). Analysis of metabolomics in Agriophyllum Squarrosum based on UPLC-MS. *Chin. J. Exp. Tradit. Med. Formulae* 21, 51–56. doi: 10.13422/j.cnki.syfjx.20181503

Zhao, C., Wang, F., Lian, Y., Xiao, H., and Zheng, J. (2020). Biosynthesis of citrus flavonoids and their health effects. *Crit. Rev. Food Sci. Nutr.* 60, 566–583. doi: 10.1080/10408398.2018.1544885

Zhao, P., Li, X., Ran, R., Sun, H., Zhao, J., and Chen, G. (2022). Precipitation and local environment shape the geographic variation of seed size across sand rice (*Agriophyllum squarrosum*) natural populations. *J. Exp. Bot.* [Epub ahead of print]. doi: 10.1093/jxb/erac231

Zhao, P., Li, X., Sun, H., Zhao, X., Wang, X., Ran, R., et al. (2021). Healthy values and de novo domestication of sand rice (*Agriophyllum squarrosum*), a comparative view against *Chenopodium quinoa*. *Crit. Rev. Food Sci. Nutr.* [Epub ahead of print]. doi: 10.1080/10408398.2021.1999202

Zhao, P., Zhang, J., Qian, C., Zhou, Q., Zhao, X., Chen, G., et al. (2017). SNP Discovery and Genetic Variation of Candidate Genes Relevant to Heat Tolerance and Agronomic Traits in Natural Populations of Sand Rice (*Agriophyllum squarrosum*). *Front. Plant Sci.* 8:536. doi: 10.3389/fpls.2017.00536

Zhou, S., Yan, X., Yang, J., Qian, C., Yin, X., Fan, X., et al. (2021a). Variations in Flavonoid Metabolites Along Altitudinal Gradient in a Desert Medicinal Plant Agriophyllum squarrosum. *Front. Plant Sci.* 12:683265. doi: 10.3389/fpls.2021.683265

Zhou, S., Yang, J., Qian, C., Yin, X., Yan, X., Fan, X., et al. (2021b). Organic acid metabolites involved in local adaptation to altitudinal gradient in *Agriophyllum squarrosum*, a desert medicinal plant. *J. Plant Res.* 134, 999–1011. doi: 10.1007/s10265-021-01325-0

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 23 September 2022
DOI 10.3389/fpls.2022.976094

OPEN ACCESS

EDITED BY
Linkai Huang,
Sichuan Agricultural University,
China

REVIEWED BY
Haidong Yan,
University of Georgia,
United States
Parviz Heidari,
Shahrood University of Technology, Iran
Cheng Mingjun,
Sichuan Grass Industry Technology
Research and Promotion Center, China

*CORRESPONDENCE
Zhipeng Liu
lzp@lzu.edu.cn
Wenxian Liu
liuwx@lzu.edu.cn

†These authors have contributed equally to this work

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

RECEIVED 23 June 2022
ACCEPTED 25 August 2022
PUBLISHED 23 September 2022

CITATION
Min X, Wang Q, Wei Z, Liu Z and
Liu W (2022) Full-length transcriptional
analysis reveals the complex relationship of
leaves and roots in responses to cold-
drought combined stress in common
vetch.
Front. Plant Sci. 13:976094.
doi: 10.3389/fpls.2022.976094

Full-length transcriptional analysis reveals the complex relationship of leaves and roots in responses to cold-drought combined stress in common vetch

Xueyang Min[1,2†], Qiuxia Wang[1†], Zhenwu Wei[2], Zhipeng Liu[1*] and Wenxian Liu[1*]

[1]State Key Laboratory of Grassland Agro-Ecosystems, Lanzhou University, Engineering Research Centre of Grassland Industry, Ministry of Education, Western China Technology Innovation Centre for Grassland Industry, College of Pastoral Agriculture Science and Technology, Lanzhou University, Lanzhou, Gansu, China, [2]College of Animal Science and Technology, Yangzhou University, Yangzhou, Jiangsu, China

Plant responses to single or combined abiotic stresses between aboveground and underground parts are complex and require crosstalk signaling pathways. In this study, we explored the transcriptome data of common vetch (*Vicia sativa* L.) subjected to cold and drought stress between leaves and roots *via* meta-analysis to identify the hub abiotic stress-responsive genes. A total of 4,836 and 3,103 differentially expressed genes (DEGs) were identified in the leaves and roots, respectively. Transcriptome analysis results showed that the set of stress-responsive DEGs to concurrent stress is distinct from single stress, indicating a specialized and unique response to combined stresses in common vetch. Gene Ontology (GO) enrichment analyses identified that "Photosystem II," "Defence response," and "Sucrose synthase/metabolic activity" were the most significantly enriched categories in leaves, roots, and both tissues, respectively. The Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis results indicated that "ABC transporters" are the most enriched pathway and that all of the genes were upregulated in roots. Furthermore, 29 co-induced DEGs were identified as hub genes based on the consensus expression profile module of single and co-occurrence stress analysis. In transgenic yeast, the overexpression of three cross-stress tolerance candidate genes increased yeast tolerance to cold-drought combined stress. The elucidation of the combined stress-responsive network in common vetch to better parse the complex regulation of abiotic responses in plants facilitates more adequate legume forage breeding for combined stress tolerance.

KEYWORDS

***Vicia sativa* L., combined abiotic stress, leaves, roots, full-length transcripts**

Introduction

Abiotic stress conditions, including drought, heat, cold, salinity, oxidative stress, and nutrient deficiency, are the primary cause of yield loss and quality reduction worldwide (Hasanuzzaman et al., 2019; Gong et al., 2020). Previous research has shown that the combination of multiple stresses could cause a more significant negative impact on crop productivity than the stress applied individually (Gargallo-Garriga et al., 2015; Correia et al., 2018). Under natural and cultivated field conditions, drought is usually combined with abiotic and biotic stresses. Although these stressors were randomly combined, the combinations of drought with cold stress occur frequently in cold, arid, and semiarid regions. When the frozen up ground tissues are subjected to the sun, their water potential will become higher than that of the surrounding air, and the frozen soil will increase water flow, which results in the leaves' water loss and restricts the water uptake and transport from the underground to the xylem (Suzuki et al., 2014; Zandalinas et al., 2018). Thus, to improve crop yield and quality, it is crucial to understand the mechanisms underlying plant responses to combined stress challenges.

Previous reports have provided a plethora of information on plants exposed to single abiotic stress, while these studies cannot be used to obtain the effects of a combination of stresses on plants (Suzuki et al., 2014; Correia et al., 2018; Mota et al., 2021). Currently, an increasing number of reports have revealed the combination of abiotic stress conditions at physiological and molecular levels (Mittler, 2006; Choudhury et al., 2017). These studies showed that the combined stress could impose a specific set of requirements, and required the tailoring of unique and shared regulatory transcript pathways including photosynthesis, hormone signaling, antioxidant mechanisms, and transcription factors (Mahalingam, 2015; Zhu, 2016). Under natural conditions in northeast and northwest China, maize often suffers from combined low temperature and drought stress at the seedling stage. The authors used transcriptomic and metabolomic analyses to systematically analyze the differences in the phenotype and photosynthetic physiology of maize seedlings under single drought, chilling and combined stress conditions and found that drought stress could significantly alleviate the damage caused by chilling stress under combined stress (Guo et al., 2021). In addition to transcriptomic-based technology, a novel system for evaluating drought-cold tolerance by using a chlorophyll fluorescence model was established in grapevine, and find that the temperature of 50% electrolyte leakage values exhibited good reliability for evaluating the grapevine tolerance to drought-cold stress (Su et al., 2015). Furthermore, the source (leaves) and sink (roots) tissues are two distinct resource pools and are linked through the vascular system. When plants suffer adverse conditions, they rapidly perceive environmental signals and then function in different organs to survive (Werner et al., 2008; Lozano-Elena et al., 2022). Various signaling molecules have been shown to play vital roles in response to abiotic stress in leaves and roots, such as phytohormones, calcium, carbohydrates, reactive oxygen species (ROS), and photosynthetic electrons (Zhang and Sonnewald, 2017; Min et al., 2020b,c). Nevertheless, systematic comparisons among leaves and roots under combined drought and cold stress are still limited. Thus, recognizing the integration of changes under combined stress, and possible intercommunication and specific mechanisms between aboveground and underground tissues, are essential to expanding our knowledge of plant responses to different stress stimuli coordinately in different tissue types.

Common vetch is an economically important annual legume species that is commonly cultivated in semiarid regions as a cool-season forage legume due to its rich nutritional value, low requirement, suitable intercrops for reducing diseases, ability to improve soil properties, and superior abiotic stress tolerance (Larbi et al., 2011; Huang et al., 2017; Kartal et al., 2020). Compared to some model plants, common vetch often has a more complex genetic basis and is likely to have additional mechanisms to adapt to stress stimuli. In this study, we investigated the transcriptional dynamics elicited by drought in combination with cold between leaves and roots *via* a meta-analysis of our previously produced common vetch reference transcriptome database to discover combined stress-responsive machines and core regulatory genes (Min et al., 2020c). We also evaluated the function of the cross-stress genes in transgenic yeast. The present study will facilitate the identification of unique and common regulators between leaf and root tissues and consequently support the improvement of abiotic stress tolerance in legumes.

Materials and methods

Sample collection and data acquisition

Experimental materials, seed pretreatment, and seed germination were performed as previously described (Min et al., 2020b). The control groups were hydroponically cultured with 1/2 MS (half-strength Murashige and Skoog, pH = 5.8) solution under plant incubator conditions (20°C, 16 h/8 h (light/dark) photoperiod, 80 μmol m^{-2} s^{-1} photosynthetically active radiation and 70% relative humidity) for 7 days. To conduct cold-drought combined stress, seven-day-old uniformly growing seedlings were transplanted into 1/2 MS solution containing 20% PEG solution to simulate drought stress while adjusting the temperature of the plant growth incubator to 4°C, the 1/2 MS treated plants at 20°C were used as control. Our previous studies found that short-term cold-drought stress could cause the aboveground of common vetch to wilt, and the malondialdehyde and soluble sugar contents were significantly increased at 24 h under 4°C in leaves (Min et al., 2020b,c). Furthermore, to avoid the influence of different photoperiods on experimental materials, samples from the control and cold-drought combined conditions were grown in parallel and harvested after 24 h of treatments. After 24 h, the third and fourth leaves from the top of the plants and root tips (~1.5 cm) of four seedlings were harvested and pooled together.

Twelve samples [(control and 4°C + 20% PEG for 24 h) × two tissues (leaves and roots) × three replications] were used for sequencing analysis, including 0 and 24 h two-time points. RNA isolation, library construction, and sequencing were performed as previously described (Min et al., 2020b). Finally, bowtie2 is used to map the clean reads from a common vetch Illumina data set to a reference transcript.[1]

Full-length reference transcriptome construct

Firstly, equal amounts of total RNA from leaves, stems, and roots under normal conditions, as well as leaves and roots treated with salt, low temperature, drought, and cold-drought combined stresses for 24 h with three biological replications were combined to construct a reference transcriptome by PacBio sequencing. Then, the raw reads were processed into error-corrected reads of insert (ROIs) using the ToFu pipeline with parameters of prediction accuracy > 0.75, full pass ≥ 0, and sequence length ≥ 300 bp. After that, the full-length, non-chimeric transcripts were determined by searching for the PolyA, and the 5′ and 3′ cDNA primers in ROIs. The Iterative Clustering for Error Correction (ICE) was used to cluster the consensus isoforms and full-length consensus sequences, and each cluster was polished with Quiver to obtain a consensus sequence, and full-length transcripts with post-correction accuracy above 99% were generated. Finally, the data of Illumina RNA-Seq were used to correct the low-quality isoforms, and the CD-HIT program was used to delete redundancy, with identity > 0.99.

Expression profile analysis

Gene expression levels of each sample were calculated by RSEM (Li and Dewey, 2011), and the read count for each transcript was obtained from the mapping results. To obtain the difference in gene expression, the parameters of false discovery rate (FDR) < 0.01 and Log_2 (fold change) > 2 were used to identify differentially expressed genes (DEGs) (0). The expression profile of DEGs was clustered by Short Time-series Expression Miner (STEM)[2] software, with $p \leq 0.05$, the maximum number of model profiles was 15, and the log normalizes the data. The specific primers were designed by Primer-BLAST progress on the NCBI website as shown in Supplementary Table S1 and synthesized by Sangon Biotech (Shanghai China). qRT–PCR was conducted as previously described (Min et al., 2020b). The *Vsactin* (*Unigene 68,614*) gene was used to calculate the relative expression levels following the $2^{-\triangle\triangle CT}$ method.

Functional annotation and network analysis

To obtain the annotation information of the transcript, BLAST (version 2.2.26, $E \leq 10^{-5}$) was used to search against the NR, SwissProt, GO, COG, KOG, Pfam, and KEGG databases (Altschul et al., 1997). The GOseq R package and KOBAS software were used to conduct GO and KEGG enrichment ($p < 0.05$; Xie et al., 2008). PlantTFDB (v5.0) was used to identify transcription factors of common vetch (Jin et al., 2016). The "best hit in *Arabidopsis thaliana*" was checked to obtain the best hits. *Arabidopsis* as a reference to predict protein–protein interactions, protein function, and subcellular localization in STRING (V11.5)[3]. Then, the key nodes of the interaction network were modulated in Cytoscape.[4]

Generating overexpression lines in yeast

The coding regions of *F01.PB9597* (*VsUMAMIT*), *F01.PB3916* (*VsPM19*), and *F01.PB22519* (*VsP5CS2*:) were amplified by primers as shown in Supplementary Table S1. The amplification product is purified by SanPrep Column DNA Gel Extraction Kit (Sangon Biotech, Shanghai). The purified cDNA products were directly cloned into the *pYES2* vector (Invitrogen, Carlsbad, United States) and then transformed into *E. coli* DH5α. The recombinant plasmid was sent to Sangon Biotech for sequencing by M13 universal primers. The following steps were performed according to the protocol described by Kawai et al. (2010). Freeing and osmotic (drought and salt) tolerance were performed in the SC-Ura medium, according to previously described methods (Min et al., 2020a). For simultaneous (cold-drought) stress treatment, yeast cells were grown in an SC-Ura liquid medium containing 30%, at −20°C for 36 h. Finally, the serial dilutions (1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, and 10^{-5}) were spotted onto SC-ura, agar plates, and carried out in triplicate at 30°C for 48 h.

Statistical analysis

Pearson correlation and principal component analysis (PCA) of the expression level among each sample were calculated *via* the R cor.test function.[5] Statistical significance was analyzed through Duncan's multiple range test at a 0.05 level. Venn diagrams, heatmaps, and clusters were generated using TBtools (Chen et al., 2020).

1 http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
2 http://www.cs.cmu.edu/~jernst/stem
3 http://www.string-db.org/
4 https://cytoscape.org/
5 http://www.r-project.org

Results

Illumina-Seq, assembly, and annotation

The 12 cDNA libraries of common vetch treated by cold-drought combined stress conditionals between leaves and roots for 24 h (CDL1, CDL2, CDL3, CDR1, CDR2, CDR3) and normal conditionals (CKL1, CKL2, CKL3, CKR1, CKR2, CKR3). In total, 311.62 million high-quality reads were identified in common vetch. The percentages of average GC content and Q30 were 43.13 and 91.75, respectively. Bowtie2 software was used to map the reads onto the previously produced common vetch full-length transcript database. Overall, the average mapped read, uniquely mapped read, and multi-mapped read percentages were 76.82%, 26.99%, and 73.01%, respectively (Supplementary Table S2). Finally, we predicted a total of 30,727 non-redundant transcripts from a min length of 304 to 14,390 bp, the mean length was 2287.35 bp with an N50 of 3.73 kb (Figure 1A). The open reading frame (ORF) length ranged from 63 to 4,674 bp, with a mean length of 1,188.58 bp (Figures 1B,C). In total, 30,250 (98.45%) transcripts were functionally annotated, and 4,879 (15.88%) were annotated in all eight databases. A total of 521 (1.7%) genes were not annotated (Supplementary Table S3). Next, the BLASTP program was used to search against the NR database, examining the evolutionary relationships among common vetch and other plants. The results indicate a close relationship among Leguminosae plants, only 4% of the transcripts were annotated to plants outside the genus *Medicago* (Figure 1D).

Identification and functional profiles of DEGs

The mapping results were combined with the RSEM method for accurate transcript quantification from RNA-Seq data (Li and Dewey, 2011). The FPKM (fragments per kilobase of transcript per million mapped reads) value was used to indicate the expression abundance of the corresponding transcript. When comparing three biological replicates, Pearson's correlation indicated significant correlations ($p < 0.05$, $R^2 \geq 0.88$). A significant positive relationship was found between normal and cold-drought combined treatments in the same tissues ($p < 0.05$, $R^2 \geq 0.64$) but not in the different tissues ($p > 0.05$, $R^2 < 0.05$). Compared to the two tissues in response to abiotic stress, a more significant positive

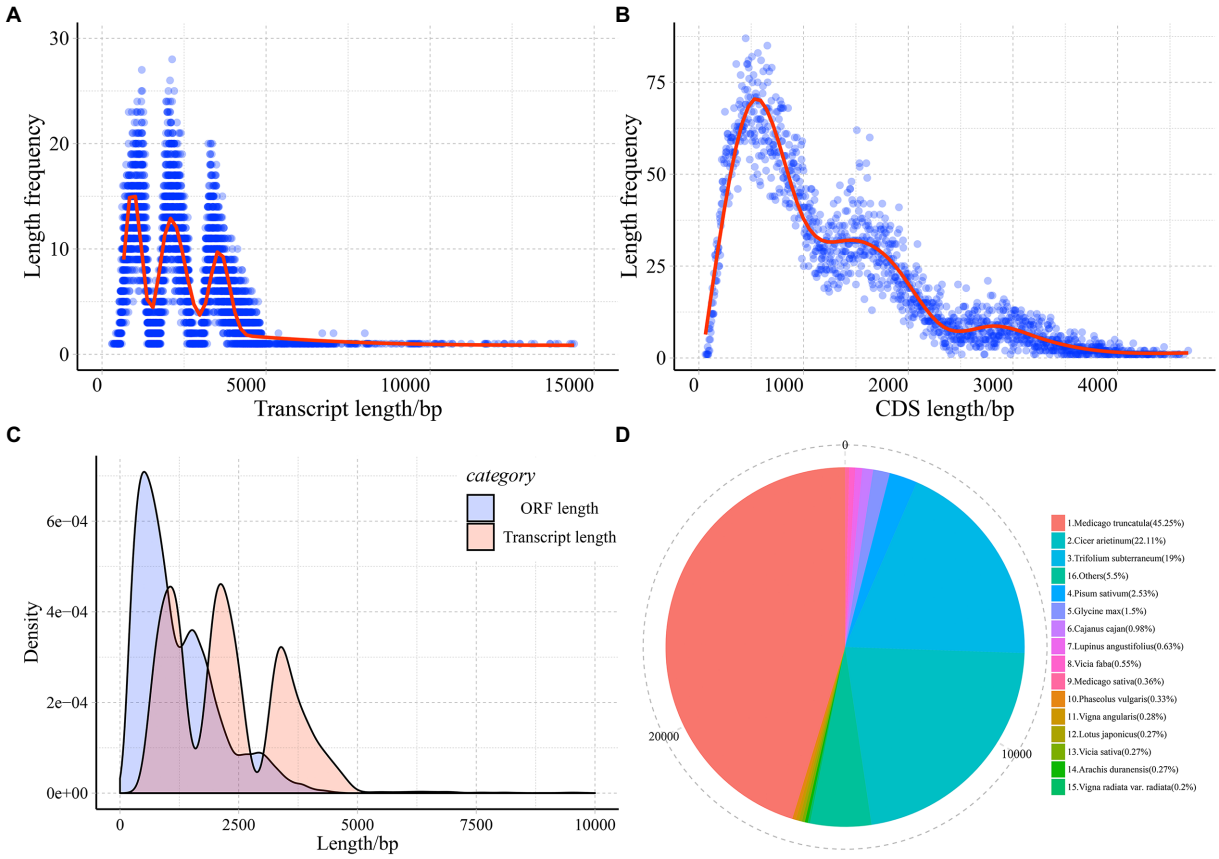


FIGURE 1
Length distribution of the assembled transcript **(A)**, CDS **(B)**, frequency distribution comparison of transcript and CDS **(C)**, and annotation distribution **(D)**.

relationship was shown in roots ($p > 0.05$, R^2 with an average of 0.77), indicating that leaves were more sensitive to abiotic stress (Figure 2A). To understand how common vetch transcripts respond to cold-drought stress between leaves and roots, the mRNA populations were compared with PCA (Figure 2B), and the results showed that the four groups were separated from each other and the CK.

A total of 4,836 and 3,103 DEGs were identified in the leaves and roots, respectively, under cold-drought combined conditions compared with the control. Among these DEGs, 2,835 and 1,755 genes were upregulated in leaves and roots, respectively, while 2,001 and 1,348 DEGs were downregulated in leaves and roots under cold-drought combined stress, respectively (Figures 3A,B). Among them, 3,549 (53.3%) and 1,816 (27.3%) DEGs were found to be leaf and root-specific, respectively (Figures 3C,D). Notably, a total of 1,187 (19.4%) genes were differentially expressed in both tissues; of them, 843 (12.7%) were co-induced, 391 (5.9%) were co-repressed and 53 (0.8%) were oppositely expressed in the leaves and roots (Figures 3C,D). For the 521 novel transcripts, their expression patterns were more likely to be downregulated under cold-drought stress (Supplementary Table S4). Among them, 41 and 18 downregulated DEGs were identified in leaves and roots, respectively. Twenty-five and 3 DEGs were classified as upregulated in leaves and roots, respectively. Furthermore, 12 co-induced and 8 co-repressed DEGs were identified in both leaves and roots (Supplementary Figure S1).

To further confirm the reliability of our RNA-Seq data, 15 candidate DEGs were selected for qRT-PCR validation (Supplementary Figure S2). As shown in Supplementary Figure S3, the expression levels of these DEGs in both leaves and roots were significantly correlated with the FPKM values ($R^2 = 0.8$). Among them, one (*F01.PB19254*) was downregulated, one (*F01.PB13497*) was oppositely expressed, and the remaining genes were upregulated after cold-drought treatment.

Functional enrichment analysis of potential DEGs

Consensus and different co-expression modules identify by GO enrichment analysis

In plants, leaves and roots can respond asymmetrically under abiotic stress, as has been observed at the physiological, morphological, and genetic levels. However, the coordinated variations in the functional genes of different tissues were also observed in common vetch. GO enrichment results showed that 12 and 14 GO enriched terms were specifically divided into leaves and roots, and five GO terms shared by leaves and roots, including "Sucrose synthase activity," "Metabolic process," "Ferric iron binding," "Ferroxidase activity," and "Integral component of membrane" (Figure 4D).

Plant leaves and roots are tightly linked and jointly respond to fluctuating environments. Even though above-and below-ground follow distinct developmental trajectories, some highly coordinated biological processes also exist. GO enrichment (top 25, $p < 0.05$) found that cold-drought combined treatments significantly affected 12 BPs (biological process), 9 MFs (molecular function), and 4 CCs (cellular component) in leaves (Figure 4A). The "Photosystem II" was the most significantly enriched category in leaves, followed by "Protein-chromophore linkage" and

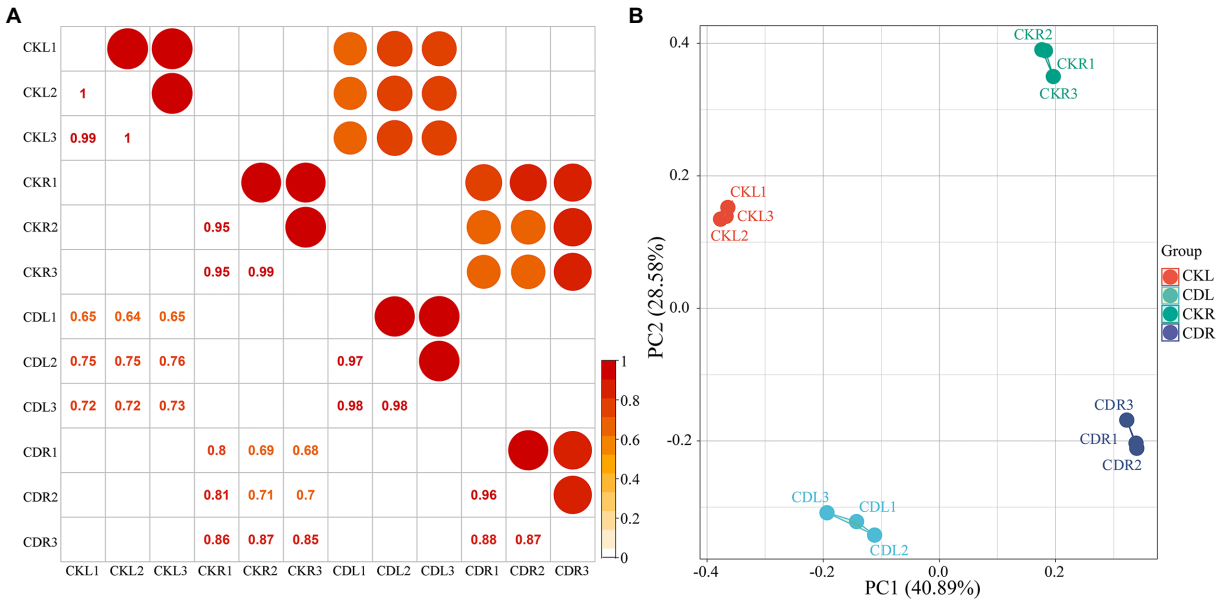


FIGURE 2
Pearson correlation **(A)** and principal component analysis **(B)** among 12 samples. CKL, leaf control groups; CKR, root control groups; CDL, PEG + cold treated leaf groups; CDR, PEG + cold treated root groups. The red background represents a high correlation coefficient.

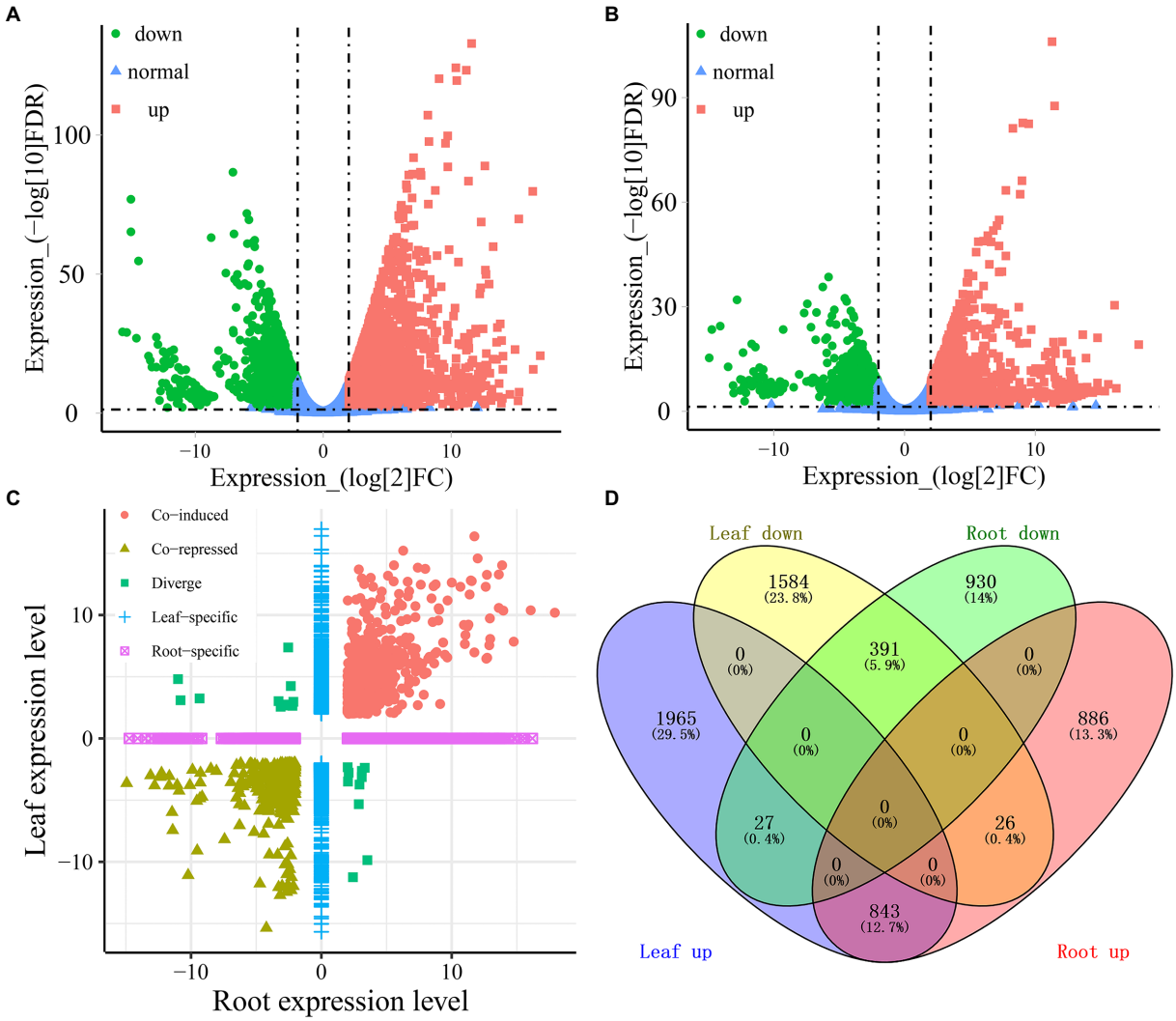


FIGURE 3
Identification of the DEGs in response to cold-drought stress between leaves and roots. Volcano plots display $\log_2$ converted fold changes and FDR values in leaves **(A)** and roots **(B)**. The expression level distribution **(C)** and the number of up-and downregulated DEGs **(D)** in each tissue.

"Response to water." Among the 67 "Photosystem II" term-enriched DEGs, all of them were upregulated in leaves and showed tissue-specific expression patterns, and 61 of them were not detected in roots (Supplementary Table S5). Notably, 20 DEGs enriched in the "Response to water" term were all upregulated, and 40% of DEGs were significantly upregulated ($Log_2FC > 10$) in leaves. One, four, nine, and six DEGs of them were annotated as "Endochitinase A2," "Embryogenic cell protein 40," "Dehydrin DHN2," and "Dehydrin-cognate," respectively. Nearly half of them were also upregulated in roots (Supplementary Table S5).

Similarly, 11 BP, 13 MF, and one CC category were enriched in the roots. The three most significantly enriched categories were "defence response," followed by "oxidation–reduction process" and "glutathione transferase activity" (Figure 4B). In total, 70 DEGs were enriched in the "defence response" term, 53 (88.6%) were downregulated, and 46 specifically expressed DEGs were found in roots, with 8 upregulated and 38 downregulated. In particular, the DEGs annotated as "TMV resistance protein" and "NB-ARC domain disease resistance protein" were mostly downregulated, implying that DEGs related to "defence response" might play a negative role in the response to multiple stresses. The DEGs belonging to "ABA-responsive protein" and "Endochitinase A2" were all upregulated. Notably, 18 DEGs enriched in the "glutathione transferase activity" term was all upregulated in roots, among which five were upregulated in the leaves and roots (Supplementary Table S6). Furthermore, there were 1,187 DEGs were enriched into 15 BP and 10 CC categories in both tissues (Figure 4C). The "Sucrose synthase/metabolic activity" terms were most significantly enriched, followed by "Glutamate-5-semialdehyde dehydrogenase activity," and "Glutamate 5-kinase activity." Interestingly, 24 DEGs were detected in the above-mentioned terms, and all of them were upregulated in the leaves

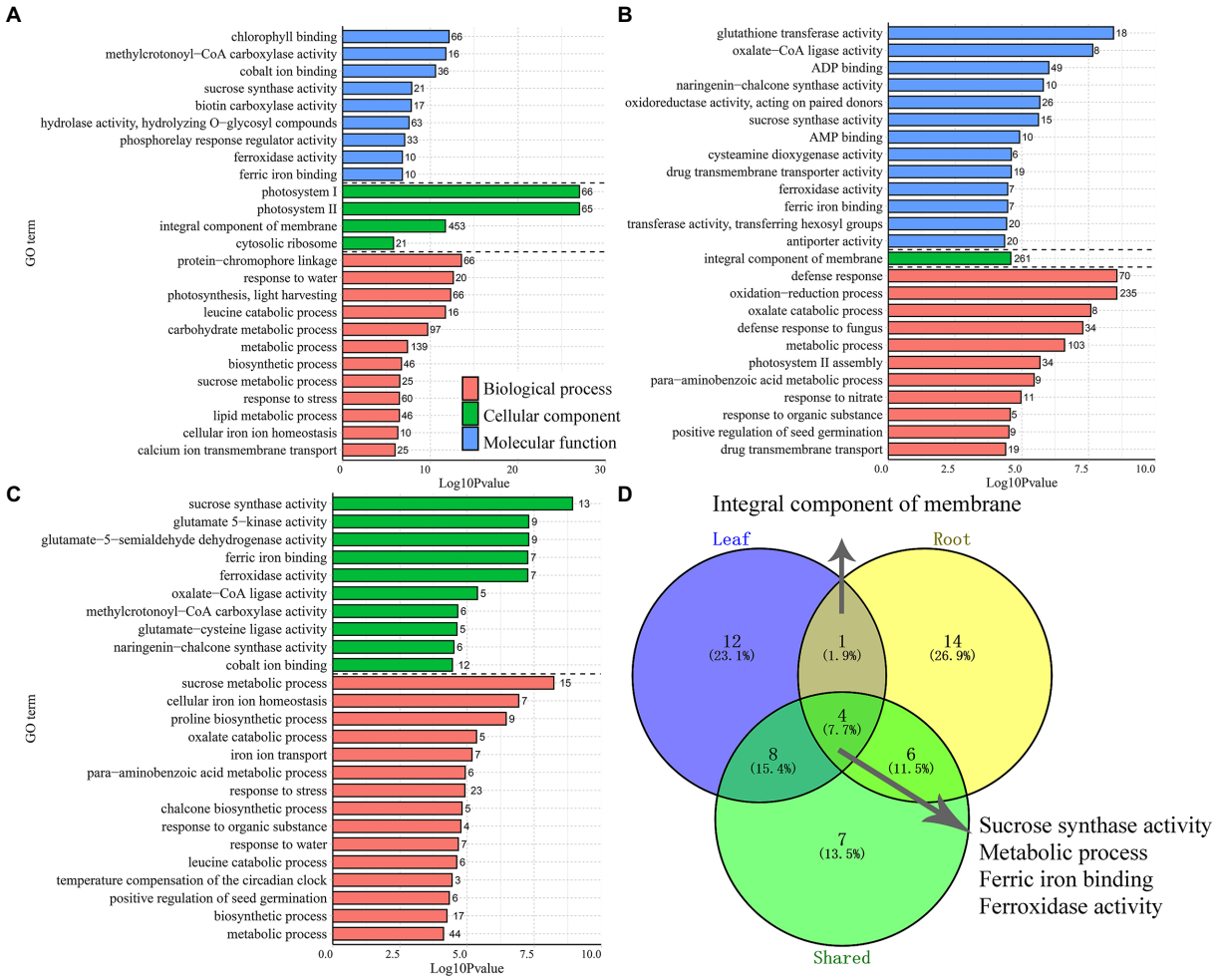


FIGURE 4
The distribution of the top 25 most represented GO categories. The DEGs obtained from leaves **(A)**, roots **(B)**, and shared between two tissues **(C)** were assigned to three main categories: biological process, cellular component, and molecular function. The number of enriched GO categories distribution **(D)**.

and roots, indicating that these DEGs may be highly coordinated in response to abiotic stress between the above-and underground tissues of common vetch (Supplementary Table S7).

Consensus and different co-expression modules identify by pathway enrichment

To explore the key pathway under combined stress, KEGG pathway analysis of DEGs obtained from leaves, roots, and co-induced (shared in both tissues) was conducted by meta-analysis. The enrichment result showed that 1,161, 881, and 340 cold-drought-induced DEGs were assigned to 121, 112, and 88 pathways in the leaves, roots, and co-induced, respectively. The significantly enriched pathways ($p < 0.05$) were further screened, and it was found that 10 common elements (23.3%), including "Circadian rhythm-plant," "Cyanoamino acid metabolism," "Valine, leucine, and isoleucine degradation," "Starch and sucrose metabolism," "Plant hormone signal transduction," "Sphingolipid metabolism," "Galactose metabolism," "Arginine and proline metabolism," "ABC transporters," and "Alanine, aspartate and glutamate metabolism," were the shared pathways in leaves (Figure 5A), roots (Figure 5B), and co-induced DEGs (Figure 5C). Except for these coenriched pathways, 18 (41.9%), 9 (20.9%), and 2 (4.7%) pathways were specifically enriched in leaves, roots, and co-induced DEGs, respectively.

KEGG pathway enrichment results showed that "ABC transporters" was the most enriched pathway among leaves, roots, and DEGs shared between both tissues. In the current study, six common vetch ABC transporters (*VsABC*) were detected in both tissues, and all of them were upregulated in roots, while, three of them had the opposite expression patterns (*F01.PB27856*, *F01.PB32848*, and *F01.PB36915*), which were upregulated in roots, but downregulated in leaves (Supplementary Table S8). Eight and seven *VsABC* genes showed tissue-specific expression profiles, which were specifically identified in leaves and roots, respectively. Compared with roots, more ABC transporters were downregulated in leaves, suggesting that the function of *VsABC* genes may have

tissue specificity, and play different functions between underground and aboveground organizations.

Previous studies have found that the *OsABCG9* gene is primarily expressed in rice leaves during vegetative growth. In mutated *OsABCG9* plants, the cuticular wax contents on the leaves are diminished by half, and this mutant exhibited growth retardation and sensitivity to drought stress (Nguyen et al., 2018). It is worth noting that the pathway of "Cutin, suberin, and wax biosynthesis" was significantly enriched in leaves. Five DEGs were identified in this pathway, two of which were upregulated in both tissues (*F01.PB14194*: *Protein ECERIFERUM 1* and *F01.PB19883*: *Cytochrome P450 94A1*), one upregulated (*F01.PB22120*: *Cytochrome P450 86A1*) in leaves, and reaming two (*F01.PB25896* and F01.PB31997 were identified as *Fatty acyl-CoA reductase 1*) downregulated in leaves (Supplementary Table S8), indicating that *VsABC* genes may play an important function by regulating the common vetch cutin, suberin, and wax biosynthesis pathway against combined abiotic stress in leaves.

Transcription factors in response to cold-drought stress

Transcription factors are important upstream regulators, directly responsible for the activation or repression of various abiotic stresses (Adcock and Caramori, 2009). For example, ERF (ethylene-responsive element binding proteins), bZIP, NAC (NAM/ATAF/CUC), GRAS (GAI/RGA/SCR), bHLH (basic helix–loop–helix), WRKY (WRKY-domain), and MYB (myeloblastosis) have been proven to participate in multiple stresses in various plants (Gujjar et al., 2014; Baillo et al., 2019). Consistent with previous studies, 189 and 130 DEGs were identified as transcription factors under cold-drought treatments in common vetch leaves and roots and were further divided into 37 and 32 families, respectively (Supplementary Table S9). ERF, bZIP, bHLH, NAC, GRAS, and WRKY represented the most abundant families. Compared with other families, the members of NAC, Dof, EIL, AP2, and MIKC-MADS were all upregulated in both tissues. Furthermore, 126 (49.2%) and 62 (24.2%) tissue-specific induced TFs were identified in leaves and roots, respectively (Figure 6). Some tissue-specifically expressed TFs have also been identified and have less-defined roles in abiotic stress responses. For example, GATA, GRF, M-type_MADS, NF-YA, NF-YC, Nin-like, and YABBY were specifically identified in leaves, FAR1 and SRS were specifically identified in roots, perhaps because functional differentiation occurred in the long-term evolutionary process, and different complementary responses of above-and belowground tissues to combined abiotic conditions.

We also identified sets of TFs with similar behaviors among leaves and roots during cold-drought combined stress, and the most abundant TF families were the NAC (9), followed by the ERF (8), bZIP (7), bHLH (6), and MYB (5) families. Among them, 63 co-responsive DEGs were identified in both tissues, which were

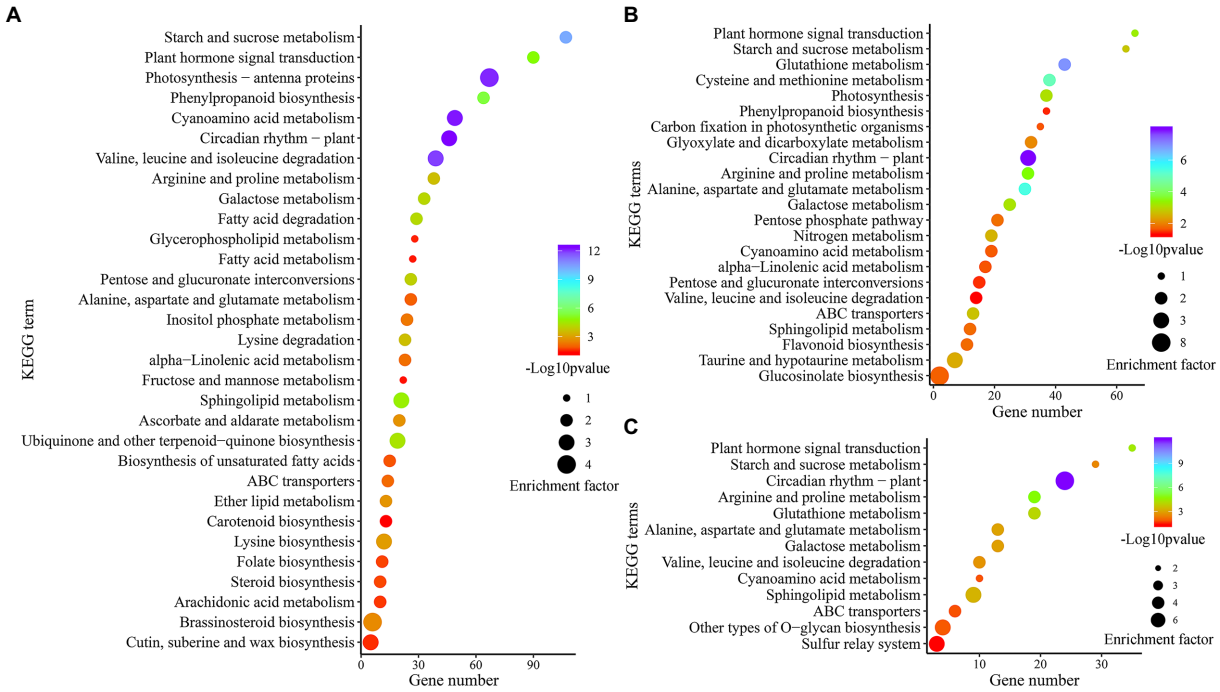


FIGURE 5
Scatter diagram of significantly (p <0.05) enriched KEGG pathways of DEGs in leaves **(A)**, roots **(B)**, and DEGs differentially induced in both tissues **(C)**.

classified into 24 TF families (Supplementary Figure S4), including 45 and 17 DEGs that were up-and downregulated, respectively. Only one gene (F01.PB17105: bZIP) showed the opposite expression pattern, which was upregulated in leaves but downregulated in roots.

Consensus module identification and hub gene protein–protein interaction network analysis

To detect the conserved modules of DEGs under multiple abiotic stresses, the consensus expression profile modules of single drought, cold, and co-occurrence of cold-drought stress were performed by combining the results we published previously (Min et al., 2020b,c). All DEGs could be clustered into six significantly enriched trends ($p < 0.05$) in leaves and roots, including two upregulated patterns (profile 19 and 17), two downregulated patterns (profile 0 and 2), one profile only upregulated under cold and cold-drought stress (profile 12), and one profile only downregulated under cold and cold-drought stress (profile 7; Figures 7A,B). Compared with other profiles, we paid more attention to profile 19, because these DEGs were upregulated under all abiotic conditions. To further discover the key genes' response to multiple stresses, 29 co-induced DEGs identified in both tissues were selected as hub genes (Figure 7C). The top two hub genes were *Delta-1-pyrroline-5-carboxylate synthase B* (*P5CSB*) and *Late embryogenesis abundant* (*LEA2*). Additionally, some proteins, including ABA-responsive protein ABR18, ERF1-3, monosaccharide-sensing protein 2, MtN21/EamA-like transporter protein, phosphatase 2C protein, and topless-related protein 2, and two hypothetical proteins were also played some crucial roles in the response to multiple abiotic stresses in whole plants. Functional enrichment results demonstrated that 29 co-induced DEGs were relevant to "glutamate-5-semialdehyde dehydrogenase activity," "glutamate 5-kinase activity," "proline biosynthetic process," "cytoplasm," and "oxidation–reduction process" (Figure 8A). Additionally, we also found that the arginine and proline metabolism pathway (ko00330), and biosynthesis of amino acids (ko01230) were the most enriched pathways among the hub genes (Supplementary Table S10).

To explore functional interactions among 29 hub DEGs obtained from common vetch, STRING analysis was conducted to present known protein coexpression by BLASTp homology searches against the *Arabidopsis* genome, and 19 unique orthologues were identified. Furthermore, Cytoscape was used to display the most overrepresented DEGs (Figure 8B). In the protein–protein interaction network, nearly half of them were enzyme proteins, including four hydrolases [*AtPP2C33* (*F01.PB22768*), *UBP2* (*F01.PB24895*), *MEE13.8* (*F01.PB29312*), *FAC1* (*F01.PB21275*)], three protein kinase genes [*AT5G58950* (*F01.PB17503*), *P5CS2* (*F01.PB22755*), *P5CS1* (*F01.PB17947*)], one oxidoreductase [*ALDH12A* (*F01.PB33234*)], one aminotransferase [*POP2* (*F01.PB20922*)] and, one ligase [*MCCA* (*F01.PB33120*)].

Functional validation of hub genes

The enrichment result showed that "arginine and proline metabolism" and "plant hormone signal transduction" were the most co-enriched pathways in leaves, roots, and co-induced DEGs, respectively (Figure 5). Previous research has shown that *Delta l-pyrroline-5-carboxylate synthetase* (*P5CS5*) catalyzes is the first and second steps in proline biosynthesis. In our study, we find that five *VsP5CS2* amino acid kinases (*F01.PB22519*, *F01.PB13370*, *F01.PB22755*, *F01.PB17947*, and *F01.PB25217*) were upregulated under all abiotic conditions in both leaves and roots. Among them, the expression of *F01.PB22519* near zero under

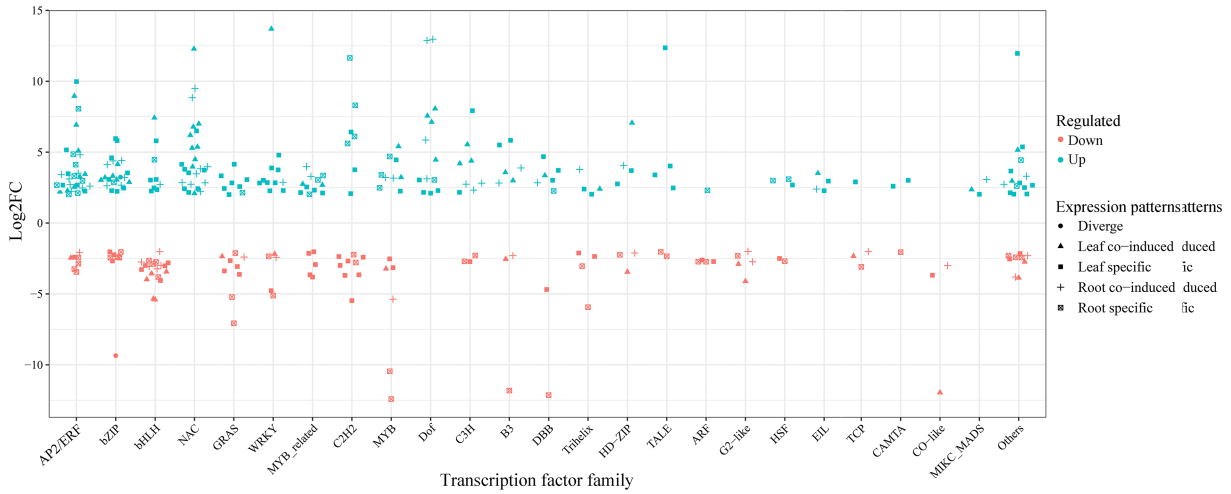


FIGURE 6
The expression pattern distribution of the most enriched transcription factors in response to cold-drought stress between leaves and roots.

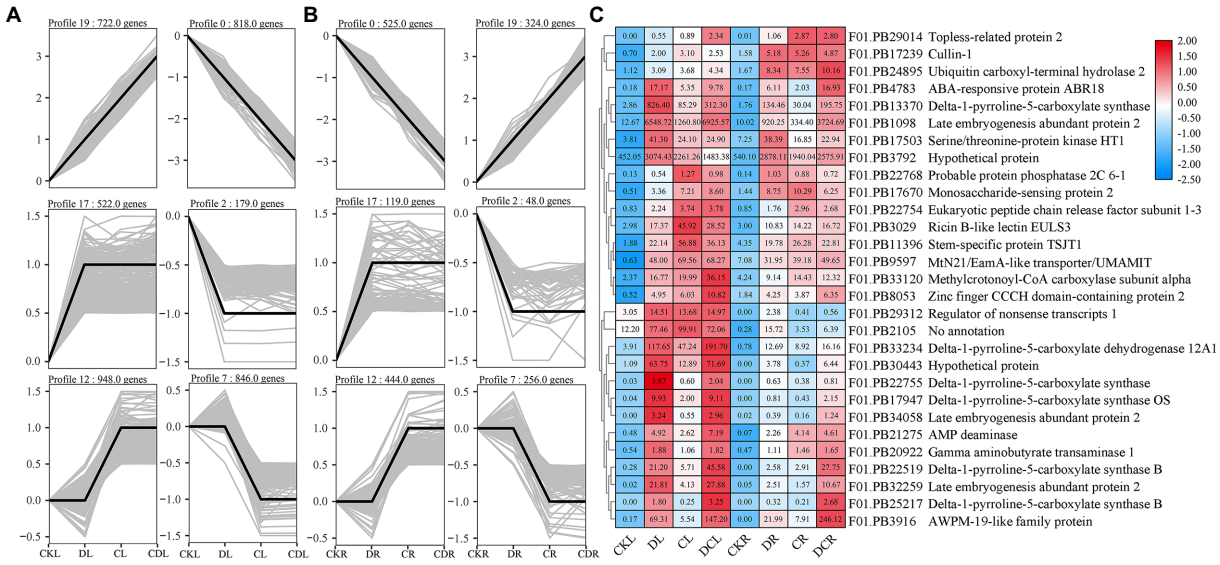


FIGURE 7
Dynamic progression of the common vetch transcriptome under cold-drought stress. **(A,B)** show the dynamic expression of DEGs in leaves and roots, respectively, by K-means clustering. **(C)** shows the expression profile of shared DEGs in profile 17 between leaves and roots. CKL and CKR are leaf and root control groups, respectively, DL, CL, and DCL are leaf under single drought, single cold, and cold-drought combined stress respectively, DR, CR, and DCR are root under single drought, single cold, and cold-drought combined stress, respectively.

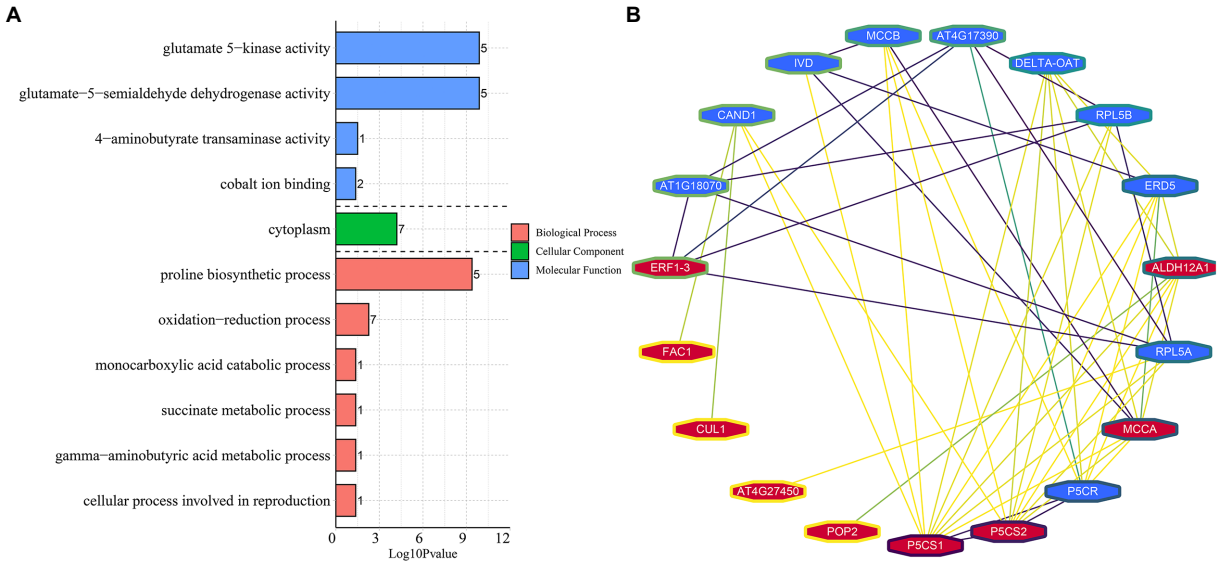


FIGURE 8
Functional enrichment **(A)** and protein−protein interaction analysis of 29 hub DEGs **(B)**.

normal conditions, but significantly upregulated in both tissues under different abiotic stress. In addition, *F01.PB9597* (*VsUMAMIT*: Usually Multiple Acids Move in and out Transporters) was identified as an amino acid and auxin transporter, which was upregulated in all abiotic stresses, especially cold-drought combined stress. Furthermore, F01.PB3916 (VsPM19) is an ABA-induced plasma membrane protein, previous research has shown that an increased level of this protein leads to greater tolerance to low temperature (Koike et al., 1997). The rice homolog, OsPM19L1, is induced by osmotic stress and may be associated with stress tolerance through an ABA-dependent pathway (Chen et al., 2015). But whether *F01.PB9597* and *F01.PB3916* responds to multiple abiotic stresses remains unclear. Therefore, *F01.PB9597*, *F01.PB22519*, and *F01.PB3916* were selected as candidate genes for further studies. To verify three hub genes in response to individual and simultaneous

stresses, we overexpressed them in the *S. cerevisiae* yeast strain INVSc1.

Under normal conditions, the *pYES2-VsPM19* recombinant plasmid grows better than the empty yeast cells, while *pYES2-VsP5CS2* were not growing well as the control. When the transgenic yeast was treated with 5 M NaCl, 30% PEG, cold (−20°C), and cold-30%PEG combined stress for 36 h, the empty and recombinant plasmid cells showed different growth states. Under single abiotic stress, *pYES2-VsPM19* and *pYES2-VsP5CS2* can still grow in 10^{-5}, but the *pYES2* plasmid transgenic yeast cells were hardly grown in a 10^{-4} dilution. Although *pYES2-VsPM19* grew well than control under single abiotic stress, did not show a better cold-30%PEG combined stress tolerance, indicating that *VsPM19* could not enhance compound stress tolerance (Figure 9). Among different abiotic stresses, *VsUMAMIT* were more sensitive to 5 M NaCl, 30% PEG and cold-30%PEG stresses. These combined data indicated that *VsUMAMIT* and *VsP5CS2* could enhance abiotic tolerance, *VsPM19* can only increase the tolerance to single stress, the precise regulatory mechanism still needs further studies. These three hub genes that have been validated will provide a foundation for improving abiotic tolerance in future common vetch genetic manipulation.

Discussion

The effects of cold and drought on transcriptional regulation have been broadly studied separately. Generally, a cascade of biochemical and molecular events was induced and led to similar responses in most individual abiotic stresses. However, in natural environments, simultaneous stress frequently occurs and always causes a more severe effect on plants. There may exist some unique, complex regulatory pathways under combined abiotic stress (Pandey et al., 2017). Additionally, involving combined stresses, plant leaves, and roots can respond to these changes in various ways. Thus, at the molecular level, understanding common vetch responses to cold-drought combined stresses among leaves and roots will be necessary to identify key molecular targets and adaptive mechanisms.

Common vetch responses to combined cold and drought stress are also unique

Our previous studies showed that the length and fresh weight of the aboveground and underground parts were significantly reduced from those of the control groups ($p < 0.05$), especially the fresh weight, which is 32% less than that of the control when the PEG concentration was up to 20% in cultivar "Lanjian No.1" (Min et al., 2020b). We also compared the physiological indicators of "Lanjian No.1" under 4°C treatment, and found that the fresh weight decreased by 4.4% after 24 h compared with that before treatment, and a significant increase in malondialdehyde and soluble sugar contents was detected at 24 h and maintained for at least 48 h under 4°C cold stress in leaves (Min et al., 2020c; Cui, 2021). At the molecular level, the responses of the leaves and roots of common vetch to combined cold and drought differ from those to each abiotic stress separately. Cold stress can alter the cell membrane fluidity and protein conformation, while drought induces hyperosmotic stress in plant cells (Zhang et al., 2022). However, cold and drought in combination alter photosynthesis, stomatal conductance, and respiration, distinctly from cold or drought alone. Under low temperature and drought conditions alone, more DEGs were identified in leaves, especially under low temperature, indicating that leaves are more sensitive to abiotic stress than roots. Consistent with previous studies, transcriptome sequencing showed that combined cold and drought stresses also have a greater impact on the leaves, and more upregulated DEGs were identified in maize and tomato leaves (Zhou et al., 2019; Guo et al., 2021).

Previous studies have indicated that plant responses to combined abiotic stress also exhibit some unique regulatory mechanisms and cannot be inferred from individual stresses. For example, a transcriptomic study in common vetch showed that Ca^{2+} signaling, hormonal signaling, photosynthesis signaling, and redox pathways play important roles through CBF-dependent or CBF-independent transcriptional mechanisms to enhance cold resistance (Min et al., 2020c). Under single drought stress alone, hormonal signaling, starch and sucrose metabolism, and arginine and proline metabolism were the most enriched pathways among leaves and roots (Min et al., 2020b). Thus, a comprehensive understanding is essential for improving common vetch resistance to combined stress. Under combined cold and drought stress, ABC transporters were the most enriched pathway among leaves and roots, followed by arginine and proline metabolism, circadian rhythm-plant, plant hormone signal transduction, starch, and sucrose metabolism (Figure 5). Cutin, suberin and wax biosynthesis, and glucosinolate biosynthesis were the most enriched pathways in leaves and roots, respectively. ABC transporters bind ATP and hydrolyze ATP, which are essential for plant growth, development, and adaptation to the harsh environment (Dahuja et al., 2021). In this study, 18 *ABCB*s and three *ABCC*s were identified as having different (positive or negative), or tissue-specific expression patterns. Some studies have shown that the expression of *ABC* genes may regulate wax and cutin transport in plant leaves, and further affect plant responses to stress, including cold and drought. In rice, compared with the wild type, the wax crystals disappeared in osabcg9-2 mutant leaves, and cuticular wax was diminished by 53%, increased leaf chlorophyll leaching, and became more sensitive to drought stress in *osabcg9-1* mutant plants (Nguyen et al., 2018). In *Arabidopsis*, *AtABCG11*, *AtABCG12*, *AtABCG13*, and *AtABCG32* are involved in cuticular wax and cutin monomer transportation (Yeats and Rose, 2013). Under natural conditions, common vetch always faces combined abiotic stresses, so it is important to understand how common and stress-specific response pathways interact.

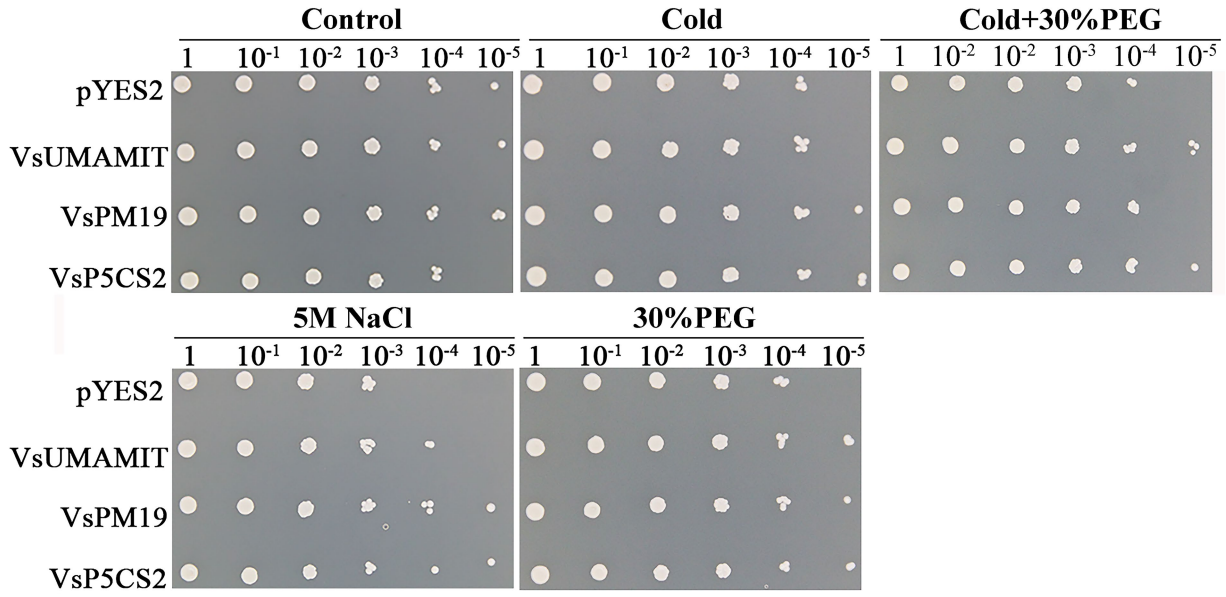


FIGURE 9
The expression of pYES2, pYES2-VsUMAMIT, pYES2-VsPM19, and pYES2-VsP5CS2 in INVSc1 cells induced under cold (−20°C), 5M NaCl, 30% PEG, and cold + 30% PEG stresses.

In general, combined cold and drought stress is more likely to disrupt the PS-II function of aboveground tissues in plants, and then significantly reduce photosynthetic activity (Guo et al., 2021). Upon abiotic stress, abiotic-stabilized phyB could promote plant tolerance by controlling the expression of some stress-responsive genes, enhancing photosynthetic efficiency (Zhang et al., 2022). The *CAB*s are the main functional components of the light-harvesting complex (LHC) in higher plants. *CAB* genes are classified into 10 gene families, and six *CAB* families, including *LHCb1*, *LHCb2*, *LHCb3*, *LHCb4*, *LHCb5*, and *LHCb6*, are related to photosystem II (PS II; Plöchinger et al., 2016). CAB genes have been identified to play a crucial role in the response to light intensity, drought, low temperature, and salinity stress. Wang et al. (2017) performed a proteomic analysis of tea leaves and found that the expression of *CAB* genes increased dramatically during drought stress (Wang et al., 2017). *Arabidopsis LHCb4* mutant plants have a disrupted PS II macrostructure and are defective in photoprotection, proving that *LHCb4* plays a unique role among PS II antenna proteins and photoprotection (de Bianchi et al., 2011). Abundant chlorophyll a-b binding proteins significantly contributed to the "photosynthesis" pathway, and all of them were significantly upregulated, indicating that photosynthesis was enhanced under abiotic stress in common vetch leaves.

Dehydrin (*DHN*) genes are highly hydrophilic and are predominantly induced by cold, drought, and salt stresses (Shekhawat et al., 2011). Previous studies have suggested that *DHN* genes can promote the accumulation of osmotic material by protecting photosynthesis, enriching chlorophyll content, enhancing water retention capacity, and activating ROS detoxification (Singh et al., 2019). For instance, when rice seedlings are exposed to severe drought conditions resulting all *DHN* genes are upregulated in leaves and roots (Verma et al., 2017). In addition, the cold and drought tolerance was improved by overexpressing *DHN* genes in transgenic tobacco, cotton, and Sorghum plants, resulting in malondialdehyde, lipid peroxidation, relative electrolyte leakage, and water loss reduction (Bao et al., 2017; Guo et al., 2017; Halder et al., 2017; Kirungu et al., 2020). Our results emphasized that DHN proteins have a positive impact on common vetch cold-drought stress tolerance. In total, 20 differentially expressed *DHN* genes were enriched to the "Response to water" category, all *DHN* genes were upregulated in leaves, and seven of them were upregulated in both tissues. Furthermore, some of *VsDHNs*, including *F01.PB13924*, *F01.PB6182*, *F01.PB8865*, *F01.PB5084*, *F01.PB3012*, and *F01.PB47* were found to be negligible in leaves and/or roots under normal conditions and was significantly upregulated by more than 100-fold under combined treatments. Abundant *CAB* and *DHN* genes were significantly induced and enriched in common vetch leaves, confirming that *DHN*s may play a vital role in maintaining the photosynthetic rates under combined stress.

TFs involved in the combined cold and drought stress

When plants face abiotic stress, many genes, especially transcription factors, can respond rapidly and regulate the expression of downstream functional genes. Previous studies found that C-repeat/dehydration-responsive element-binding (CBF/DREB) transcription factors could be activated by AP2/ERF, CAMTA, MYB, and nuclear factor Y (NF-Y), and play a key role in low temperatures and/or water deficit (Wang et al., 2012;

Park et al., 2015; Hu et al., 2020). AP2/ERF family, as one of the largest gene families in plants, plays a diverse role in response to abiotic stresses and phytohormones and can control a wide network of downstream genes in the cellular signaling pathway (Faraji et al., 2020; Heidari et al., 2021). Consistent with previous studies, 18 and 22 AP2/ERF genes were identified under combined stress in common vetch leaves and roots. Among them, nine AP2/ERF genes were upregulated in both tissues. Some AP2/ERF genes showed tissue-specific expression patterns. Two *CAMTA* (*F01.PB25251*, *F01.PB28165*) were specifically upregulated in leaves, and one (*F01.PB23756*) was specifically downregulated in roots. Several transcription factor families have shown tissue-specific expression patterns were obtained. For example, *NF-YA* (*F01.PB13198*) was upregulated in leaves, and *NF-YB* (*F01.PB3279*) was downregulated in leaves. The CBF transcriptional level response is quite different between leaf and root tissues, which might interact and/or inhibit each other.

The NAC (NAM, ATAF, and CUC) TF family members play an important role in plant growth, development, and stress responses. In *Arabidopsis*, *ANAC019* (*AT1G52890.1*), *ANAC055* (*AT3G15500*), and *ANAC072* (*AT4G27410.2*) were induced by drought, salinity, and/or low temperature, and transgenic plants overexpressing either of them showed significantly increased stress tolerance compared to the wild type (Tran et al., 2004). In this study, the NAC transcription factors were observed to be the most significantly upregulated, including *ANAC072* orthologous *F01.PB6545*, *F01.PB8282*, and *F01.PB9565*, which are upregulated in both tissues. Dof (DNA binding with one finger) has been reported to participate in the regulation of gene expression in plant defense processes (Waqas et al., 2020). The expression levels of Dof transcription factors (*F01.PB31167* and *F01.PB10374*) changed significantly in both tissues in common vetch. Some of the previously uncharacterized abiotic transcription factors also identified in this study might represent novel regulators of combined stress tolerance in whole plants. Examples include HSI2-like 1 (*F01.PB35573*) and DBB (*F01.PB6342*) transcription factors.

Conclusion

The present study designed to compare the transcription profiles of cold-drought combined stress between aboveground and underground tissues revealed the operation of mechanisms that (a) the responses of leaves and roots of common vetch to combined cold and drought differ from that subjected to single stress, as well as some unique genes related to stress tolerance, (b) cold-drought combined stress has a greater impact on leaves, and more upregulated DEGs were identified, (c) improve tolerance to combined stress by upregulating a large amount of chlorophyll a-b binding protein and water responsive genes (e.g., ECP40, DHN2 and Endochitinase A2) in leaves, many d response proteins (e.g., TMV resistance protein, NB-ARC domain disease resistance protein, ABA-responsive protein, and Endochitinase A2) in roots, and multifunctional enzymes (e.g., GABA-TP1, SUS2, and P5CS) were upregulated in both tissues. Additionally, the hub genes involved in the single and combined cold-drought stress were identified and functionally validated. Further metabonomic and transgenic studies are needed to study the effect of hub genes under both single and combined stresses.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: https://doi.org/10.6084/m9.figshare.20629467.v1.

Author contributions

XM: methodology, software, validation, formal analysis, investigation, data curation, visualization, and writing—original draft. QW: formal analysis, supervision, and validation. ZW: resources and project administration. ZL: methodology, funding acquisition, and supervision. WL: investigation, methodology, conceptualization, funding acquisition, visualization, and project administration. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by the National Natural Science Foundation of China (32071862), the Natural Science Foundation of Gansu Province (20JR10RA622), and the Fundamental Research Funds for the Central Universities (lzujbky-2021-ct21, lzujbky-2021-it05, and lzujbky-2021-ct14), the Nature Science Foundation of the Jiangsu Higher Education Institution of China (22KJB230011), and the Natural Science Foundation of Jiangsu Province (BK20220583).

Acknowledgments

We thank Jiaoyang Tian and Xiaoshan Lin for their assistance with the qRT-PCR analysis.

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.2022.976094/full#supplementary-material

SUPPLEMENTARY FIGURE S1
Characteristics of novel transcripts. **(A)** The expression pattern distribution of novel transcripts. **(B)** The expression level of novel transcripts under cold-drought combined stress in leaves and roots.

SUPPLEMENTARY FIGURE S2
Validation of the expression profiles (log2-fold change) of 15 selected DEGs from leaves and roots using qRT–PCR. The green bars represent qRT-PCR-expressed leaves (left y-axis), and the blue trend lines represent the FPKM value changes of the DEGs (right y-axis).

SUPPLEMENTARY FIGURE S3
The expression (Log2-fold change) correlation of 15 DEGs based on RNA-Seq via qRT-PCR.

SUPPLEMENTARY FIGURE S4
The expression profile of 63 co-induced transcription factors in leaves and roots, respectively.

References

Adcock, I. M., and Caramori, G. (2009). "Chapter 31 – transcription factors," in *Asthma and COPD. 2nd Edn*. eds. P. J. Barnes, J. M. Drazen, S. I. Rennard and N. C. Thomson *Second* ed (Oxford: Academic Press), 373–380.

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, 3389–3402.

Baillo, E. H., Kimotho, R. N., Zhang, Z., and Xu, P. (2019). Transcription factors associated with abiotic and biotic stress tolerance and their potential for crops improvement. *Genes* 10:771. doi: 10.3390/genes10100771

Bao, F., Du, D., An, Y., Yang, W., Wang, J., Cheng, T., et al. (2017). Overexpression of Prunus mume Dehydrin genes in tobacco enhances tolerance to cold and drought. *Front. Plant Sci.* 8: 151. doi: 10.3389/fpls.2017.00151

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., Lan, H., Huang, P., Zhang, Y., Yuan, X., Huang, X., et al. (2015). Characterization of OsPM19L1 encoding an AWPM-19-like family protein that is dramatically induced by osmotic stress in rice. *Genet. Mol. Res.* 14, 11994–12005. doi: 10.4238/2015.october.5.12

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

Correia, B., Hancock, R. D., Amaral, J., Gomez-Cadenas, A., Valledor, L., and Pinto, G. (2018). Combined drought and heat activates protective responses in *Eucalyptus globulus* that are not activated when subjected to drought or heat stress alone. *Front. Plant Sci.* 9: 819. doi: 10.3389/fpls.2018.00819

Cui, Y. (2021). *Transcrptome and metabonomic analysis of cold resistance of "Lanjian" cultivars in common vetch (Vicia sative) [D]*. Lanzhou: Lanzhou University.

Dahuja, A., Kumar, R. R., Sakhare, A., Watts, A., Singh, B., Goswami, S., et al. (2021). Role of ATP-binding cassette transporters in maintaining plant homeostasis under abiotic and biotic stresses. *Physiol. Plant.* 171, 785–801. doi: 10.1111/ppl.13302

de Bianchi, S., Betterle, N., Kouril, R., Cazzaniga, S., Boekema, E., Bassi, R., et al. (2011). *Arabidopsis* mutants deleted in the light harvesting protein LHCB4 have a disrupted photosystem II macrostructure and are defective in photoprotection. *Plant Cell* 23, 2659–2679. doi: 10.1105/tpc.111.087320

Faraji, S., Filiz, E., Kazemitabar, S. K., Vannozzi, A., Palumbo, F., Barcaccia, G., et al. (2020). The *AP2/ERF* gene family in *Triticum durum*: genome-wide identification and expression analysis under drought and salinity stresses. *Genes* 11:1464. doi: 10.3390/genes11121464

Gargallo-Garriga, A., Sardans, J., Pérez-Trujillo, M., Oravec, M., Urban, O., Jentsch, A., et al. (2015). Warming differentially influences the effects of drought on stoichiometry and metabolomics in shoots and roots. *New Phytol.* 207, 591–603. doi: 10.1111/nph.13377

Gong, Z., Xiong, L., Shi, H., Yang, S., Herrera-Estrella, L. R., Xu, G., et al. (2020). Plant abiotic stress response and nutrient use efficiency. *Sci. China Life Sci.* 63, 635–674. doi: 10.1007/s11427-020-1683-x

Gujjar, R. S., Akhtar, M., and Singh, M. (2014). Transcription factors in abiotic stress tolerance. *Indian J. Plant Physiol.* 19, 306–316. doi: 10.1007/s40502-014-0121-8

Guo, Q., Li, X., Niu, L., Jameson, P. E., and Zhou, W. (2021). Transcription-associated metabolomic adjustments in maize occur during combined drought and cold stress. *Plant Physiol.* 186, 677–695. doi: 10.1093/plphys/kiab050

Guo, X., Zhang, L., Zhu, J., Liu, H., and Wang, A. (2017). Cloning and characterization of SiDHN, a novel dehydrin gene from Saussurea involucrata Kar. Et Kir. That enhances cold and drought tolerance in tobacco. *Plant Sci.* 256, 160–169. doi: 10.1016/j.plantsci.2016.12.007

Halder, T., Upadhyaya, G., and Ray, S. (2017). YSK2 type Dehydrin (SbDhn1) from Sorghum bicolor showed improved protection under high temperature and osmotic stress condition. *Front. Plant Sci.* 8: 918. doi: 10.3389/fpls.2017.00918

Hasanuzzaman, M., Hakeem, K. R., Nahar, K., and Alharby, H. F. (2019). *Plant abiotic stress tolerance: Agronomic, Molecular and Biotechnological Approaches* (Switzerland: Springer).

Heidari, P., Entazari, M., Ebrahimi, A., Ahmadizadeh, M., Vannozzi, A., Palumbo, F., et al. (2021). Exogenous EBR ameliorates endogenous hormone contents in tomato species under low-temperature stress. *Horticulturae* 7:84. doi: 10.3390/horticulturae7040084

Hu, Z., Ban, Q., Hao, J., Zhu, X., Cheng, Y., Mao, J., et al. (2020). Genome-wide characterization of the C-repeat binding factor (CBF) gene family involved in the response to abiotic stresses in tea plant (Camellia sinensis). *Front. Plant Sci.* 11: 921. doi: 10.3389/fpls.2020.00921

Huang, Y., Gao, X., Nan, Z., and Zhang, Z. (2017). Potential value of the common vetch (Vicia sativa L.) as an animal feedstuff: a review. *J. Anim. Physiol. Anim. Nutr.* 101, 807–823. doi: 10.1111/jpn.12617

Jin, J., Tian, F., Yang, D.-C., Meng, Y.-Q., Kong, L., Luo, J., et al. (2016). PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res.* 45, D1040–D1045. doi: 10.1093/nar/gkw982

Kartal, G. K., Senbek, G., Karaca, M., and Acikgoz, E. (2020). Hybridization studies in Vicia sativa complex. *Euphytica* 216:29. doi: 10.1007/s10681-020-2566-3

Kawai, S., Hashimoto, W., and Murata, K. (2010). Transformation of Saccharomyces cerevisiae and other fungi. *Bioeng. Bugs* 1, 395–403. doi: 10.4161/bbug.1.6.13257

Kirungu, J. N., Magwanga, R. O., Pu, L., Cai, X., Xu, Y., Hou, Y., et al. (2020). Knockdown of Gh_A05G1554 (GhDHN_03) and Gh_D05G1729 (GhDHN_04) Dehydrin genes, reveals their potential role in enhancing osmotic and salt tolerance in cotton. *Genomics* 112, 1902–1915. doi: 10.1016/j.ygeno.2019.11.003

Koike, M., Takezawa, D., Arakawa, K., and Yoshida, S. (1997). Accumulation of 19-kDa plasma membrane polypeptide during induction of freezing tolerance in wheat suspension-cultured cells by abscisic acid. *Plant Cell Physiol.* 38, 707–716. doi: 10.1093/oxfordjournals.pcp.a029224

Larbi, A., Abd El-Moneim, A., Nakkoul, H., Jammal, B., and Hassan, S. (2011). Intra-species variations in yield and quality determinants in Vicia species: 3. Common vetch (Vicia sativa ssp. sativa L.). *Anim. Feed Sci. Technol.* 164, 241–251. doi: 10.1016/j.anifeedsci.2010.09.002

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

Lozano-Elena, F., Fàbregas, N., Coleto-Alcudia, V., and Caño-Delgado, A. I. (2022). Analysis of metabolic dynamics during drought stress in Arabidopsis plants. *Sci. Data* 9:90. doi: 10.1038/s41597-022-01161-4

Mahalingam, R. (2015). "Consideration of combined stress: a crucial paradigm for improving multiple stress tolerance in plants," in *Combined stresses in plants: Physiological, molecular, and biochemical aspects*. ed. R. Mahalingam (Cham: Springer International Publishing), 1–25.

Min, X., Jin, X., Zhang, Z., Wei, X., Ndayambaza, B., Wang, Y., et al. (2020a). Genome-wide identification of NAC transcription factor family and functional analysis of the abiotic stress-responsive genes in Medicago sativa L. *J. Plant Growth Regul.* 39, 324–337. doi: 10.1007/s00344-019-09984-z

Min, X., Lin, X., Ndayambaza, B., Wang, Y., and Liu, W. (2020b). Coordinated mechanisms of leaves and roots in response to drought stress underlying full-length transcriptome profiling in Vicia sativa L. *BMC Plant Biol.* 20:165. doi: 10.1186/s12870-020-02358-8

Min, X., Liu, Z., Wang, Y., and Liu, W. (2020c). Comparative transcriptomic analysis provides insights into the coordinated mechanisms of leaves and roots response to cold stress in common vetch. *Ind. Crop Prod.* 158:112949. doi: 10.1016/j.indcrop.2020.112949

Mittler, R. (2006). Abiotic stress, the field environment and stress combination. *Trends Plant Sci.* 11, 15–19. doi: 10.1016/j.tplants.2005.11.002

Mota, A. P. Z., Brasileiro, A. C. M., Vidigal, B., Oliveira, T. N., da Cunha Quintana Martins, A., Saraiva, M. A. D. P., et al. (2021). Defining the combined stress response in wild Arachis. *Sci. Rep.* 11:11097. doi: 10.1038/s41598-021-90607-7

Nguyen, V. N. T., Lee, S. B., Suh, M. C., An, G., and Jung, K.-H. (2018). OsABCG9 is an important ABC transporter of Cuticular wax deposition in Rice. *Front. Plant Sci.* 9: 960. doi: 10.3389/fpls.2018.00960

Pandey, P., Irulappan, V., Baga Vathiannan, M. V., and Senthil-Kumar, M. (2017). Impact of combined abiotic and biotic stresses on plant growth and avenues for crop improvement by exploiting Physio-morphological traits. *Front. Plant Sci.* 8: 537. doi: 10.3389/fpls.2017.00537

Park, S., Lee, C. M., Doherty, C. J., Gilmour, S. J., Kim, Y., and Thomashow, M. F. (2015). Regulation of the Arabidopsis CBF regulon by a complex low-temperature regulatory network. *Plant J.* 82, 193–207. doi: 10.1111/tpj.12796

Plöchinger, M., Schwenkert, S., von Sydow, L., Schröder, W. P., and Meurer, J. (2016). Functional update of the auxiliary proteins PsbW, PsbY, HCF136, PsbN, TerC and ALB3 in maintenance and assembly of PSII. *Front. Plant Sci.* 7: 423. doi: 10.3389/fpls.2016.00423

Shekhawat, U. K., Srinivas, L., and Ganapathi, T. R. (2011). MusaDHN-1, a novel multiple stress-inducible SK(3)-type dehydrin gene, contributes affirmatively to drought-and salt-stress tolerance in banana. *Planta* 234, 915–932. doi: 10.1007/s00425-011-1455-3

Singh, P. K., Srivastava, D., Tiwari, P., Tiwari, M., Verma, G., and Chakrabarty, D. (2019). "Chapter 6- drought tolerance in plants: molecular mechanism and regulation of signaling molecules" in *Plant signaling molecules*. eds. M. I. R. Khan, P. S. Reddy, A. Ferrante and N. A. Khan (UK: Woodhead Publishing), 105–123.

Su, L., Dai, Z., Li, S., and Xin, H. (2015). A novel system for evaluating drought-cold tolerance of grapevines using chlorophyll fluorescence. *BMC Plant Biol.* 15:82. doi: 10.1186/s12870-015-0459-8

Suzuki, N., Rivero, R. M., Shulaev, V., Blumwald, E., and Mittler, R. (2014). Abiotic and biotic stress combinations. *New Phytol.* 203, 32–43. doi: 10.1111/nph.12797

Tran, L. S., Nakashima, K., Sakuma, Y., Simpson, S. D., Fujita, Y., Maruyama, K., et al. (2004). Isolation and functional analysis of Arabidopsis stress-inducible NAC transcription factors that bind to a drought-responsive cis-element in the early responsive to dehydration stress 1 promoter. *Plant Cell* 16, 2481–2498. doi: 10.1105/tpc.104.022699

Verma, G., Dhar, Y. V., Srivastava, D., Kidwai, M., Chauhan, P. S., Bag, S. K., et al. (2017). Genome-wide analysis of rice dehydrin gene family: its evolutionary conservedness and expression pattern in response to PEG induced dehydration stress. *PLoS One* 12:e0176399. doi: 10.1371/journal.pone.0176399

Wang, Y., Fan, K., Wang, J., Ding, Z.-T., Wang, H., Bi, C.-H., et al. (2017). Proteomic analysis of Camellia sinensis (L.) reveals a synergistic network in the response to drought stress and recovery. *J. Plant Physiol.* 219, 91–99. doi: 10.1016/j.jplph.2017.10.001

Wang, Y., Jiang, C.-J., Li, Y.-Y., Wei, C.-L., and Deng, W.-W. (2012). CsICE1 and CsCBF1: two transcription factors involved in cold responses in Camellia sinensis. *Plant Cell Rep.* 31, 27–34. doi: 10.1007/s00299-011-1136-5

Waqas, M., Shahid, L., Shoukat, K., Aslam, U., Azeem, F., and Atif, R. M. (2020). "Chapter 1 – role of DNA-binding with one finger (Dof) transcription factors for abiotic stress tolerance in plants" in *Transcription factors for abiotic stress tolerance in plants*. ed. S. H. Wani (UK: Academic Press), 1–14.

Werner, T., Holst, K., Pörs, Y., Guivarc'h, A., Mustroph, A., Chriqui, D., et al. (2008). Cytokinin deficiency causes distinct changes of sink and source parameters in tobacco shoots and roots. *J. Exp. Bot.* 59, 2659–2672. doi: 10.1093/jxb/ern134

Xie, C., Mao, X., Huang, J., Ding, Y., Wu, J., Dong, S., et al. (2008). KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res.* 39 (Web Server issue), W316–W322. doi: 10.1093/nar/gkr483

Yeats, T. H., and Rose, J. K. (2013). The formation and function of plant cuticles. *Plant Physiol.* 163, 5–20. doi: 10.1104/pp.113.222737

Zandalinas, S. I., Mittler, R., Balfagón, D., Arbona, V., and Gómez-Cadenas, A. (2018). Plant adaptations to the combination of drought and high temperatures. *Physiol. Plant.* 162, 2–12. doi: 10.1111/ppl.12540

Zhang, H., and Sonnewald, U. (2017). Differences and commonalities of plant responses to single and combined stresses. *Plant J.* 90, 839–855. doi: 10.1111/tpj.13557

Zhang, H., Zhu, J., Gong, Z., and Zhu, J.-K. (2022). Abiotic stress responses in plants. *Nat. Rev. Genet.* 23, 104–119. doi: 10.1038/s41576-021-00413-0

Zhou, R., Yu, X., Zhao, T., Ottosen, C.-O., Rosenqvist, E., and Wu, Z. (2019). Physiological analysis and transcriptome sequencing reveal the effects of combined cold and drought on tomato leaf. *BMC Plant Biol.* 19:377. doi: 10.1186/s12870-019-1982-9

Zhu, J.-K. (2016). Abiotic stress signaling and responses in plants. *Cells* 167, 313–324. doi: 10.1038/s41576-021-00413-0

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 04 October 2022
DOI 10.3389/fpls.2022.995074

OPEN ACCESS

EDITED BY
Linkai Huang,
Sichuan Agricultural University,
China

REVIEWED BY
Min Sun,
Sichuan Agricultural University,
China
Mingshu Cao,
AgResearch Ltd.,
New Zealand

*CORRESPONDENCE
Yong Ding
dingyong@caas.cn

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

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

RECEIVED 15 July 2022
ACCEPTED 26 August 2022
PUBLISHED 04 October 2022

CITATION
Wan D, Wan Y, Zhang T, Wang R and Ding Y (2022) Multi-omics analysis reveals the molecular changes accompanying heavy-grazing-induced dwarfing of *Stipa grandis*. *Front. Plant Sci.* 13:995074. doi: 10.3389/fpls.2022.995074

Multi-omics analysis reveals the molecular changes accompanying heavy-grazing-induced dwarfing of *Stipa grandis*

Dongli Wan[1†], Yongqing Wan[2†], Tongrui Zhang[1], Ruigang Wang[2] and Yong Ding[1*]

[1]Institute of Grassland Research, Chinese Academy of Agricultural Sciences, Hohhot, China, [2]College of Life Sciences, Inner Mongolia Agricultural University, Hohhot, China

Heavy grazing significantly reduces *Stipa grandis* growth. To enhance our understanding of plant responses to heavy grazing, we conducted transcriptomic, proteomic, and metabolic analyses of the leaves of non-grazed plants (NG) and heavy-grazing-induced dwarf plants (HG) of *S. grandis*. A total of 101 metabolites, 167 proteins, and 1,268 genes differed in abundance between the HG and NG groups. Analysis of Kyoto Encyclopedia of Genes and Genomes pathways among differentially accumulated metabolites (DAMs) revealed that the most enriched pathways were flavone and flavonol biosynthesis, tryptophan metabolism, and phenylpropanoid biosynthesis. An integrative analysis of differentially expressed genes (DEGs) and proteins, and DAMs in these three pathways was performed. Heavy-grazing-induced dwarfism decreased the accumulation of DAMs enriched in phenylpropanoid biosynthesis, among which four DAMs were associated with lignin biosynthesis. In contrast, all DAMs enriched in flavone and flavonol biosynthesis and tryptophan metabolism showed increased accumulation in HG compared with NG plants. Among the DAMs enriched in tryptophan metabolism, three were involved in tryptophan-dependent IAA biosynthesis. Some of the DEGs and proteins enriched in these pathways showed different expression trends. The results indicated that these pathways play important roles in the regulation of growth and grazing-associated stress adaptions of *S. grandis*. This study enriches the knowledge of the mechanism of heavy-grazing-induced growth inhibition of *S. grandis* and provides valuable information for restoration of the productivity in degraded grassland.

KEYWORDS

***Stipa grandis*, heavy grazing, transcriptomic, proteomic, metabolic**

Introduction

Stipa grandis (Poaceae, $2n = 44$) is a wind-pollinated C_3 perennial bunchgrass that flowers in mid- to late-July and the seeds ripen in late August or early September (Zhao et al., 2008; Wu et al., 2010b). The mature plants form dense tussocks that are approximately 30 cm high with long and thin leaves (Zhao et al., 2008). *S. grandis* is among the dominant

plant species in typical steppe of the Inner Mongolian Plateau (Wu et al., 2010b). *S. grandis*-dominated grassland, the most common and representative community in the Eurasian steppe (Gao et al., 2018b), has been degraded to varying degrees (Xiao et al., 1995; Wu et al., 2010a), which greatly affects grassland productivity.

Overgrazing and climate change (such as aridification) are the main factors responsible for degradation of grassland of the Inner Mongolian Plateau (Gao et al., 2018a). Livestock grazing is a major anthropogenic disturbance of grasslands on a global scale (Zhao et al., 2009). To respond to grazing, plants alter their morphological and functional traits (Zhao et al., 2009). For instance, the individual plant height and biomass (of leaves, stems, and the whole plant) of *Leymus chinensis* are significantly restricted under overgrazing (Li et al., 2015). The size of individual *S. grandis* plants is minimized to cope with frequent grazing or overgrazing (Wan et al., 2015; Li et al., 2018). The leaf photosynthetic activity, as indicated by net photosynthetic rate, stomatal conductance, intercellular carbon dioxide concentration, and transpiration rate, is markedly decreased following overgrazing (Ren et al., 2017). In recent years, several studies have been conducted to explore the molecular mechanism of plant dwarfism resulting from overgrazing at the transcription and protein levels (Wan et al., 2015; Ren et al., 2018a,b). However, little information is available on the protein and metabolic mechanisms responsible for the response of *S. grandis* to overgrazing.

Proteins are macromolecular compounds that function as the biochemical units in all cellular processes (Mann et al., 2013). Proteomics is the large-scale study of proteins and is crucial for understanding biological processes at the molecular level (Pandey and Mann, 2000; Zhang et al., 2014). Proteomics is widely used to study the processes of plant development and stress responses (Chen et al., 2015; Liu et al., 2015; Jiang et al., 2019; Zhan et al., 2019; Li et al., 2020; Lv et al., 2021), and provides essential tools to study global protein expression and dissect the unique functions underlying the many plant-specific biological processes (Job et al., 2011).

Plant metabolites, which are essential for humans as a nutritional source, play vital roles in the interaction of plants and the surrounding environment (Chen et al., 2020). It has been estimated that 200,000 to 1,000,000 metabolites exist in the plant kingdom (Saito and Matsuda, 2010). Metabolomics, which is the analysis of almost all metabolites in a biological sample using qualitative and quantitative methods, acts as a bridge to link the genome and phenome (Shi et al., 2020). Metabolomic approaches have been used to study metabolism in plant developmental processes and responses to environmental stimuli, such as low temperature (Yang et al., 2019b), salt (Li and Song, 2019), high light (Zhang et al., 2019), fruit development and ripening (Xu et al., 2020), and color variation (Yang et al., 2019a; Li et al., 2021b). Combined analysis of omics data has become a powerful tool to explain diverse developmental processes, environmental responses, and regulatory mechanisms (Chen et al., 2021; Guo et al., 2021; Xu et al., 2021; Li et al., 2021a; Qin et al., 2022). For instance, comparative transcriptomic and metabolomic analysis revealed an abscisic acid-dependent acclimation mechanism to drought and cold stress in maize (Guo et al., 2021). Combined transcriptomic, proteomic, and metabolomic analyses revealed that core metabolic processes were influenced in seeds of transgenic maize engineered for enhanced carotenoid synthesis (Decourcelle et al., 2015). Multi-omics analysis has highlighted distinct roles of central carbon metabolism to assist high productivity of specialized metabolites in glandular trichomes of tomato (Balcke et al., 2017), uncovered sequential roles associated with abscisic acid during seed maturation (Chauffour et al., 2019), and provided new insights into the decline in fruit quality of apple grown under high nitrogen fertilization (Wang et al., 2021b).

In this study, we performed transcriptomic, tandem mass tag (TMT) label-based proteomic, and widely targeted metabolic profiling in leaves of *S. grandis* grown under non-grazing and heavy-grazing treatments. The important metabolic pathways involved in the grazing response were evaluated by combined analysis of the omics data. The aim of this study was to elucidate the mechanism by which heavy grazing depresses *S. grandis* growth from different molecular aspects. The results will improve our understanding of the impact of grazing on metabolite and protein accumulation in *S. grandis*, and provide valuable information for restoration of productivity in degraded grassland.

Materials and methods

Study area and sampling

Sampling was conducted at the Inner Mongolia Typical Grassland Ecological Protection and Restoration Research Station of the Chinese Academy of Agricultural Sciences, located in Xilinhot, Inner Mongolia, China (116°32′E, 44°15′N), on August 2, 2018. The study region has a temperate semiarid continental climate, with a mean annual temperature of 0.7°C, and lowest and highest temperatures of −41.1°C in January and 38.5°C in August, respectively (Liu et al., 2019). The mean annual precipitation is 300–360 mm with maximum precipitation in the period from June to August (Zhang et al., 2020). The dominant plant species in the study area comprised *S. grandis* and *L. chinensis*.

Grazing had been prohibited from the study area from 2007 to 2013, and the grazing experiments were conducted since 2014. The non-grazed (NG) plot was continually protected from grazing from 2014 to 2018. The heavy-grazing (HG) plot was grazed from 2014 to 2017 by 12 sheep. Grazing started in the middle of June and ended in the middle of September each year. The NG and HG plots were 1.33 Ha in area. The grazing pressure of HG is 1.4 standard sheep unit per hectare per year (SSU·ha^{-1}·y^{-1}), which is apparently higher than that of locally allowed standards (0.84 SSU·ha^{-1}·y^{-1}). Each grazing treatment was established with three replicate plots. In 2018, prior to grazing, half of each HG plot was fenced to exclude grazing disturbance, and the individuals of *S. grandis* in this area retained a dwarf phenotype compared with the plants in the NG area (Supplementary Figure S1); the HG samples for subsequent analysis

were collected from plants growing in the exclusion area. The leaves from three bunches of *S. grandis*, sampled at a vigorous growth stage from one plot for each treatment, were respectively pooled as one biological replicate, immediately frozen in liquid nitrogen, and stored at −80°C for subsequent total RNA, protein, and metabolite extraction. Three biological replicates were taken from one sampling plot for each treatment.

Transcriptome analysis

Total RNA of each sample was isolated from leaves of three bunches of *S. grandis* using TRIzol Reagent. The RNA integrity was examined using agarose gel electrophoresis and an Agilent 2,100 Bioanalyzer. The RNA concentration and purity were determined with a NanoDrop spectrophotometer based on the OD_{260}/OD_{280} and OD_{260}/OD_{230} ratios, respectively.

For cDNA library construction, mRNAs were enriched from the total RNAs using Oligo (dT) magnetic beads, after fragmentation into short segments using fragmentation buffer, used as templates to synthesize the first-strand cDNA with random hexamer primers and reverse transcriptase. Following second-strand cDNA synthesis, the fragments were end-repaired, a single adenine nucleotide was added to the 3′ end, and then ligated with sequencing adapters. Subsequently, the cDNA fragments were selected and PCR amplification was performed to obtain the final cDNA library. In total, six libraries were sequenced using an Illumina NovaSeq 6,000 platform.

Following processing for quality control, transcriptome assembly was performed with Trinity (Grabherr et al., 2011). Gene functions were annotated using the Kyoto Encyclopedia of Genes and Genomes (KEGG),[1] Gene Ontology (GO),[2] National Center for Biotechnology Information (NCBI) non-redundant (NR),[3] Swiss-Prot,[4] UniProtKB/TrEMBL,[5] and Clusters of Orthologous Groups of proteins (COG/KOG)[6] databases. The transcript levels of unigenes were calculated using the fragments per kilobase per million fragments (FPKM) method. Differential expression analysis between the HG and NG treatments was performed using DESeq2 (Love et al., 2014; Varet et al., 2016). The significantly differentially expressed genes (DEGs) were determined with the thresholds of $|\log_2(HG/NG)| \geq 1$ and FDR ≤ 0.05. Analyses of GO term and KEGG pathway enrichment of the DEGs were conducted with the R version 3.5.1 software using R package "clusterProfiler" (Yu et al., 2012) using the threshold of a corrected value of $p \leq 0.05$ to determine significant enrichment.

1 http://www.genome.jp/kegg/
2 www.geneontology.org
3 http://www.ncbi.nlm.nih.gov
4 http://www.expasy.ch/sprot
5 https://www.uniprot.org/uniprotkb?facets=reviewed%3Afalse&query=%2A
6 https://www.ncbi.nlm.nih.gov/COG/

Proteome analysis

Protein extraction and TMT labeling

Protein extraction from leaves of each sample for proteomic analysis was performed following a previously described method (Wang et al., 2021a). The protein concentration was calculated using the Bradford assay with bovine serum albumin as the standard (Bradford, 1976). Protein quality was examined using SDS-PAGE. The proteins extracted from each sample were digested with Trypsin Gold (Promega, Madison, WI, United States) at 37°C for 16 h. The peptide was desalted using a C18 cartridge and dried by vacuum centrifugation. The TMT labeling of peptides was performed using TMTsixplex reagents (TMTsixplex™ Isobaric Label Reagent Set, Thermo Fisher, Waltham, MA, Unites States) in accordance with the manufacturer's instructions.

HPLC fractionation and LC–MS/MS analysis

Fractionation of the TMT-labeled peptide mix was performed on a Rigol L3000 HPLC using a C18 column (4.6 mm × 250 mm, 5 μm), with a column temperature of 50°C. Mobile phases A and B were water solutions containing 2% and 98% acetonitrile (pH 10.0), respectively. The TMT-labeled peptide mix was dissolved in 1 ml A and the supernatant was used as the injection sample. The peptides were separated into 15 fractions based on the following elution gradient: 3% B, 5 min; 3–8% B, 0.1 min; 8–18% B, 11.9 min; 18–32% B, 11 min; 32–45% B, 7 min; 45–80% B, 3 min; 80% B, 5 min; 80%–5% B, 0.1 min; and 5% B, 6.9 min. Following drying under vacuum, the samples were reconstituted in 0.1% (v/v) formic acid (FA) in water and used for LC–MS/MS analyses.

For proteomic analyses, the EASY-nLC™ 1200 UHPLC system (Thermo Fisher) was employed to analyze the samples. A Q Exactive™ HF-X Orbitrap mass spectrometer (Thermo Fisher) was used to obtain the peptides in the data-dependent acquisition (DDA) mode. Briefly, 2 μg peptides were dissolved in 0.1% FA and injected into a C18 Nano-Trap column (2 cm × 100 μm, 5 μm). The peptides were separated with a ReproSil-Pur 120 C18-AQ analytical column (15 cm × 150 μm, 1.9 μm), with a linear gradient from 5% to 100% eluent B (0.1% FA in 80% acetonitrile) in eluent A (0.1% FA in H_2O) for 90 min with a flow rate of 600 nl/min.

For DDA, the MS conditions were set as follows: spray voltage 2.3 kV, capillary temperature 320°C, range of full MS scans 350–1,500 *m/z*, resolution 60,000 (200 *m/z*), automatic gain control (AGC) target value 3×10^6, and maximum ion injection time 20 ms. Based on the full MS scan, the 40 most abundant precursor ions were selected for higher-energy collisional dissociation fragment analysis. Resolution was set to 15,000, with AGC target value 1×10^5, maximum ion injection time 45 ms, normalized collision energy of peptide fragmentation 32%, intensity threshold 8.3×10^3, and dynamic exclusion parameter 60 s.

Protein identification and bioinformatics analysis

For protein identification, Proteome Discoverer (version 2.2) was used to search the raw data derived from MS analysis against the *S. grandis* transcriptome database (NCBI Sequence Read Archive database: PRJNA867365). The reporter quantification type was set as TMTsixplex. Trypsin was specified as an enzyme with a maximum of two missed cleavage sites. Mass tolerance for precursor and fragment was 10 ppm and 0.02 Da, respectively. Dynamic modifications were set as oxidation of methionine and TMTsixplex labeling lysine. The fixed modification was set as carbamidomethyl on cysteine. The parameters used for N-terminal modifications were acetylation of peptide N-terminus and TMTsixplex labeling N-terminus. Peptides were quantified according to the peak areas of corresponding mass reporters. The sequence of the peptides was determined by the mass-to-charge ratio of the peptide fragment ion peak. To reduce the frequency of false positives, the searched data were subsequently filtered through Proteome Discoverer. A peptide–spectrum match of 95% confidence was identified as credible, and the proteins containing at least one unique peptide were identified as credible. The false discovery rate (FDR) value was set at <5%.

Protein annotation was conducted using the GO, KEGG, and COG databases. For TMT quantification of proteins, the fold change (FC) between HG and NG samples was determined from the mean ratios of the TMT reporter ion intensities in the MS/MS spectra (*m/z* 126, 127-130C, 131) from all raw data sets. The significance of the differences was calculated using the corresponding value of p of each protein in two compared samples determined using Student's t-test. The thresholds of $FC \geq 1.2$ or $FC \leq 0.83$ and value of $p \leq 0.05$ were applied to detect differentially expressed proteins (DEPs). The functional subgroups and metabolic pathway enrichment for the DEPs were determined using the GO and KEGG databases.

Metabolite extraction and profiling

Metabolite extraction and profiling of the HG and NG samples was conducted by Wuhan Metware Biotechnology Co., Ltd. (Wuhan, China). Metabolite extraction, metabolite data acquisition, and assessment followed standard protocols as described in full previously (Chen et al., 2013; Li et al., 2021a). Briefly, the freeze-dried leaves were ground to powder using a mixer mill (MM 400, Retsch) with a zirconia bead (30 Hz, 1.5 min). One hundred milligram powder was extracted with 1.0 ml 70% aqueous methanol overnight at 4°C. After centrifugation (10,000 *g*, 10 min), the supernatant was filtrated using a micropore filter membrane (0.22 μm pore size) for LC–MS analysis. A LC-ESI-MS/MS system (UPLC, Shim-pack UFLC SHIMADZU CBM30A system, Kyoto, Japan; MS, Applied Biosystems 6,500 QTRAP, Framingham, MA, USA) was used for the UPLC–MS analysis. The chromatographic conditions were as follows: column, Waters ACQUITY UPLC HSS T3 C18 (1.8 μm, 2.1 mm × 100 mm); solvent system, water (0.04% acetic acid), acetonitrile (0.04% acetic acid); elution gradient, water: acetonitrile was 95:5 (V/V) at 0 min, 5:95 (V/V) at 11.0 and 12.0 min, 95:5 (V/V) at 12.1 and 15.0 min; flow rate 0.4 ml min^{-1}; column temperature 40°C, sample injection volume 2 μl. The mass spectrometry conditions: ESI temperature 500°C; ion spray voltage 5,500 V, curtain gas 25 psi; collision-activated dissociation was set to high. Instrument tuning was performed with 10 μmol/l polypropylene glycol solutions in triple quadrupole mode, and mass calibration was performed with 100 μmol/l polypropylene glycol solutions in linear ion trap mode. Metabolite identification was performed using the MWDB database.[7] Quantitative metabolite analysis was conducted using the multiple reaction monitoring mode (Fraga et al., 2010).

Pearson's correlation analysis was used to analyze the correlation among samples within a group to assess biological repeatability. Hierarchical cluster analysis was conducted with R software[8] to determine the metabolite accumulation patterns among different samples. Unsupervised principal component analysis (PCA) was employed to analyze the variability between the HG and NG groups as well as the three replications in each group.

The metabolite FC was calculated as the ratio of the mean metabolite abundance of HG plants relative to NG plants. The differentially accumulated metabolites (DAMs) were determined by a combination of the FC values and variable influence on projection (VIP) values. Partial least squares–discriminant analysis (PLS-DA), which is a supervised multivariate method, was used to maximize the metabolome differences between a pair of samples. The VIP value was used to check the relative importance of each metabolite in the PLS-DA model. Metabolites with $VIP \geq 1$ and $FC \geq 2$ or $FC \leq 0.5$ were identified as significantly differentially accumulated. The KEGG database was accessed for metabolite classification and pathway enrichment analysis.

Weighted correlation network analysis

Co-expression networks of DAMs and genes were constructed with R version 3.5.1 software using the weighted gene co-expression network analysis (WGCNA) package (version1.69), the parameter were set as follow: powerEstimate18, mergeCutHeight 0.25, and minModuleSize 50.

Results

Transcriptome profile of *Stipa grandis*

To investigate the molecular basis of the impact of heavy grazing on *S. grandis*, transcriptome profiling was performed and the transcript abundances were compared between the HG and

7 http://www.metwarebio.com/list/27.html
8 https://www.r-project.org/

NG samples (Supplementary Table S1). In total, 1,268 DEGs were detected, of which 860 were up-regulated and 408 were down-regulated (Figure 1A; Supplementary Table S2). Seventy-two GO terms were significantly enriched among the DEGs (Supplementary Figure S2). The most enriched KEGG pathways among these DEGs were the ribosome, plant–pathogen interaction, limonene and pinene degradation, lysine biosynthesis, ascorbate and aldarate metabolism, stilbenoid, diarylheptanoid, and gingerol biosynthesis; and histidine metabolism pathways (Figure 1B). The ribosome and plant–pathogen interaction pathways were significantly enriched ($p \leq 0.05$; Figure 1B).

TMT label-based proteome profile of *Stipa grandis*

The proteomic differences between the HG and NG groups were analyzed using TMT-based LC–MS/MS. Detailed information on the quality control is presented in Supplementary Figure S3. In total, 7,456 proteins were detected, of which 7,392 were annotated with functional terms from at least one of the GO, KEGG, and COG databases (Figure 2A; Supplementary Table S3). The coefficient of variation for most proteins was less than 20% (Figure 2B), which demonstrated high repeatability among the three biological replicates in each HG and NG group.

The abundance of 167 proteins differed significantly between the HG and NG groups with the thresholds FC ≥ 1.2 or FC ≤ 0.83, and value of $p \leq 0.05$. The expression of 135 DEPs was upregulated, whereas that of 32 DEPs was downregulated, in HG compared with NG (Figure 2C). Functional enrichment analysis showed that the DEPs were significantly enriched with 28 GO terms ($p \leq 0.05$, Figure 2D; Supplementary Table S4). The five most enriched terms in the molecular function category were catalase activity, peroxidase activity, heme binding, oxidoreductase activity, and D-aminoacyl-tRNA deacylase activity. The five most enriched terms associated with the biological process category were oxidation–reduction process, response to oxidative stress, metabolic process, response to stress, and regulation of translation. Only one enriched term, extracellular region, was associated with the cellular component category. The KEGG pathway analysis showed that the DEPs were enriched with 36 pathways (Figure 2E; Supplementary Table S4). It was noteworthy that glyoxylate and dicarboxylate metabolism, tryptophan metabolism, MAPK signaling pathway-plant, peroxisome, other types of *O*-glycan biosynthesis, and ribosome pathway were significantly enriched ($p \leq 0.05$).

Widely targeted metabolic profile analysis of *Stipa grandis*

To investigate the impact of heavy grazing on metabolic changes in *S. grandis*, the metabolome was detected in leaves from HG and NG samples of *S. grandis* by LC–MS/MS analysis. In total, 717 metabolites were identified (Supplementary Table S5). The metabolites were assigned to 32 types (Figure 3A). The five most abundant metabolite types were organic acids (10.32%), flavones (9.76%), amino acid derivatives (8.79%), nucleotides and derivatives (7.81%), and flavone C-glycosides (5.72%). The PCA revealed a clear separation between HG and NG plants (Figure 3B). Similarly, the heatmap cluster analysis clearly divided the six samples into two groups (Figure 3C). In addition, the Pearson's correlation analysis showed that the intragroup samples were significantly correlated (Supplementary Figure S4).

A total of 101 metabolites differed significantly between the HG and NG groups (Figure 4A). The metabolites were classified into 21 categories, and comprised 78 upregulated and 23 downregulated metabolites (Table 1). The top 20 DAMs were mainly associated with flavones, flavone C-glycosides, quinate and its derivatives, catechin derivatives, carbohydrates, benzoic acid derivatives, organic acids, indole derivatives, hydroxycinnamoyl

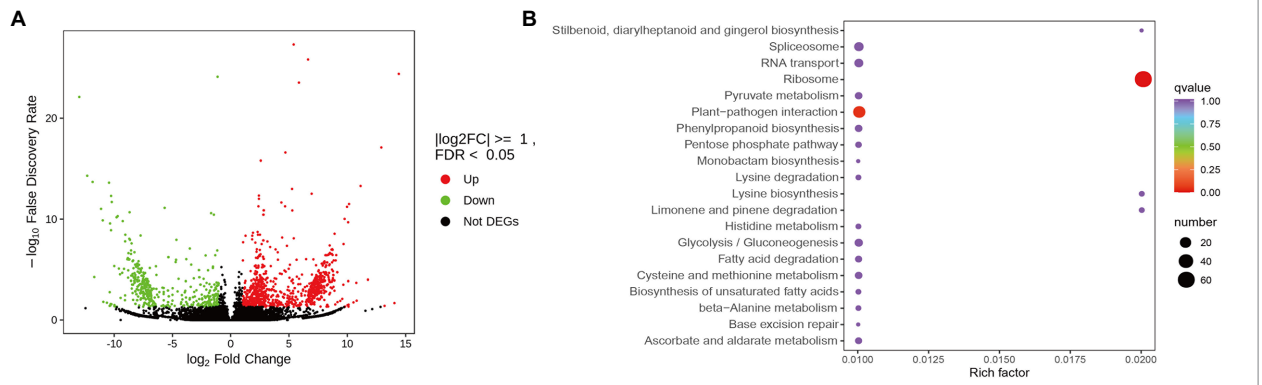


FIGURE 1
Differentially expressed genes (DEGs) between non-grazed (NG) and heavy-grazing (HG) plants of *Stipa grandis*. **(A)** Heatmap analysis of DEGs. **(B)** KEGG pathway enrichment analysis of DEGs.

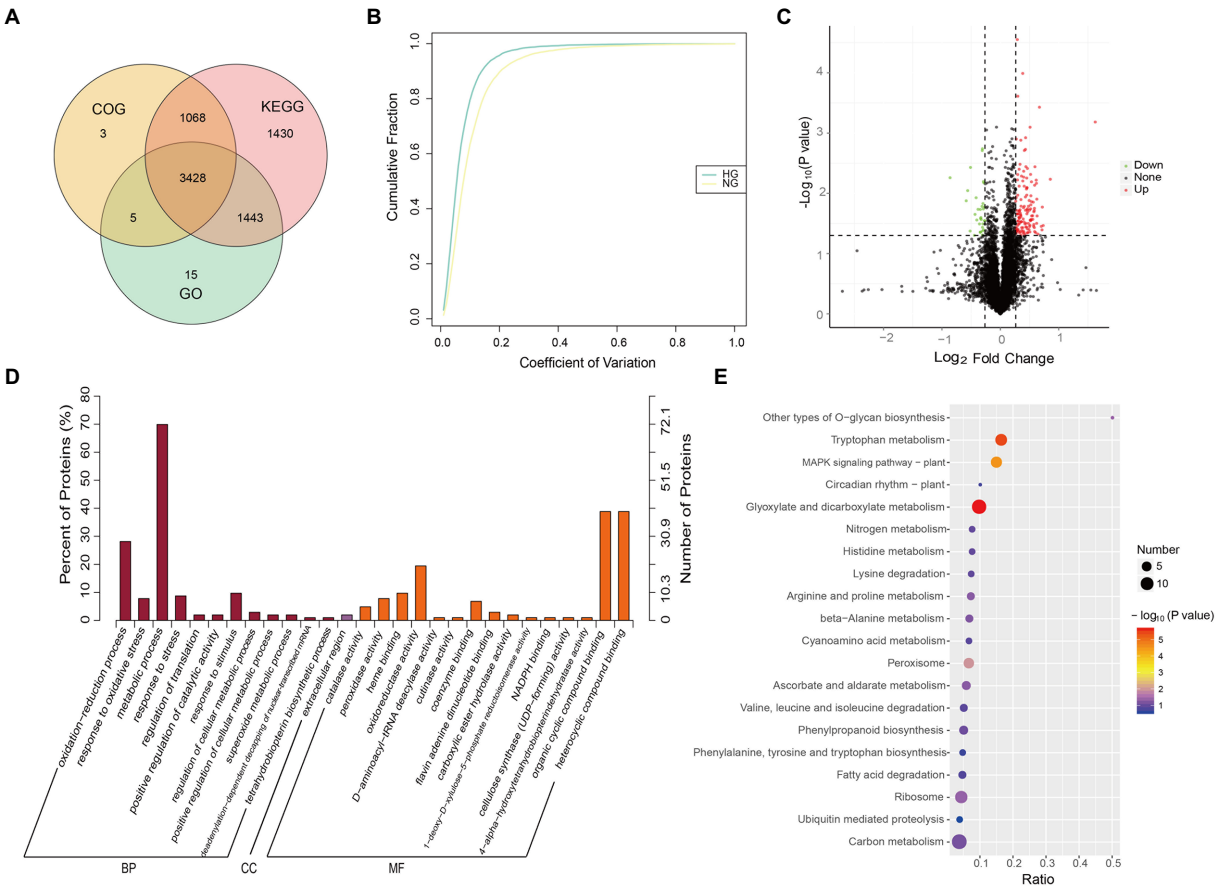


FIGURE 2
Proteome profile and differentially expressed proteins (DEPs) between NG and HG plants of *S. grandis*. **(A)** Venn diagram of proteins annotated from the GO, KEGG, and COG databases. **(B)** Coefficient of variation cumulative curve for NG and HG samples. **(C)** Quantitative analysis of the proteome of NG and HG samples. Green, red, and black represent down-regulated, up-regulated, and non-significantly changed proteins, respectively. **(D)** GO term enrichment analysis of DEPs between NG and HG samples. BP, biological process; CC, cellular component; and MF, molecular function. **(E)** KEGG pathway enrichment analysis of DEPs between NG and HG samples.

derivatives, amino acid derivatives, alcohols and polyols, and phenolamides (Supplementary Figure S5).

Based on KEGG pathway classification, the DAMs were involved in 64 pathways (Supplementary Table S6). The ten most enriched metabolic pathways were (in decreasing order) flavone and flavonol biosynthesis, tryptophan (Trp) metabolism, phenylpropanoid biosynthesis, stilbenoid, diarylheptanoid, and gingerol biosynthesis, styrene degradation, linoleic acid metabolism, renal cell carcinoma, serotonin receptor agonists/antagonists, galactose metabolism, and African trypanosomiasis (Figure 4B). Notably, the flavone and flavonol biosynthesis pathway was significantly enriched ($p < 0.05$) in the comparison of HG and NG.

KEGG pathways shared among DEGs, DEPs, and DAMs

The KEGG pathway enrichment analysis of the DEGs and DAMs showed that 26 KEGG pathways were shared among the significantly changed metabolites and transcripts (Figure 5A). Relatively high enrichment was observed among the stilbenoid, diarylheptanoid and gingerol biosynthesis, phenylpropanoid biosynthesis, tryptophan metabolism, pyruvate metabolism, degradation of aromatic compounds, indole alkaloid biosynthesis, pentose phosphate, and flavonoid biosynthesis pathways. Combined analysis of KEGG pathway enrichment among DEGs and DEPs revealed that the top shared pathways were glyoxylate and dicarboxylate metabolism, ribosome, biosynthesis of secondary metabolites, RNA degradation, and endocytosis (Figure 5B).

Joint analysis of phenylpropanoid biosynthesis, and flavone and flavonol biosynthesis

Given that flavone and flavonol biosynthesis as well as phenylpropanoid biosynthesis were among the three most highly

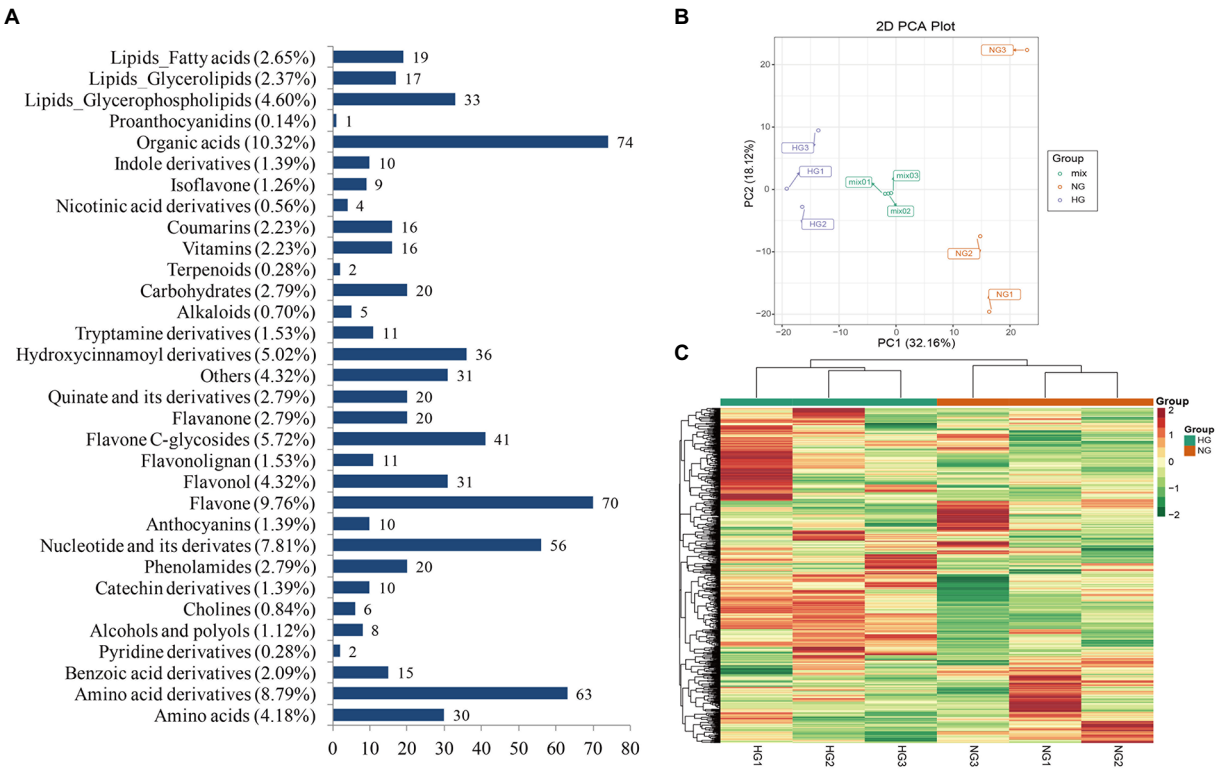


FIGURE 3
Qualitative and quantitative analysis of metabolites in leaves of *S. grandis*. **(A)** Classification of all identified metabolites. **(B)** Principal component analysis scatterplot of leaf samples from NG and HG plants. **(C)** Cluster analysis heatmap of all identified metabolites. The color scale indicates the abundance of each metabolite.

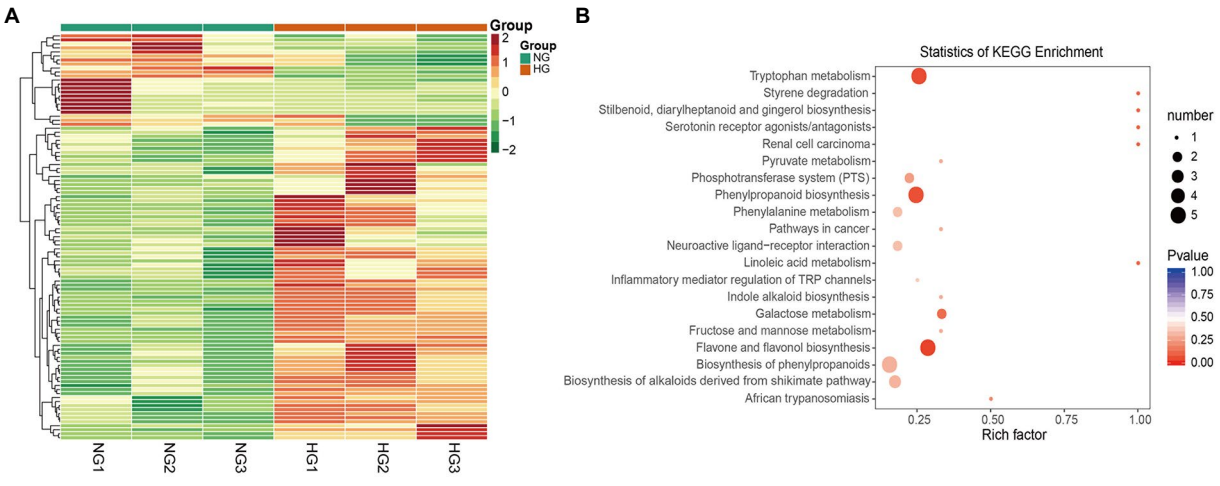


FIGURE 4
Differentially accumulated metabolites (DAMs) between NG and HG plants of *S. grandis*. **(A)** Cluster analysis heatmap of DAMs. **(B)** KEGG pathway enrichment analysis of DAMs.

enriched pathways in the metabolic profile analysis, an integrated analysis of DAMs, DEGs, and DEPs was performed (Figure 6). In the phenylpropanoid biosynthesis pathway, all of the five DAMs, namely *p*-coumaryl alcohol, chlorogenic acid, coniferylaldehyde, coniferyl alcohol, and isoeugenol, showed decreased accumulation in HG compared with NG samples. Four proteins,

comprising two peroxidases (POD), one beta-glucosidase (BGL), and one cinnamyl-alcohol dehydrogenase (CAD), showed enhanced expression in HG compared with NG leaves. Six genes were differentially expressed in HG compared with the NG group, among which two *trans*-cinnamate 4-monooxygenase-encoding genes (*CYP73A*) and two POD-encoding genes were down-regulated, and one CAD-encoding gene and one cinnamyl-alcohol dehydrogenase-encoding gene (*REF1*) were up-regulated. In the flavone and flavonol biosynthesis pathway, five metabolites, namely luteolin 7-*O*-glucoside, ayanin, isovitexin, kaempferide, and rhoifolin, showed differential accumulation patterns between the HG and NG groups, with higher accumulation observed in HG than in NG samples. However, no enrichment of significant DEGs or DEPs was observed. In addition, two DAMs (neohesperidin and catechin) that were mapped to the flavonoid biosynthesis pathway showed altered accumulation in HG samples. WGCNA was performed to identify potential genes co-expression with these DAMs. The *CAD* and *REF1*were distributed to module of yellow, and negatively correlated with DAMs of phenylpropanoid biosynthesis; while *CYP73As* and *PODs* were distributed to module grey60, and positively correlated with DAMs of phenylpropanoid biosynthesis (Supplementary Figure S6).

TABLE 1 Classification of significantly differentially accumulated metabolites between NG and HG plants of *S. grandis*.

Class	Total	Up-regulation	Down-regulation
Amino acid derivatives	8	7	1
Benzoic acid derivatives	3	2	1
Alcohols and polyols	3	3	0
Catechin derivatives	2	1	1
Phenolamides	3	3	0
Nucleotide and its derivates	2	0	2
Anthocyanins	1	1	0
Flavone	22	22	0
Flavonol	6	6	0
Flavone C-glycosides	17	17	0
Flavanone	2	2	0
Quinate and its derivatives	4	0	4
Others	2	0	2
Hydroxycinnamoyl derivatives	4	0	4
Tryptamine derivatives	4	4	0
Carbohydrates	3	1	2
Indole derivatives	1	1	0
Organic acids	5	1	4
Lipids_Glycerophospholipids	5	3	2
Lipids_Glycerolipids	1	1	0
Lipids_Fatty acids	3	3	0

Joint analysis of Trp metabolism

To evaluate the impact of heavy grazing on Trp metabolism, an integrated analysis of DAMs, DEPs, and DEGs associated with auxin biosynthesis was conducted (Figure 7). All five enriched metabolites involved in Trp metabolism were more highly accumulated in HG compared with NG samples, among which tryptamine (TAM), *N*-hydroxytryptamine, and

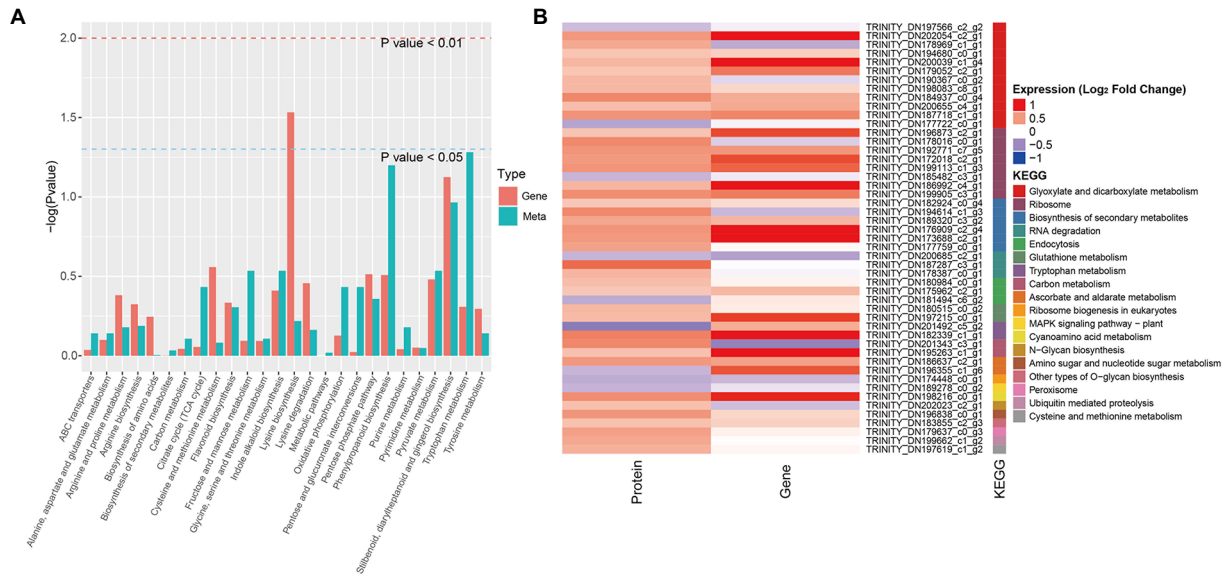


FIGURE 5
Enriched KEGG pathways common to DAMs, DEPs, and DEGs between NG and HG plants of *S. grandis*. **(A)** Histogram of KEGG pathways common to DAMs and DEGs. **(B)** Heatmap of KEGG pathways common to DEPs and DEGs.

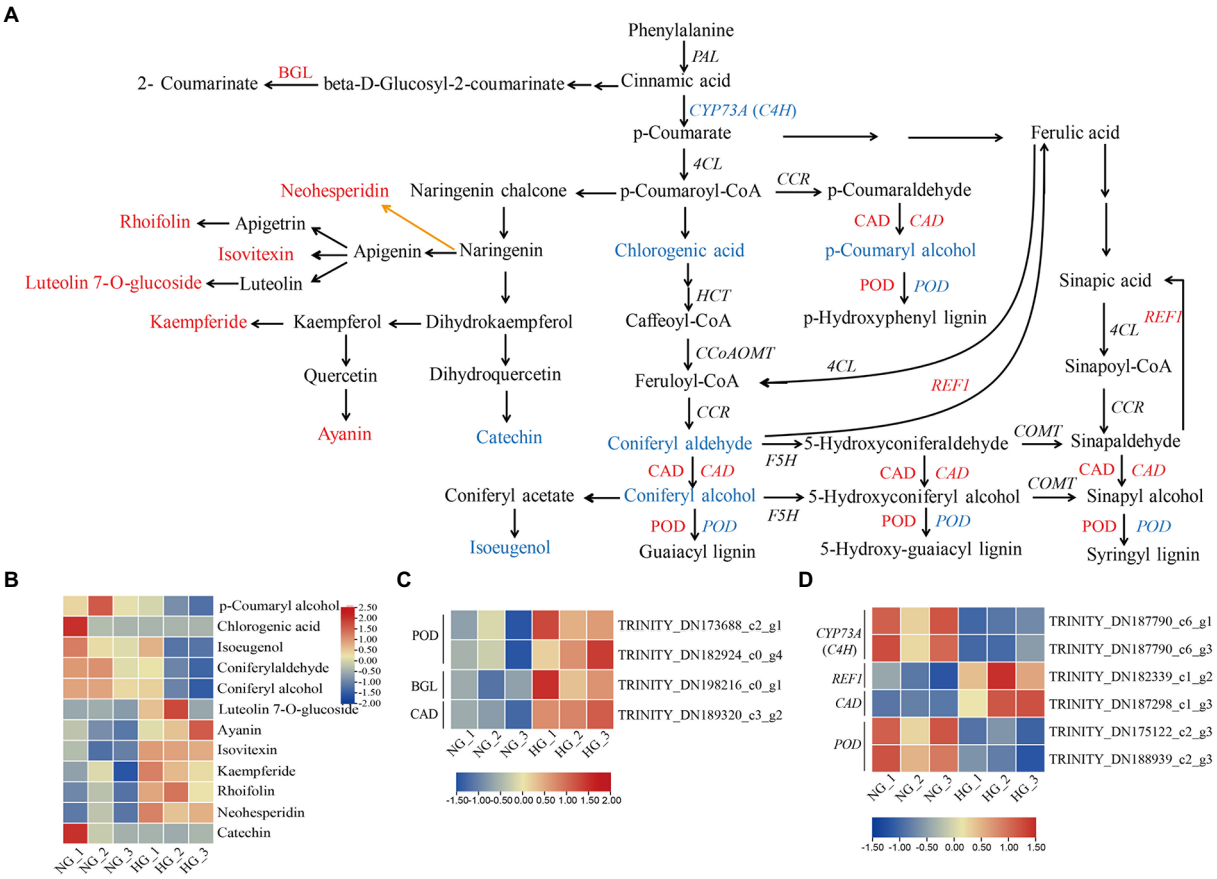


FIGURE 6
Phenylpropanoid biosynthesis, and flavone and flavonol biosynthesis pathways in *S. grandis* in response to heavy grazing. **(A)** Joint analysis of the DAMs, DEPs, and DEGs. Red and blue letters represent up- and down-regulation, respectively. Upper-case letters indicate proteins and italicized upper-case letters indicate genes. Orange line indicate biosynthetic step is not clear. The heatmap represents the corresponding expression levels of **(B)** DAMs, **(C)** DEPs, and **(D)** DEGs in the NG and HG groups. The color scale indicates the relative expression level.

indole-3-acetonitrile (IAN) are involved in Trp-dependent indole-3-acetic acid (IAA) biosynthesis. Three DEPs were associated with IAA biosynthesis, of which two aldehyde dehydrogenase (ALDH) family 2 proteins were up-regulated, and one amidase (AMI), which is involved in the conversion of indole-3-acetamide (IAM) into active IAA, was down-regulated in HG compared with NG samples. Three DEGs associated with *ALDH* genes were up-regulated in HG compared with NG samples. WGCNA analysis showed that one and two *ALDHs* distributed to yellow and green modules, respectively, and positively correlated with the DAMs of Trp metabolism (Supplementary Figure S7).

Discussion

Grazing is the most widespread and important land use in Inner Mongolia. The plant morphology, grassland productivity, and vegetation dynamics are affected by different grazing intensities (Liu et al., 2019). In this study, we investigated the mechanism for heavy-grazing-induced plant dwarfism by conducting an integrated analysis of the transcriptome, proteome, and metabolome. Based on the metabolic profile analysis, 101 metabolites were significantly differentially accumulated (VIP ≥ 1) between the HG and NG groups. These DAMs were assigned to 64 KEGG pathways among which the three most enriched pathways were flavone and flavonol biosynthesis, tryptophan metabolism, and phenylpropanoid biosynthesis.

Phenylpropanoid biosynthesis, and flavone and flavonol biosynthesis response to heavy grazing

The phenylpropanoid pathway is among the most extensively investigated specialized metabolic pathways (Dong and Lin, 2021). It gives rise to a vast variety of aromatic metabolites (Zhang and Liu, 2015). The lignin and flavonoid pathways are two major branches among the diverse branches of phenylpropanoid metabolism derived from the general phenylpropanoid pathway

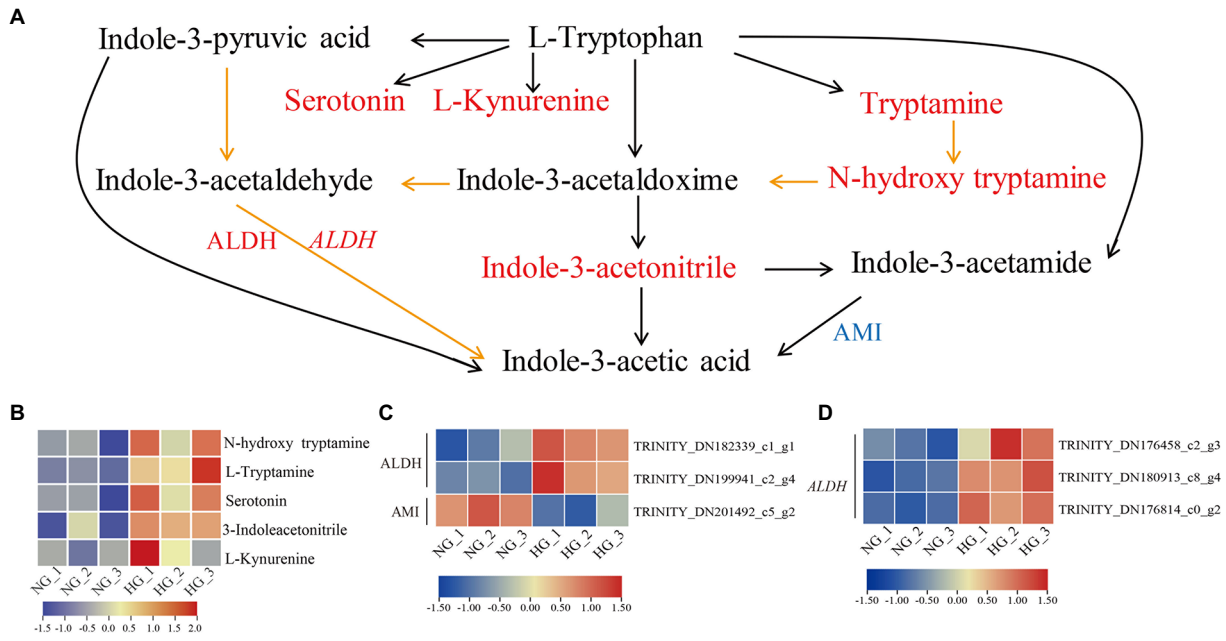


FIGURE 7
Tryptophan metabolism pathway in *S. grandis* in response to heavy grazing. **(A)** Joint analysis of the DAMs, DEPs, and DEGs. Red and blue letters represent up- and down-regulation, respectively. Upper-case letters indicate proteins and italicized upper-case letters indicate genes. Orange lines represent biosynthetic steps that are not clear. The heatmap represents the corresponding expression levels of **(B)** DAMs, **(C)** DEPs, and **(D)** DEGs in the NG and HG groups. The color scale indicates the relative expression level.

(Dong and Lin, 2021). In which, phenylalanine is catalyzed by phenylalanine ammonia lyase to form *trans*-cinnamic acid, which is hydroxylated by cinnamic acid 4-hydroxylase (C4H) to generate *p*-coumaric acid. In turn, *p*-coumaric acid is converted into *p*-coumaroyl-CoA catalyzed by 4-coumarate-CoA ligase (4CL; Dong and Lin, 2021).

Lignin, primarily a component of secondarily thickened cell walls, is essential for mechanical support during plant growth, facilitates the transportation of water and nutrients, and plays roles in plant resistance to pathogen attack (Boerjan et al., 2003). Lignin is derived mainly from three monolignols, namely *p*-coumaryl ($M1_H$), coniferyl ($M1_G$), and sinapyl ($M1_S$) alcohols, which are used to produce the three basic units of lignin, *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S), respectively (Boerjan et al., 2003; Dong and Lin, 2021). Alteration of monolignol metabolism strongly impacts on plant growth (Meents et al., 2018). However, the mechanism of dwarfism caused by suppression of monolignol biosynthesis is not completely elucidated, but several factors may contribute, such as water transport disruption, lack of cell proliferation or expansion resulting from failed synthesis of a phenylpropanoid-derived compound, or accumulation of a toxic pathway intermediate (Bonawitz and Chapple, 2010). In the present study, four of the five DAMs enriched in phenylpropanoid biosynthesis that were decreased in abundance in HG were involved in lignin biosynthesis, namely chlorogenic acid, coniferylaldehyde, and the monolignols coniferyl alcohol and *p*-coumaryl alcohol. Inconsistent with this finding, three enzymes were more highly accumulated in HG compared with NG. One of these enzymes was CAD, which functions in the final step of monolignol biosynthesis and is responsible for catalyzing the cinnamaldehydes to the corresponding alcohols, namely *p*-coumaryl alcohol, caffeyl alcohol, coniferyl alcohol, 5H coniferyl alcohol, and sinapyl alcohol (Boerjan et al., 2003; Dong and Lin, 2021). The other two enzymes were PODs, which are involved in the dehydrogenative polymerization of monolignols to form lignin (Fraser and Chapple, 2011). One CAD-encoding gene was up-regulated and two POD-encoding genes were downregulated. The inconsistent expression patterns of POD proteins and *POD* genes may result from the lower correlation between transcript and protein level due to post-transcriptional and post-translational regulations (Kumar et al., 2016; Xu et al., 2021). Alternatively, because these belong to a different member of POD family that may play different roles in lignin biosynthesis. C4H is a unique member of CYP73A, termed as CYP73A5 in *Arabidopsis thaliana* (Mizutani et al., 1997; Fraser and Chapple, 2011). Mutant of *C4H* displayed a dwarf phenotype in *A. thaliana* (Schilmiller et al., 2009). In the present study, down-regulated expression of two *C4H* members, i.e the *CYP73A* genes, in HG compared with NG, may be positively related to the heavy-grazing-induced dwarfism of *S. grandis*. Taken together, these results indicated that monolignol biosynthesis was inactivated in the heavy-grazing-induced dwarfism of *S. grandis*, and decreased production of monolignols may be associated with the plant morphological phenotype and reduction of lignin-mediated defense responses. But the regulatory mechanism of reduced monolignol

biosynthesis associated with the dwarfism of *S. grandis* remains unclear. Additionally, the activated expression of proteins or genes in HG may be associated with restored promotion of growth following the exclusion of grazing after the spring regreening stage. Further research on the molecular regulation of monolignol biosynthesis activity to adapt to heavy grazing is necessary.

Flavonoids, a major class of phenolic compounds, possess a basic C6–C3–C6 skeleton and are divided into numerous subgroups, such as flavanones, flavanols, flavones, and flavonols (Nakayama et al., 2019; Nabavi et al., 2020). Flavonoids perform a wide range of roles in plant development and environmental adaptation (Peer and Murphy, 2007; Thompson et al., 2010; Hichri et al., 2011). In particular, flavonoid biosynthesis is induced by oxidative stress (Nabavi et al., 2020), and flavonoids act as antioxidants to decrease oxidative damage triggered by reactive oxygen species under exposure to stresses (Nakabayashi and Saito, 2015; Dong and Lin, 2021). In the present study, all five DAMs involved in flavone and flavonol biosynthesis were more highly accumulated in the HG group than in NG samples. Except for ayanin, which has antimicrobial activity (French et al., 1991), the other four metabolites, namely luteolin 7-*O*-glucoside, isovitexin, kaempferide, and rhoifolin, exhibit antioxidant capabilities (Picerno et al., 2003; Eldahshan and Azab, 2012; Bian et al., 2013; Lama-Muñoz et al., 2019). This finding indicated that heavy-grazing-induced dwarfism may enhance tolerance to environmental stresses and might be mediated by flavonoid compounds through decreased oxidative damage. However, little is known about the underlying regulatory mechanism. Grazing caused multi-stress, involved in trampling/wounding, defoliation, saliva deposition, and manure deposition (Chen et al., 2009). Over or heavy grazing resulted in grassland degradation, including but not limited to reduced vegetation coverage and species diversity, declined plant growth, decreased soil coverage and water conservation, as well as changed soil microbial structure (Han et al., 2008; Yang et al., 2016). Whether these factors contribute to the changed accumulation pattern of flavonoids in dwarf plants, and which is the driving factor, is worth further confirmation by conducting the single or cross-stress experiments in *S. grandis*.

Trp metabolism response to heavy grazing

The most abundant and intensively studied endogenous auxin is IAA, which is essential for plant growth and development (Yue et al., 2014; Kasahara, 2016). In the Trp metabolism pathway, Trp is a precursor of auxin and various secondary metabolites (Maeda and Dudareva, 2012). The IAA is mainly synthesized by Trp-dependent and Trp-independent pathways (Mano and Nemoto, 2012). In the Trp-dependent pathway, IAA can be derived from the following pathways: Trp, IAM and IAA in the IAM pathway; Trp, indole-3-pyruvic acid (IPA), indole-3-acetaldehyde (IAAld) and IAA in the IPA pathway; Trp, indole-3-acetaldoxime (IAOx), IAN or IAAld, and IAA in the IAOx pathway; and Trp, TAM, *N*-hydroxytryptamine, IAOx, IAAld, and IAA in the TAM pathway (Woodward and Bartel, 2005; Mano and Nemoto, 2012). In the present study, TAM, *N*-hydroxytryptamine, and IAN showed elevated accumulation in heavy-grazing-induced dwarfed plants compared with NG plants, but the IAA content showed no obvious difference between the two groups. In addition, the expression level of AMI1, which converts IAM to IAA, was decreased in dwarfed plants. Fine-tuning of cellular auxin contents is important for plant survival (Casanova-Saez and Voss, 2019). We hypothesized that auxin contents are fine-tuned in heavily grazed plants and may play important roles in the regulation of plant development and responses to environmental stresses. In further studies, it will be important to explore the cross-talk of IAA with other phytohormones in the regulation of heavy grazing response.

Data availability statement

The mass spectrometry proteomics data presented in the study are deposited in the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the iProX partner repository, accession number PXD035914. The transcriptomic data presented in the study are deposited in the NCBI Sequence Read Archive (SRA) repository, accession number PRJNA867365.

Author contributions

DW and YD conceived and designed the experiments. DW and YW performed the experiments. DW, YW, TZ, and RW analyzed the data. DW drafted the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research was funded by the Natural Science Foundation of Inner Mongolia Autonomous Region (2020ZD06 and 2022LHMS03007), Key technology Project of Science and Technology Plan in Inner Mongolia Autonomous (2019GG009), Major special projects of science and technology in Inner Mongolia Autonomous Region (2019ZD008), and the Central Public-interest Scientific Institution Basal Research Fund (1610332020004).

Acknowledgments

We thank Robert McKenzie, PhD, from Liwen Bianji (Edanz) (www.liwenbianji.cn) for editing the English text of a draft of this 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.2022.995074/full#supplementary-material

References

Balcke, G. U., Bennewitz, S., Bergau, N., Athmer, B., Henning, A., Majovsky, P., et al. (2017). Multi-Omics of tomato glandular Trichomes reveals distinct features of central carbon metabolism supporting high productivity of specialized metabolites. *Plant Cell* 29, 960–983. doi: 10.1105/tpc.17.00060

Bian, Q. Y., Wang, S. Y., Xu, L. J., Chan, C. O., Mok, D. K., and Chen, S. B. (2013). Two new antioxidant diarylheptanoids from the fruits of Alpinia oxyphylla. *J. Asian Nat. Prod. Res.* 15, 1094–1099. doi: 10.1080/10286020.2013.816297

Boerjan, W., Ralph, J., and Baucher, M. (2003). Lignin biosynthesis. *Annu. Rev. Plant Biol.* 54, 519–546. doi: 10.1146/annurev.arplant.54.031902.134938

Bonawitz, N. D., and Chapple, C. (2010). The genetics of lignin biosynthesis: connecting genotype to phenotype. *Annu. Rev. Genet.* 44, 337–363. doi: 10.1146/annurev-genet-102209-163508

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248–254. doi: 10.1016/0003-2697(76)90527-3

Casanova-Saez, R., and Voss, U. (2019). Auxin metabolism controls developmental decisions in land plants. *Trends Plant Sci.* 24, 741–754. doi: 10.1016/j.tplants.2019.05.006

Chauffour, F., Bailly, M., Perreau, F., Cueff, G., Suzuki, H., Collet, B., et al. (2019). Multi-omics analysis reveals sequential roles for ABA during seed maturation. *Plant Physiol.* 180, 1198–1218. doi: 10.1104/pp.19.00338

Chen, W., Gong, L., Guo, Z., Wang, W., Zhang, H., Liu, X., et al. (2013). A novel integrated method for large-scale detection, identification, and quantification of widely targeted metabolites: application in the study of rice metabolomics. *Mol. Plant* 6, 1769–1780. doi: 10.1093/mp/sst080

Chen, J., Han, G., Shang, C., Li, J., Zhang, H., Liu, F., et al. (2015). Proteomic analyses reveal differences in cold acclimation mechanisms in freezing-tolerant and freezing-sensitive cultivars of alfalfa. *Front. Plant Sci.* 6:105. doi: 10.3389/fpls.2015.00105

Chen, J., Hu, X., Shi, T., Yin, H., Sun, D., Hao, Y., et al. (2020). Metabolite-based genome-wide association study enables dissection of the flavonoid decoration pathway of wheat kernels. *Plant Biotechnol. J.* 18, 1722–1735. doi: 10.1111/pbi.13335

Chen, S., Li, X. Q., Zhao, A., Wang, L., Li, X., Shi, Q., et al. (2009). Genes and pathways induced in early response to defoliation in rice seedlings. *Curr. Issues Mol. Biol.* 11, 81–100. doi: 10.21775/cimb.011.081

Chen, C., Zhou, G., Chen, J., Liu, X., Lu, X., Chen, H., et al. (2021). Integrated metabolome and transcriptome analysis unveils novel pathway involved in the formation of yellow Peel in cucumber. *Int. J. Mol. Sci.* 22:1494. doi: 10.3390/ijms22031494

Decourcelle, M., Perez-Fons, L., Baulande, S., Steiger, S., Couvelard, L., Hem, S., et al. (2015). Combined transcript, proteome, and metabolite analysis of transgenic maize seeds engineered for enhanced carotenoid synthesis reveals pleotropic effects in core metabolism. *J. Exp. Bot.* 66, 3141–3150. doi: 10.1093/jxb/erv120

Dong, N. Q., and Lin, H. X. (2021). Contribution of phenylpropanoid metabolism to plant development and plant-environment interactions. *J. Integr. Plant Biol.* 63, 180–209. doi: 10.1111/jipb.13054

Eldahshan, O. A., and Azab, S. S. (2012). Anti-inflammatory effect of Apigenin-7-neohesperidoside (Rhoifolin) in Carrageenin-induced rat Oedema model. *J. Appl. Pharm. Sci.* 2, 74–79. doi: 10.7324/JAPS.2012.2811

Fraga, C. G., Clowers, B. H., Moore, R. J., and Zink, E. M. (2010). Signature-discovery approach for sample matching of a nerve-agent precursor using liquid chromatography-mass spectrometry, XCMS, and Chemometrics. *Anal. Chem.* 82, 4165–4173. doi: 10.1021/ac1003568

Fraser, C. M., and Chapple, C. (2011). The phenylpropanoid pathway in Arabidopsis. *Arabidopsis Book.* 9:e0152. doi: 10.1199/tab.0152

French, C. J., Elder, M., Leggett, F., Ibrahim, R. K., and Neil Towers, G. H. (1991). Flavonoids inhibit infectivity of tobacco mosaic virus. *Can. J. Plant Pathol.* 13, 1–6. doi: 10.1080/07060669109500959

Gao, S. B., Mo, L. D., Zhang, L. H., Zhang, J. L., Wu, J. B., Wang, J. L., et al. (2018b). Phenotypic plasticity vs. local adaptation in quantitative traits differences of *Stipa grandis* in semi-arid steppe. *China Sci. Rep.* 8:3148. doi: 10.1038/s41598-018-21557-w

Gao, S., Zheng, Z., Wang, Y., Liu, L., Zhao, N., and Gao, Y. (2018a). Drought and grazing drive the retrogressive succession by changing the plant-plant interaction of the main species in Inner Mongolia steppe. *Ecol. Evol.* 8, 11954–11963. doi: 10.1002/ece3.4652

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883

Guo, Q., Li, X., Niu, L., Jameson, P. E., and Zhou, W. (2021). Transcription-associated metabolomic adjustments in maize occur during combined drought and cold stress. *Plant Physiol.* 186, 677–695. doi: 10.1093/plphys/kiab050

Han, J. G., Zhang, Y. J., Wang, C. J., Bai, W. M., Wang, Y. R., Han, G. D., et al. (2008). Rangeland degradation and restoration management in China. *Rangel. J.* 30, 233–239. doi: 10.1071/RJ08009

Hichri, I., Barrieu, F., Bogs, J., Kappel, C., Delrot, S., and Lauvergeat, V. (2011). Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway. *J. Exp. Bot.* 62, 2465–2483. doi: 10.1093/Jxb/Erq442

Jiang, Z., Jin, F., Shan, X., and Li, Y. (2019). iTRAQ-based proteomic analysis reveals several strategies to cope with drought stress in maize seedlings. *Int. J. Mol. Sci.* 20:5956. doi: 10.3390/ijms20235956

Job, D., Haynes, P. A., and Zivy, M. (2011). Plant proteomics. *Proteomics* 11, 1557–1558. doi: 10.1002/pmic.201190035

Kasahara, H. (2016). Current aspects of auxin biosynthesis in plants. *Biosci. Biotechnol. Biochem.* 80, 34–42. doi: 10.1080/09168451.2015.1086259

Kumar, D., Bansal, G., Narang, A., Basak, T., Abbas, T., and Dash, D. (2016). Integrating transcriptome and proteome profiling: strategies and applications. *Proteomics* 16, 2533–2544. doi: 10.1002/pmic.201600140

Lama-Muñoz, A., Del Mar Contreras, M., Espínola, F., Moya, M., De Torres, A., Romero, I., et al. (2019). Extraction of oleuropein and luteolin-7-O-glucoside from olive leaves: optimization of technique and operating conditions. *Food Chem.* 293, 161–168. doi: 10.1016/j.foodchem.2019.04.075

Li, X., Bi, X., An, M., Xia, Z., and Wu, Y. (2020). iTRAQ-based proteomic analysis of watermelon fruits in response to cucumber green mottle mosaic virus infection. *Int. J. Mol. Sci.* 21:2541. doi: 10.3390/ijms21072541

Li, S. P., Deng, B. L., Tian, S., Guo, M. X., Liu, H. X., and Zhao, X. S. (2021a). Metabolic and transcriptomic analyses reveal different metabolite biosynthesis profiles between leaf buds and mature leaves in *Ziziphus jujuba* mill. *Food Chem.* 347:129005. doi: 10.1016/j.foodchem.2021.129005

Li, X., Huang, Q., Mi, X., Bai, Y., Zhang, M., and Li, X. (2018). Grazing every month minimizes size but boosts photosynthesis in *Stipa grandis* in the steppe of Inner Mongolia. *China J. Arid Land.* 10, 601–611. doi: 10.1007/s40333-018-0011-4

Li, X. L., Liu, Z. Y., Wang, Z., Wu, X. H., Li, X. L., Hu, J., et al. (2015). Pathways of *Leymus chinensis* individual aboveground biomass decline in natural semiarid grassland induced by overgrazing: a study at the plant functional trait scale. *PLoS One* 10:e0124443. doi: 10.1371/journal.pone.0124443

Li, Q., and Song, J. (2019). Analysis of widely targeted metabolites of the euhalophyte Suaeda salsa under saline conditions provides new insights into salt tolerance and nutritional value in halophytic species. *BMC Plant Biol.* 19:388. doi: 10.1186/s12870-019-2006-5

Li, W., Wen, L., Chen, Z., Zhang, Z., Pang, X., Deng, Z., et al. (2021b). Study on metabolic variation in whole grains of four proso millet varieties reveals metabolites important for antioxidant properties and quality traits. *Food Chem.* 357:129791. doi: 10.1016/j.foodchem.2021.129791

Liu, M., Gong, J., Yang, B., Ding, Y., Zhang, Z., Wang, B., et al. (2019). Differences in the photosynthetic and physiological responses of *Leymus chinensis* to different levels of grazing intensity. *BMC Plant Biol.* 19:558. doi: 10.1186/s12870-019-2184-1

Liu, B., Zhang, N., Zhao, S., Chang, J., Wang, Z., Zhang, G., et al. (2015). Proteomic changes during tuber dormancy release process revealed by iTRAQ quantitative proteomics in potato. *Plant Physiol. Biochem.* 86, 181–190. doi: 10.1016/j.plaphy.2014.12.003

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

Lv, J., Zheng, T., Song, Z., Pervaiz, T., Dong, T., Zhang, Y., et al. (2021). Strawberry proteome responses to controlled hot and cold stress partly mimic post-harvest storage temperature effects on fruit quality. *Front. Nutr.* 8:812666. doi: 10.3389/fnut.2021.812666

Maeda, H., and Dudareva, N. (2012). The Shikimate pathway and aromatic amino acid biosynthesis in plants. *Annu. Rev. Plant Biol.* 63, 73–105. doi: 10.1146/annurev-arplant-042811-105439

Mann, M., Kulak, N. A., Nagaraj, N., and Cox, J. (2013). The coming age of complete, accurate, and ubiquitous proteomes. *Mol. Cell* 49, 583–590. doi: 10.1016/j.molcel.2013.01.029

Mano, Y., and Nemoto, K. (2012). The pathway of auxin biosynthesis in plants. *J. Exp. Bot.* 63, 2853–2872. doi: 10.1093/jxb/ers091

Meents, M. J., Watanabe, Y., and Samuels, A. L. (2018). The cell biology of secondary cell wall biosynthesis. *Ann. Bot.* 121, 1107–1125. doi: 10.1093/aob/mcy005

Mizutani, M., Ohta, D., and Sato, R. (1997). Isolation of a cDNA and a genomic clone encoding cinnamate 4-hydroxylase from Arabidopsis and its expression manner in planta. *Plant Physiol.* 113, 755–763. doi: 10.1104/pp.113.3.755

Nabavi, S. M., Samec, D., Tomczyk, M., Milella, L., Russo, D., Habtemariam, S., et al. (2020). Flavonoid biosynthetic pathways in plants: versatile targets for metabolic engineering. *Biotechnol. Adv.* 38:107316. doi: 10.1016/j.biotechadv.2018.11.005

Nakabayashi, R., and Saito, K. (2015). Integrated metabolomics for abiotic stress responses in plants. *Curr. Opin. Plant Biol.* 24, 10–16. doi: 10.1016/j.pbi.2015.01.003

Nakayama, T., Takahashi, S., and Waki, T. (2019). Formation of flavonoid Metabolons: functional significance of protein-protein interactions and impact on flavonoid Chemodiversity. *Front. Plant Sci.* 10:821. doi: 10.3389/fpls.2019.00821

Pandey, A., and Mann, M. (2000). Proteomics to study genes and genomes. *Nature* 405, 837–846. doi: 10.1038/35015709

Peer, W. A., and Murphy, A. S. (2007). Flavonoids and auxin transport: modulators or regulators? *Trends Plant Sci.* 12, 556–563. doi: 10.1016/j.tplants.2007.10.003

Picerno, P., Mencherini, T., Lauro, M. R., Barbato, F., and Aquino, R. (2003). Phenolic constituents and antioxidant properties of Xanthosoma violaceum leaves. *J. Agric. Food Chem.* 51, 6423–6428. doi: 10.1021/jf030284h

Qin, X., Yin, Y., Zhao, J., An, W., Fan, Y., Liang, X., et al. (2022). Metabolomic and transcriptomic analysis of *Lycium chinese* and *L. ruthenicum* under salinity stress. *BMC Plant Biol.* 22:8. doi: 10.1186/s12870-021-03375-x

Ren, W., Hou, X., Wu, Z., Kong, L., Guo, H., Hu, N., et al. (2018a). De novo transcriptomic profiling of the clonal *Leymus chinensis* response to long-term overgrazing-induced memory. *Sci. Rep.* 8:17912. doi: 10.1038/s41598-018-35605-y

Ren, W., Hu, N., Hou, X., Zhang, J., Guo, H., Liu, Z., et al. (2017). Long-term overgrazing-induced memory decreases photosynthesis of clonal offspring in a perennial grassland plant. *Front. Plant Sci.* 8:419. doi: 10.3389/fpls.2017.00419

Ren, W., Xie, J., Hou, X., Li, X., Guo, H., Hu, N., et al. (2018b). Potential molecular mechanisms of overgrazing-induced dwarfism in sheepgrass (*Leymus chinensis*) analyzed using proteomic data. *BMC Plant Biol.* 18:81. doi: 10.1186/s12870-018-1304-7

Saito, K., and Matsuda, F. (2010). Metabolomics for functional genomics, systems biology, and biotechnology. *Annu. Rev. Plant Biol.* 61, 463–489. doi: 10.1146/annurev.arplant.043008.092035

Schilmiller, A. L., Stout, J., Weng, J. K., Humphreys, J., Ruegger, M. O., and Chapple, C. (2009). Mutations in the cinnamate 4-hydroxylase gene impact metabolism, growth and development in Arabidopsis. *Plant J.* 60, 771–782. doi: 10.1111/j.1365-313X.2009.03996.x

Shi, T. T., Zhu, A. N., Jia, J. Q., Hu, X., Chen, J., Liu, W., et al. (2020). Metabolomics analysis and metabolite-agronomic trait associations using kernels of wheat (Triticum aestivum) recombinant inbred lines. *Plant J.* 103, 279–292. doi: 10.1111/tpj.14727

Thompson, E. P., Wilkins, C., Demidchik, V., Davies, J. M., and Glover, B. J. (2010). An Arabidopsis flavonoid transporter is required for anther dehiscence and pollen development. *J. Exp. Bot.* 61, 439–451. doi: 10.1093/jxb/erp312

Varet, H., Brillet-Gueguen, L., Coppee, J. Y., and Dillies, M. A. (2016). SARTools: a DESeq2-and EdgeR-based R pipeline for comprehensive differential analysis of RNA-Seq data. *PLoS One* 11:e0157022. doi: 10.1371/journal.pone.0157022

Wan, D. L., Wan, Y. Q., Hou, X. Y., Ren, W. B., Ding, Y., and Sa, R. L. (2015). De novo assembly and Transcriptomic profiling of the grazing response in *Stipa grandis*. *PLoS One* 10:e0122641. doi: 10.1371/journal.pone.0122641

Wang, F., Chen, N., and Shen, S. (2021a). iTRAQ-based quantitative proteomics analysis reveals the mechanism of Golden-yellow leaf mutant in hybrid paper mulberry. *Int. J. Mol. Sci.* 23:127. doi: 10.3390/ijms23010127

Wang, F., Ge, S. F., Xu, X. X., Xing, Y., Du, X., Zhang, X., et al. (2021b). Multiomics analysis reveals new insights into the apple fruit quality decline under high nitrogen conditions. *J. Agric. Food Chem.* 69, 5559–5572. doi: 10.1021/acs.jafc.1c01548

Woodward, A. W., and Bartel, B. (2005). Auxin: regulation, action, and interaction. *Ann. Bot.* 95, 707–735. doi: 10.1093/aob/mci083

Wu, J. B., Gao, Y. B., Bao, X. Y., Gao, H., Jia, M. Q., Li, J., et al. (2010a). Genetic diversity of *Stipa grandis* P Smirn populations across the species range in the Inner Mongolia plateau of China. *Biochem. Syst. Ecol.* 38, 471–477. doi: 10.1016/j.bse.2010.04.008

Wu, J. B., Gao, Y. B., Bao, X. Y., Gao, H., Jia, M. Q., Li, J., et al. (2010b). Genetic variation among *Stipa grandis* P. Smirn populations with different durations of fencing in the inner Mongolian steppe. *Rangel. J.* 32, 427–434. doi: 10.1071/Rj09038

Xiao, X. M., Wang, Y. F., Jiang, S., Ojima, D. S., and Bonham, C. D. (1995). Interannual variation in the climate and aboveground biomass of Leymus-Chinense steppe and Stipa-Grandis steppe in the Xilin River basin, inner-Mongolia. *China J. Arid Environ.* 31, 283–299. doi: 10.1016/S0140-1963(05)80033-3

Xu, X., Legay, S., Sergeant, K., Zorzan, S., Leclercq, C. C., Charton, S., et al. (2021). Molecular insights into plant desiccation tolerance: transcriptomics, proteomics and targeted metabolite profiling in *Craterostigma plantagineum*. *Plant J.* 107, 377–398. doi: 10.1111/tpj.15294

Xu, J. D., Yan, J. J., Li, W. J., Wang, Q. Y., Wang, C. X., Guo, J. X., et al. (2020). Integrative analyses of widely targeted metabolic profiling and Transcriptome data reveals molecular insight into Metabolomic variations during apple (*Malus domestica*) fruit development and ripening. *Int. J. Mol. Sci.* 21:4797. doi: 10.3390/ijms21134797

Yang, Z., Xiong, W., Xu, Y., Jiang, L., Zhu, E., Zhan, W., et al. (2016). Soil properties and species composition under different grazing intensity in an alpine meadow on the Eastern Tibetan plateau. *China Environ. Monit. Assess.* 188:678. doi: 10.1007/s10661-016-5663-y

Yang, F., Yang, T., Liu, K., Yang, Q., Wan, Y., Wang, R., et al. (2019a). Analysis of metabolite accumulation related to pod color variation of Caragana intermedia. *Molecules* 24:717. doi: 10.3390/molecules24040717

Yang, M., Yang, J., Su, L., Sun, K., Li, D., Liu, Y., et al. (2019b). Metabolic profile analysis and identification of key metabolites during rice seed germination under low-temperature stress. *Plant Sci.* 289:110282. doi: 10.1016/j.plantsci.2019.110282

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

Yue, J., Hu, X., and Huang, J. (2014). Origin of plant auxin biosynthesis. *Trends Plant Sci.* 19, 764–770. doi: 10.1016/j.tplants.2014.07.004

Zhan, Y., Wu, Q., Chen, Y., Tang, M., Sun, C., Sun, J., et al. (2019). Comparative proteomic analysis of okra (*Abelmoschus esculentus* L.) seedlings under salt stress. *BMC Genomics* 20:381. doi: 10.1186/s12864-019-5737-7

Zhang, Z., Gong, J., Wang, B., Li, X., Ding, Y., Yang, B., et al. (2020). Regrowth strategies of *Leymus chinensis* in response to different grazing intensities. *Ecol. Appl.* 30:e02113. doi: 10.1002/eap.2113

Zhang, Y., Li, Y., Li, W., Hu, Z., Yu, X., Tu, Y., et al. (2019). Metabolic and molecular analysis of nonuniform anthocyanin pigmentation in tomato fruit under high light. *Hortic. Res.* 6:56. doi: 10.1038/s41438-019-0138-2

Zhang, X., and Liu, C. J. (2015). Multifaceted regulations of gateway enzyme phenylalanine ammonia-lyase in the biosynthesis of phenylpropanoids. *Mol. Plant* 8, 17–27. doi: 10.1016/j.molp.2014.11.001

Zhang, Z., Wu, S., Stenoien, D. L., and Paša-Tolić, L. (2014). High-throughput proteomics. *Annu. Rev. Anal. Chem.* 7, 427–454. doi: 10.1146/annurev-anchem-071213-020216

Zhao, W., Chen, S. P., Han, X. G., and Lin, G. H. (2009). Effects of long-term grazing on the morphological and functional traits of *Leymus chinensis* in the semiarid grassland of Inner Mongolia. *Ecol. Res.* 24, 99–108. doi: 10.1007/s11284-008-0486-0

Zhao, N. X., Gao, Y. B., Wang, J. L., and Ren, A. Z. (2008). Population structure and genetic diversity of *Stipa grandis* P. Smirn, a dominant species in the typical steppe of Northern China. *Biochem. Syst. Ecol.* 36, 1–10. doi: 10.1016/j.bse.2007.08.004

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 11 October 2022
DOI 10.3389/fpls.2022.995754

OPEN ACCESS

EDITED BY
Linkai Huang,
Sichuan Agricultural University, China

REVIEWED BY
Danmei Liu,
Shanxi University, China
Peizhi Yang,
Northwest A&F University, China

*CORRESPONDENCE
Yuxiang Wang
wyx9868@163.com
Yongzhen Pang
pangyongzhen@caas.cn

†These authors have contributed equally to this work

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

RECEIVED 16 July 2022
ACCEPTED 01 September 2022
PUBLISHED 11 October 2022

CITATION
Li Q, Jiang W, Jiang Z, Du W, Song J, Qiang Z, Zhang B, Pang Y and Wang Y (2022) Transcriptome and functional analyses reveal *ERF053* from *Medicago falcata as* key regulator in drought resistances.
Front. Plant Sci. 13:995754.
doi: 10.3389/fpls.2022.995754

Transcriptome and functional analyses reveal *ERF053* from *Medicago falcata as* key regulator in drought resistances

Qian Li[1,2†], Wenbo Jiang[2†], Zhihu Jiang[1], Wenxuan Du[2], Jiaxing Song[3], Zhiquan Qiang[3], Bo Zhang[1], Yongzhen Pang[2*] and Yuxiang Wang[1*]

[1]Key Laboratory of Grassland Resources and Ecology of Western Arid Region, Ministry of Education, Key Laboratory of Grassland Resources and Ecology of Xinjiang, College of Grassland Science, Xinjiang Agricultural University, Urumqi, China, [2]Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing, China, [3]College of Grassland Agriculture, Northwest A&F University, Shanxi, China

Medicago falcata L. is an important legume forage grass with strong drought resistant, which could be utilized as an important gene pool in molecular breed of forage grass. In this study, *M. falcata* seedlings were treated with 400 mM mannitol to simulate drought stress, and the morphological and physiological changes were investigated, as well as the transcriptome changes of *M. falcata* seedlings at different treatment time points (0 h, 2 h, 6 h, 12 h, 24 h, 36 h and 48 h). Transcriptome analyses revealed four modules were closely related with drought response in *M. falcata* by WGCNA analysis, and four *ERF* transcription factor genes related with drought stress were identified (*MfERF053*, *MfERF9*, *MfERF034* and *MfRAP2.1*). Among them, *MfERF053* was highly expressed in roots, and MfERF053 protein showed transcriptional activation activity by transient expression in tobacco leaves. Overexpression of *MfERF053* in *Arabidopsis* improved root growth, number of lateral roots and fresh weight under drought, salt stress and exogenous ABA treatments. Transgenic *Arabidopsis* over-expressing *MfERF053* gene grew significantly better than the wild type under both drought stress and salt stress when grown in soil. Taken together, our strategy with transcriptome combined WGCNA analyses identified key transcription factor genes from *M. falcata*, and the selected *MfERF053* gene was verified to be able to enhance drought and salt resistance when over-expressed in *Arabidopsis*.

KEYWORDS

Medicago falcata*, Drought, RNA-Seq, WGCNA, *MfERF053

Introduction

Drought is one of the most serious abiotic factors that can cause significant damage to both agriculture, human and livestock populations (Gupta et al., 2020). Forage grass are subjected to damage caused by drought stress due to the lack of rainfall and insufficient water supply over a long period of time, which eventually leads to a significant decrease in yield (Ray et al., 2015). Therefore, in order to largely prevent environmental damage, plants themselves have evolved specific regulatory protection mechanisms (Pinhero et al., 1997). During drought stress, dramatic changes occur from biochemical and physiological to gene expression and metabolic processes (Rao et al., 2020; Waititu et al., 2021). Drought stress system promotes the production of metabolites such as proline, initiates the antioxidant defense system internal to scavenge increased reactive oxygen species (ROS), prevents cell damage by scavenging free radicals, reduces the degree of membrane lipid peroxidation, and maintains membrane integrity (Wang et al., 2009; Anjum et al., 2011; Wei et al., 2019). Many drought related genes had been identified and used as candidate genes in genetic engineering, such as *EDT1* (Yu et al., 2016), *MfNACsa* (Duan et al., 2017), *CYT75B1* (Rao et al., 2020), *MYB30* (Wen et al., 2021), *CBF4* (Haake et al., 2002) and *ERF172* (Zhao et al., 2020). Among them, many of them are transcription factors, and they play key roles in regulating the expression of downstream targeted genes and metabolic pathway.

AP2/ERF superfamily transcription factors are one of the largest plant-specific transcriptional regulator groups in plants, with a conserved AP2/ERF DNA-binding structural domain of 57-66 amino acids in size (Okamuro et al., 1997). Ethylene responsive factors (ERFs) belong to AP2/ERF superfamily, which participate in plant response to hormone and abiotic stress (Qiang et al., 2010; Gibbs et al., 2015; Jung et al., 2017). In rice, overexpression of *JERF3* and *OsERF115/AP2/EREBP110* can increase the soluble sugar and proline content of transgenic plants, up-regulated the expression of *P5CS* gene under drought stress, and improve the tolerance of crops to drought and osmotic stress (Thoenes et al., 2004; Zhang and Huang, 2010). In addition, overexpression of tomato ethylene response factor ERF (TSRF1) in rice can improve permeability and drought resistance through binding GCC box, and up-regulated the expression of MYB, MYC, proline synthesis and photosynthesis-related genes (Quan et al., 2010), activated the expression of the abscisic acid (ABA) synthesis gene *SDR*, thereby enhanced the sensitivity of transgenic rice to ABA. However, overexpression of *OsDERF1* down-regulates ethylene synthesis and negatively affects drought tolerance (Zhai et al., 2012). *OsERF71* positively regulates ABA signaling to alter root structure and impart drought tolerance (Dong-Keun et al., 2016; Li et al., 2018). *NtERF172* acts as a positive factor in drought stress tolerance, transgenic tobacco showed higher oxidase activity, and lower H_2O_2 accumulation, in part by regulating the dynamic balance of CAT-mediated H_2O_2, thereby exhibited greater drought tolerance (Zhao et al., 2020).

M. falcata is widely distributed in northern China, and most of them are wild resources with good drought resistance and good palatability, providing rich nutrients for cattle, sheep and other livestock (Yue and Zhou, 2004). By cross-pollination, *M. falcata* can be crossed with alfalfa to produce *Medicago varia* with stronger resistance and higher utilization value, therefore *M. falcata* is an important gene source for *Medicago* breeding (Wang et al., 2008; Kang et al., 2011). In *M. falcata*, some genes have been reported to be associated with abiotic stress response, including *MfNACsa* (Duan et al., 2017), *MfNAC3* (Qu et al., 2016), *MfUSP1* (Gou et al., 2020), and galactinol synthase gene 1 (*MfGolS1*) (Zhuo et al., 2013). At present, the functional study on *ERF* genes in response to drought are not clear in *M. falcata*.

In this study, we explored the physiological and molecular responses to drought stress in *M. falcata* seedlings, and the correlation analysis of the clustered modules with physiological indicators analyzed by WGCNA. Combined with these analysis, an *ERF* gene was found to be strongly correlated with drought-related module, suggesting a potential role in drought stress. Furthermore, we reported the functional characterization of *MfERF053* in conferring multiple resistances to abiotic stresses by over-expression in *Arabidopsis*.

Materials and methods

Plant materials and sample collection

Seeds of *M. falcata* were provided by the Key Laboratory of Grassland Resources and Ecology of Western Arid Region, Ministry of Education, College of Grass Industry, Xinjiang Agricultural University. The seedlings were grown at the Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing. To ensure the consistency of seed germination, seeds with uniform size and fullness were selected and gently scratched with knife. The seeds were then sterilized with 75% ethanol for 10 min, 5% sodium hypochlorite for 10 min, followed by wash with sterile water for 4-5 times. The seeds were then sow on 1/2 MS medium, and placed at 4 °C for 3 days, and then in an light incubator at 22 °C for germination (16 h light/8 h darkness). Seedlings were transferred into flasks containing different concentrations of mannitol for drought treatment. The concentrations of mannitol were 200 mM, 300 mM, 400 mM, 500 mM, 600 mM, and the treatment without mannitol was set as the control group (CK) in this study.

In this experiment, the treatment with 400 mM mannitol were selected for physiological index determination, and seven different treatment time points were selected with 15-day-old seedlings, these samples were collected at 0 h (CK), 2 h, 6 h, 12 h, 24 h, 36 h and 48 h. For each treatment time point, three

biological replicates were performed with 15 plants with whole seedlings for each replicate, samples were quickly collected and frozen in liquid nitrogen, and stored in a -80°C refrigerator.

Measurement of physiological index

The content of MDA, Proline, SOD, CAT and POD in 15-day-old seedlings as for transcriptome sequencing were measured according to the instruction manual as provided at the website (https://www.solarbio.com/), and they were measured by using spectrophotometry methods. Statistical analysis was performed by one-way analysis of variance (ANOVA) and Duncan multiple tests using SPSS 22.0.

RNA extraction and library construction for transcriptome analysis

The Eastep Super Total RNA Extraction Kit was used for RNA extraction (Promega Biotech, Shanghai, China). RNA quality was assessed on an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA) and checked using RNase free agarose gel electrophoresis. Twenty-one cDNA libraries were constructed, then the cDNA fragments were purified with QiaQuick PCR extraction kit (Qiagen, Venlo, The Netherlands), end repaired, poly(A) added, and ligated to Illumina sequencing adapters to sequence. The ligation products were size selected by agarose gel electrophoresis, PCR amplified, and sequenced using Illumina HiSeq2500 by Gene Denovo Biotechnology Co. (Guangzhou, China).

Raw data processing, sequence assembly and functional annotation

Raw read containing adapters or low quality bases will affect the following assembly and analysis, which were uploaded in NCBI SRA (http://www.ncbi.nlm.nih.gov/sra): SRR19146603-SRR19146623. Thus, to get high quality clean reads, reads were further filtered by fastp (Chen et al., 2018) (version 0.18.0). An index of the reference genome was built (Chen et al., 2020), and paired-end clean reads were mapped to the reference genome using HISAT2. 2.4 (Kim et al., 2015) with "-rna-strandness RF" and other parameters set as a default.

Identification of differentially expressed genes and PCA analysis

Analyses on differentially expressed genes were performed by DESeq2 software (Love et al., 2014) between two different groups, and by edgeR (Smyth, 2010) between two samples. The criteria of differentially expressed genes/transcripts (DEGs) screening was set as FDR<0.05 and $|\log_2 FC|>1$. Principal component analysis (PCA) was performed with R package models (http://www.r-project.org/).

WGCNA analysis

WGCNA (Weighted gene co-expression network analysis) is an analytical method to analyze the gene expression patterns of multiple samples (Zhang and Horvath, 2005), which allows clustering genes with similar expression patterns, and analyzing the association between modules and specific traits or phenotypes. Therefore, in this study, physiological indicator traits were analyzed in association with gene modules, using the R language package (Langfelder and Horvath, 2008). For annotation of the biological functions of the DEGs, GO and KEGG pathway enrichment analyses were performed with agriGO 2.0 (https://systemsbiology.cau.edu.cn/agriGOv2/) and KOBAS 3.0 (https://kobas.cbi.pku.edu.cn/), respectively.

Quantitative real-time PCR

Four *MfERF* genes were selected for validation using qRT-PCR. *Actin* gene was amplified as internal standard gene, and AlleleID 6 Tool was used to design the gene-specific primers (Table S1), relative expression level was normalized by comparing with control and calculated using $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001), qRT-PCR analysis program was as follows: one cycle at 95°C for 15 min, followed by 40 cycles at 95°C for 10 s, 60°C for 20 s and 72°C for 30 s.

Gene cloning and analyses on sequences and phylogenetic relationship

The full-length coding DNA sequence (CDS) of *MsERF053* was isolated from the roots of *Medicago falcata*, and cloned into pENTR vector for sequencing. The protein sequence of the homologous gene was selected by blast and multiple sequence comparisons. DNAMAN software was used to perform multiple sequence alignment. The phylogenetic tree was developed using protein sequences from *Medicago truncatula*, *Medicago sativa*, *Medicago ruthenica*, *Mucuna pruriens*, *Arabidopsis thaliana*, *Glycine sopa*, *Vigna angularis* and *Vigna radiata* with MEGA 7.0 (http://www.megasoftware.net) and visualize by using Evolview with bootstrap value of 1000 replications.

Transactivation assay

The open reading frame of *MfERF053* was also cloned with gene-specific primers by seamless cloning with KOD

polymerase. The transactivation construct was generated by inserting the full-length sequence of *MfERF053* into the *Kpn* I and *Xba* I sites of vector pCAMBIA1300BD. The BDGAL4 plasmid was used as control, and BDGAL4-MfERF053 recombinant plasmid were transformed into *A. tumefaciens* strain GV3101 and infiltrated into leaves of *N. benthamiana* as previously reported (Yang et al., 2021). After 48 h, the leave samples were taken separately, the proteins were extracted and the activity was determined.

Phenotypic analysis of transgenic *Arabidopsis* over-expression *MfERF053*

MfERF053 gene was cloned into the plant overexpression vector pB2GW7 and transformed into the *Agrobacterium tumefaciens* strain GV3101 for transformation in *Arabidopsis* by using floral dipping method. Three over-expression lines (OE19, OE20 and OE33) of T3 generation and wild type Columbia-0 (Col) plants were used for subsequent phenotype analyses. Seedlings were grown in soil at 24°C (16 h light/8 h darkness), 70%-80% relative humidity, and 400 $\mu mol \cdot m^{-2} \cdot s^{-2}$ light intensity.

We transferred four-day-old transgenetic and wild type *Arabidopsis* plants that were germinated on plates containing 1/2 MS medium to plates containing 1/2 MS medium supplied with different concentrations of mannitol (300 mM and 400 mM), NaCl (100 mM, 125 mM, 150 mM and 200 mM) and ABA (50 μM, 100 μM, 150 μM and 200 μM). Root length, lateral root number and fresh weight were measured after 10-day treatment. Each measurement contained 10 seedlings with triplicates. Seedlings grown in soil under normal conditions for 30 days were used for stress treatment. For drought stress, seedlings were grown without water for 15 days, then rewatered for 5 days. For salt stress, transgenic and wild type *Arabidopsis* plants were treated with 300 mM NaCl for 7 and 12 days, and the seedlings were photographed, respectively.

Results

Analysis of phenotypic and physiological indicators of *M. falcata* under drought stress

Initially, four-week-old *M. falcata* seedlings were subjected to mannitol treatment with concentration of 200 mM, 300 mM, 400 mM, 500 mM, and 600 mM. Mannitol treatment inhibited the growth of *M. falcata*, and the root length gradually decreased and the number of lateral roots significantly decreased with the increase of mannitol concentration (Figure S1A). These seedlings were more severely stressed and showed purplish-red root color and wilted leaves at mannitol concentration of 500 mM and 600 mM (Figure S1A), thus a relatively lower concentration of 400 mM mannitol was selected for subsequent experiments.

Seedlings were subjected to mannitol treatment (400 mM) at different time points (0 h, 2 h, 6 h, 12 h, 24 h, 36 h and 48 h). The seedlings grew normally at 2 h, and the leaves appear slightly wilted at 6 h, and started to lose water at 24 h, and finally the leaves wrinkled and severely wilted at 48 h (Figure S1B). It was clear that leaf wilting became more severe with stress time, and *M. falcata* responded to mannitol stress in a relatively short period of time.

To study the effect of drought stress on physiology changes in *M. falcata*, seedlings stressed with 400 mM mannitol at different time points were subjected to determine physiological indicators, including content of MDA, proline, and activity of SOD, POD and CAT (Figure 1). MDA content gradually increased from 2 h to about 55 μmol/g, and then gradually increased from 6 h to 48 h and reached the highest content at 48 h, indicating that the cells were most severely damaged by the treatment (Figure 1A). Drought stress significantly affected proline content in *M. falcata* plants at later stage (Figure 1B), with 10-fold increase ($p<0.01$) to 200 μg/g at 24 h, reaching maximum level at 48 h compared to untreated samples (Figure 1C). SOD activity of *M. falcata* plants increased with treatment time, began to increase significantly ($p<0.05$) after 2 h, and reached maximum level of approximately 280 U/g at 48 h, an increase of 186% compared to the control (Figure 1C). In terms of POD activity, the greatest increase was observed after 2 h treatment, with a slight increase at 6 h followed by a decrease at 12 h, with no significant differences from 24 h to 48 h (Figure 1D). CAT activity increased from 2 h to 36 h with its maximum value of 4200 $U/g^{-1}/min^{-1}$ at 12 h (Figure 1E). In conclusion, these physiological indicators of *M. falcata* responded to mannitol stress at different treatment time points with different degree.

De novo transcriptome assembly and functional annotation of unigenes

The 21 cDNA libraries yielded 8,624,521,500 bp raw reads (Table S2), the clean data is 8,537,671,848 bp, the Q30 before and after filtration was relatively high (around 94%), and the GC content was around 44% (Table S2), indicating good sequencing quality. In order to assemble the sequencing data, the reference genome of alfalfa ecotype Xinjiang Daye was selected as the reference genome for comparison, and the sequencing results were assembled and annotated as shown in Table S3. For all samples, the unmapped reads were only about 6.92-10.78%, the unique mapped reads were 39.86-42.11%, and the total mapped reads accounted for 89.22-93.68%, thus these data clearly indicated that the genome sequences of *M. falcata* have a very high matching rate with that of *M. sativa*, which can be used for subsequent analysis.

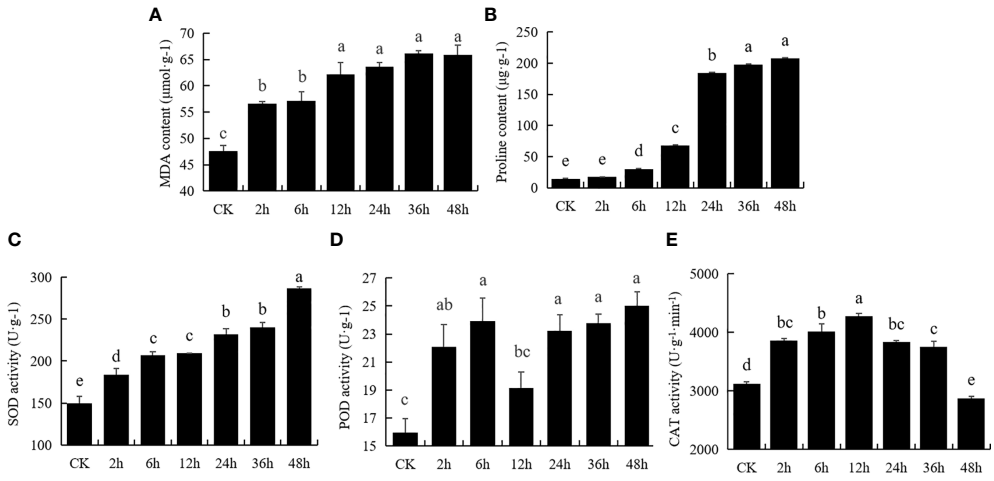


FIGURE 1
Physiological response of *Medicago falcata* to 400 mM mannitol treatment. Plant samples were collected under normal conditions and treated with 400 mM mannitol for 2 h, 6 h, 12 h, 24 h, 36 h and 48 h. The values are the average and error of three biological replicates. The same letters mean no significant difference, and different letters mean significant difference. **(A)**, MDA content; **(B)**, Proline content; **(C)**, SOD activity; **(D)**, POD activity; **(E)**, CAT activity.

Identification and analysis of expression pattern of DEGs

Principal component analysis (PCA) showed that samples from CK were clustered into a separate category and they were separated away from the samples under stress treatments (Figure S2). Samples from 2 h and 6 h treatments were closer than with other treatments, samples from 12 h treatments were clustered into a separate category, and samples from 24 h, 36 h and 48 h with long duration stress treatments were clustered into one category (Figure S2). These data together indicated that samples from different treatments tend to cluster differently.

The FPKM values for different samples were analyzed to investigate the changes in gene expression and to identify critical genes involved in drought stress in *M. falcata*. The volcano plot can be used to visualize the differentially expressed genes between treatment groups and control group. A total of 16,304 DEGs were obtained from transcriptome of *M. falcata* at 7 treatment time points. Among them, total 3,426, 3,632, 4,543, 3,944, 2,898, and 3,408 genes were down-regulated at 2 h, 6 h, 12 h, 24 h, 36 h and 48 h when compared with the control group at 0 h, respectively (Figure 2A). Meanwhile, 1,232, 2,165, 3,723, 4,156, 2,932, and 3,916 genes were down-regulated at 2 h, 6 h, 12 h, 24 h, 36 h and 48 h when compared with the control group at 0 h, respectively (Figure 2A).

Among these DEGs, 88 of them were found to be transcription factor genes belonging to 16 TF families (Figure 2B), including 14 *bHLH* genes, 11 *ERF* genes, 11 *bZIP* genes, and 10 *NAC* genes. Heat maps for the expression profiles of these 88 transcription factor genes under drought treatment showed that 73 of them were up-regulated compared with the control group (0 h) (Figure 2C, top), whereas 15 of them were down-regulated (Figure 2C, bottom). Among them, 10 out of 11 *ERF* genes were up-regulated (Figure 2C), indicating they may play leading roles in drought resistance in *M. falcata*.

WGCNA of common DEGs in drought stress

In order to further investigate potential key genes involved in drought response in *M. falcata*, the weighted gene co-expression network was constructed using WGCNA, resulting in eight modules (Figure S3). In order to explore the correlation of these clustered modules with the above-mentioned physiological indicators, correlation analysis was performed between the module eigenvalues with activity of POD, CAT, SOD, content of Pro and MDA. Heat map was used to display the top correlation coefficient (Figure 3). The correlation coefficient between the black module and three indicators (MDA, SOD and Pro) were relatively high with values of 0.83, 0.88 and 0.92 (significance of 4e-06, 2e-07 and 3e-09), respectively (Figure 3). Grey60 module was significantly correlated with SOD activity and Pro content with correlations coefficient of 0.72 and 0.88, respectively (Figure 3). Both the lightcyan and darkgreen modules showed highly significant positive correlations with CAT activity, with correlation coefficients of 0.75 and 0.64 (significance of 1e-04 and 0.002), respectively (Figure 3). These data clearly indicated that four modules, namely black module, Grey60 module, lightcyan

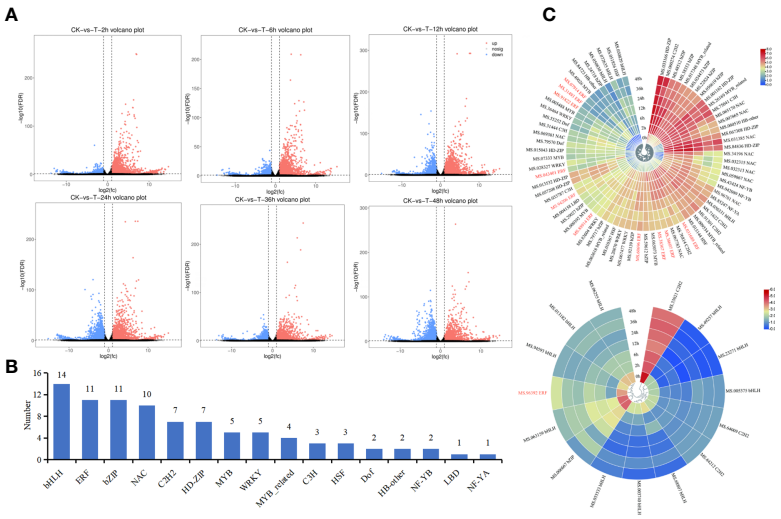


FIGURE 2
Analysis of differentially expressed genes at each time point under mannitol treatment. **(A)**, Volcano of difference among treatments. The horizontal coordinate indicates the $\log_2$(FC) of the difference between two groups, and the vertical coordinate indicates the negative $\log_{10}$ value of the FDR of the difference between the two groups. Red (up-regulated expression of group_2 relative to group_1) and blue (down-regulated expression) points indicate difference in gene expression (judged by FDR < 0.05, and more than two-fold difference), and black points indicate no difference. **(B)**, Statistic analysis of the number of different types of transcription factors. **(C)**, Heat map clustering of transcription factor expression. Red represents high expression and blue represents low expression. *MfERF* genes were highlighted in red.

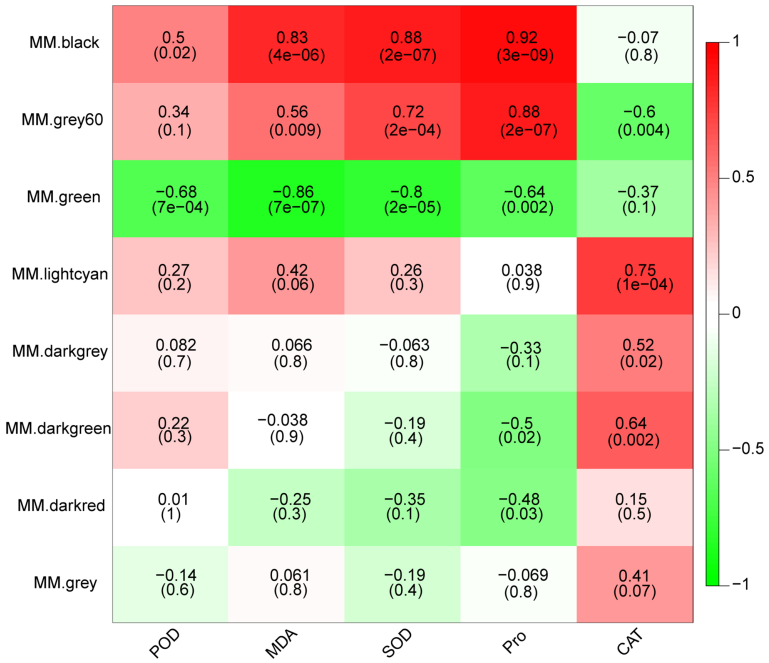


FIGURE 3
Character correlation diagram of physiological indicators. The horizontal coordinates indicate the character and the vertical coordinates indicate the module, which is plotted with Pearson correlation coefficient. Red represents positive correlation and green represents negative correlation. The darker the color, the stronger the correlation. The number in the brackets below represents significance *P*-value. The smaller the value, the stronger the significance.

module and darkgreen module, were likely contain genes involved in drought responsive in *M. falcata*.

Screening of candidate gene among DEGs

The four modules (black, Grey60, lightcyan and darkgreen) were up-regulated modules, but they differed in expression patterns, which was consistent with the expected expression pattern of concern. The darkgreen module was up-regulated in early stage (2 h and 6 h) but down-regulated in the late stage (12 h, 36 h and 48 h, Figure 4A). The lightcyan module was up-regulated in the early and middle stage (2 h, 6 h, 12 h) and down-regulated in the late stage (36 h and 48 h, Figure 4B). The black and grey60 modules showed a higher expression in the late stage (36 h and 48 h) than the earlier time points (Figure 4C, D). Furthermore, venn diagram with these four modules and *ERF* genes showed that 6 *ERF* genes were expressed in two out of four modules (Figure 4E). Among these 6 genes, both *MS.gene31493* and *MS.gene043401* from the darkgreen modules were two different transcripts for the same gene *MfERF053*, and *MS.gene38367* and *MS.gene38697* were for *MfRAP2-1* gene (Table S4), therefore, the expression of these four genes, namely *MfERF053*, *MfERF9*, *MfRAP2-1* and *MfERF034* were further verified by qRT-PCR. It was revealed that the expression of these four genes were up-regulated to different levels under drought treatment at different time points (Figure 4F), which is consistent the transcriptome sequencing data as verified by correlation analyses (Figure S4). Among these four genes, *MfERF053* showed the highest expression level, which was thus selected for further investigation.

Cloning, multiple sequence alignment and evolutionary tree analysis

The full-length open reading frame of *MfERF053* was cloned and the sequence was submitted to the National Center for Biotechnology Information (NCBI) under GenBank accession number of OM970125. Multiple sequence alignment of *MfERF053* with *ERF053* from other plant species showed that *MfERF053* shared 95%, 56%, 54% similarity with its homology genes from *M. truncatula*, *Vigna angularis* and *Glycine soja* at amino acid level (Figure 5A). All proteins from *M. truncatula*, *G. soja*, *V. angularis*, *Arabidopsis thaliana*, *Vigna radiata* var.

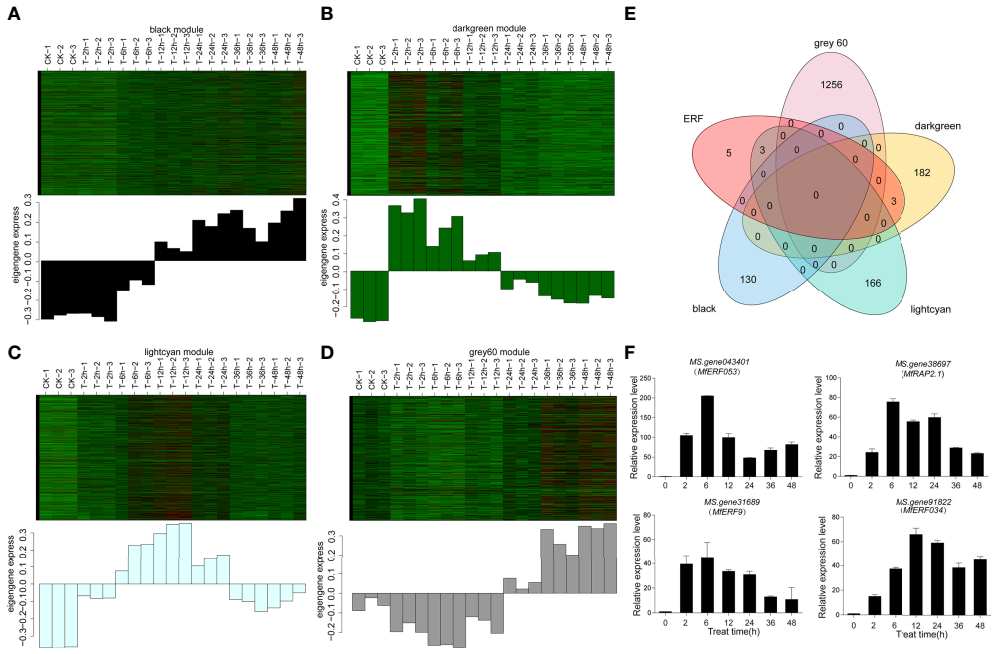


FIGURE 4
Differential candidate gene analysis. **(A–D)**, Heat map of gene expression pattern of each module. The above figure showed the expression level map of genes in modules in different samples; The following figure shows the characteristic values of modules in different samples. **(A–D)** represent darkgreen, lightcyan, black and gray60 module, respectively. **(E)**, The number of transcription factors differentially expressed in the five candidate combinations were screened by Venn map. Black, lightcyan, darkgreen and grey60 represent the three candidate modules screened by WGCNA, and *ERF* represents the number of *ERF* genes common shared by all treatments and controls. **(F)**, The expression level of four *MfERF* genes under mannitol treatment at different time point as detected by qRT-PCR.

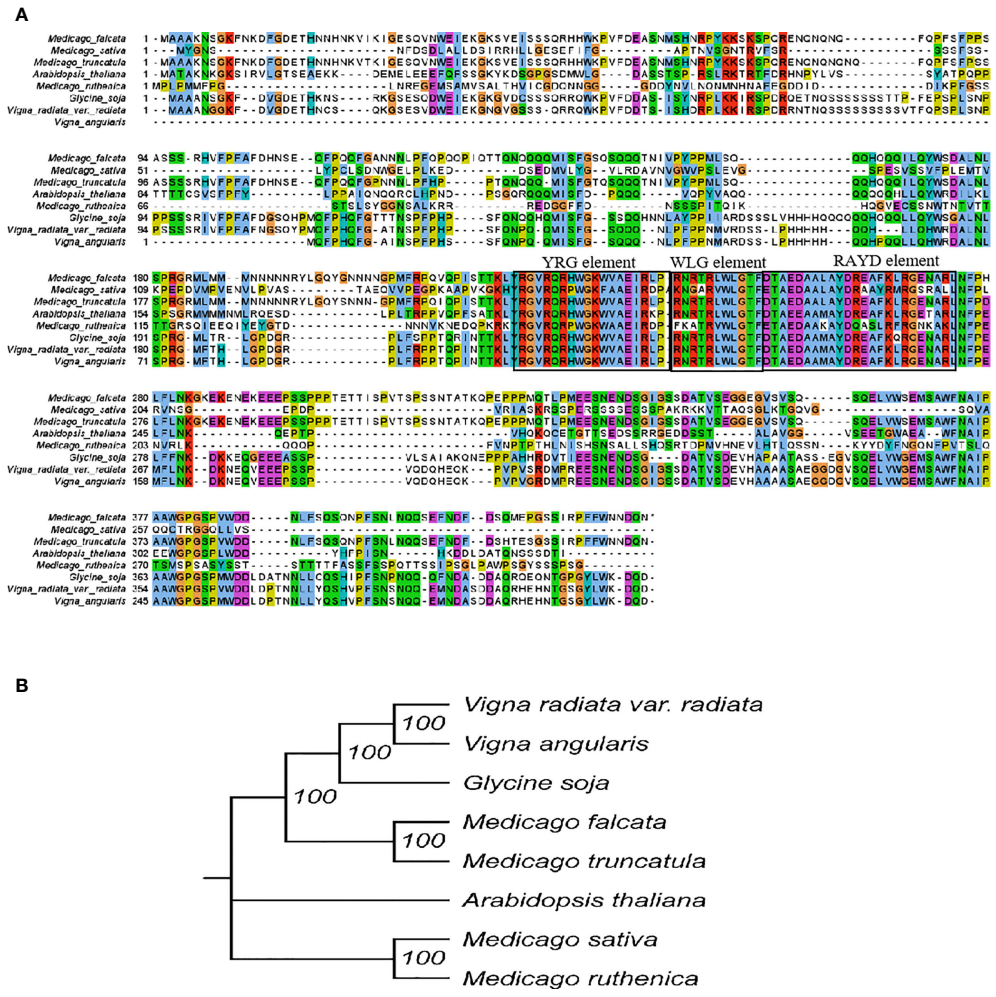


FIGURE 5
Sequence analysis of plant *ERF* genes. **(A)**, Multiple sequence alignment. The alignment were constructed by MEGA-X and visualized by Jalview. Residues with more than 50% similarity were shaded. Conserved regions (YRG element, WLG element and RAYD element) were indicated at the top. **(B)**, Phylogenetic analysis *ERF053* genes from different plant species.

radiata, *M. sativa*, *Medicago ruthenica* shared the characteristics features of DREB proteins, with conserved YRG, WLG, and RAYD motifs (Figure 5A).

Phylogenetic analyses revealed that *MfERF053* was clustered with *ERF053* from *M. truncatula* and *G. soja*, which was separated from those of *A. thaliana*, *M. sativa*, *M. ruthenica* with relatively low sequence similarity (51%, 49% and 48%, respectively, Figure 5B).

Expression profile and transcriptional activity of MfERF053

We analyzed relative expression level of *MfERF053* gene by qRT-PCR in roots, stems, leaves, flowers, branches and inflorescences of seedlings of *M. falacata* under normal growth condition. It was found that *MfERF053* was most highly expressed in roots, followed by in leaves and stems (Figure 6A). This finding suggests that *MfERF053* may be involved in mediating drought stress signaling through roots.

To investigate whether MfERF053 possesses transactivation activity, we generated a transactivation construct (35S:: GAL4DB- MfERF053) and expressed it in tobacco epidermal cells by *Agrobacterium*-medicated transformation, using GAL4DB as a negative control (Figure 6B). It was revealed that MfERF053 could activate the GAL4-responsive expression of the LUC reporter protein, and the relative luciferase activity for MfERF053 were about 9 times higher than the control (GALDB), indicating that MfERF053 acts as a transcription activator (Figure 6C).

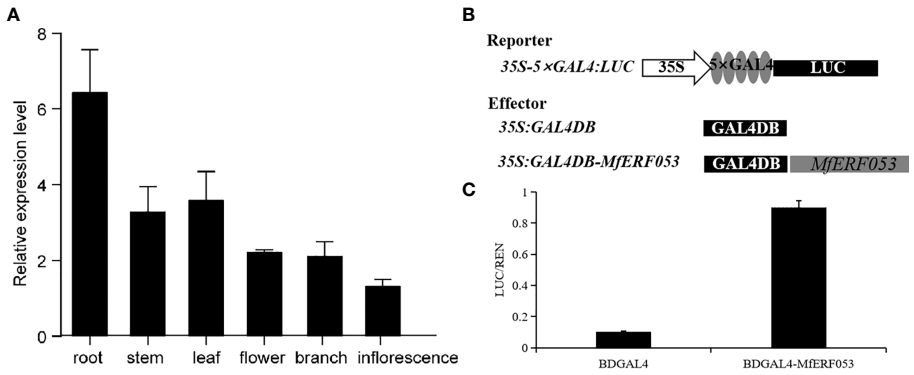


FIGURE 6
Characteristic of *MfERF053* gene. **(A)**, Validation of expression pattern of *M. falcata* in various tissues by qRT-PCR. **(B)**, Schematic diagram of reporter and effector for *MfERF053* gene for transactivation assay. **(C)**, *MfERF053* transcriptional activity analysis.

Over-expression of *MfERF053* in *Arabidopsis* enhanced resistance under plate culture condition

qRT-PCRs were performed to measure the expression levels of *MfERF053* in leaves of the transgenic *Arabidopsis* of the T3 generation. Among twenty-five lines that were detected, three independent transgenic lines with relatively high expression level were selected for further analyses (lines OE19, OE 20, and OE 33). No significant difference were observed between the transgenic line and the wild type control under normal plate culture condition (Figure S5). Both the transgenic lines and the wild-type *Arabidopsis* seedlings were grown on plates supplied with mannitol of different concentrations (300 mM and 400 mM), NaCl (100 mM, 125 mM, 150 mM, and 200 mM) and ABA (50 μM, 100 μM, 150 μM and 200 μM) (Figure 7, Figure S6, Figure S7, Figure S8).

When treated with 300 mM mannitol and 400 mM mannitol for 10 days, the root length of the overexpression lines did not show any difference from the control (Figure 7A, Figure S6), but the number of lateral roots and fresh weight increased significantly compared with the control ($p < 0.05$). Moreover, the increase in both lateral root number and fresh weight under 400 mM mannitol treatment were greater than those with 300 mM mannitol treatment (Figures 7B–D, Figure S6).

After 10 days of treatment with 100 mM NaCl, overexpression lines grew significantly better than the wild type, showing longer root length and increased number of lateral roots (Figure S7). After 10 days of treatment with 125 mM NaCl stress, the chlorophyll content of the leaves in the wild-type plants decreased with yellow leaves, while the overexpression lines showed green leaves and increased in root length, number of lateral roots and fresh weight (Figures 7E–H). After 10 days of treatment with 150 mM NaCl stress, the wild-type plants turned white, but the overexpression lines remained green, with significantly more green leaves and fibrous root than the wild-type (Figure S7). After 10 days of treatment with 200 mM NaCl stress, both the wild-type and overexpression lines turned white, and their growth were severely inhibited, but leave color of the overexpression lines changed to light purple and some leaves turned white (Figure S7).

In addition, the wild-type and overexpression *Arabidopsis* were grown on plates supplied with different concentrations of ABA treatment for 10 d (Fig 7I, Fig S7). It was found that the leaves of the overexpression lines showed darker green and the number of lateral root increased after ABA stress with low concentration (50 and 100 μM) for 10 d (Figure 7I, Figure S8). However, high ABA concentration inhibited the growth of *Arabidopsis* roots, but the transgenic lines grew better than the wild-type (Figure S8). The root length, lateral root number and fresh weight of the overexpression lines were significantly increased compared with the wild type under ABA treatment ($p < 0.01$), indicating that overexpression of *MfERF053* gene could improve the sensitivity to ABA on root growth in *Arabidopsis* under certain concentration (Figures 7J–L, Figure S8).

Over-expression of *MfERF053* in *Arabidopsis* enhanced resistance grown in soil

Three transgenic *Arabidopsis* strains (OE19, OE20 and OE33) showed no significant difference with the wild type control when plants grew in soil under natural watering condition (Figure 8A). But after 12 days of drought stress,

10.3389/fpls.2022.995754

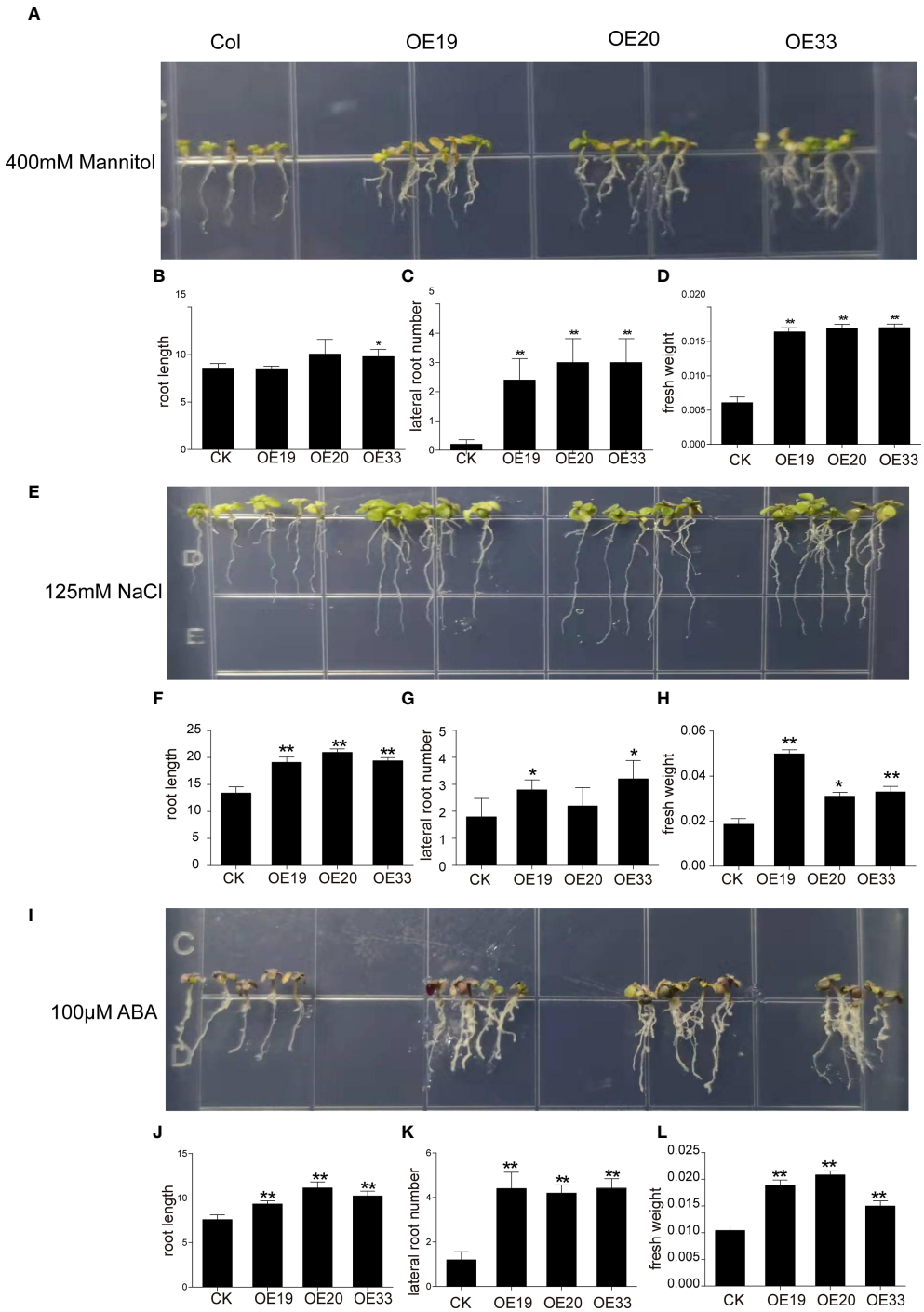


FIGURE 7
Evaluation of different resistance of *MfERF053* transgenic *Arabidopsis*. **(A, E, I)**, Plants overexpression *MfERF053* were treated with 400 mM mannitol, 100 mM NaCl and 100 μM ABA with wide type, respectively. **(B, F, J)**, Root length of different plant lines. **(C, G, K)**, Lateral root number of different plant lines. **(D, H, L)**, Fresh weight of different plant lines. Three replicates per treatment, 6 plants per replicate, * $p < 0.05$, ** $p < 0.01$, Duncan's *t*-test.

10.3389/fpls.2022.995754

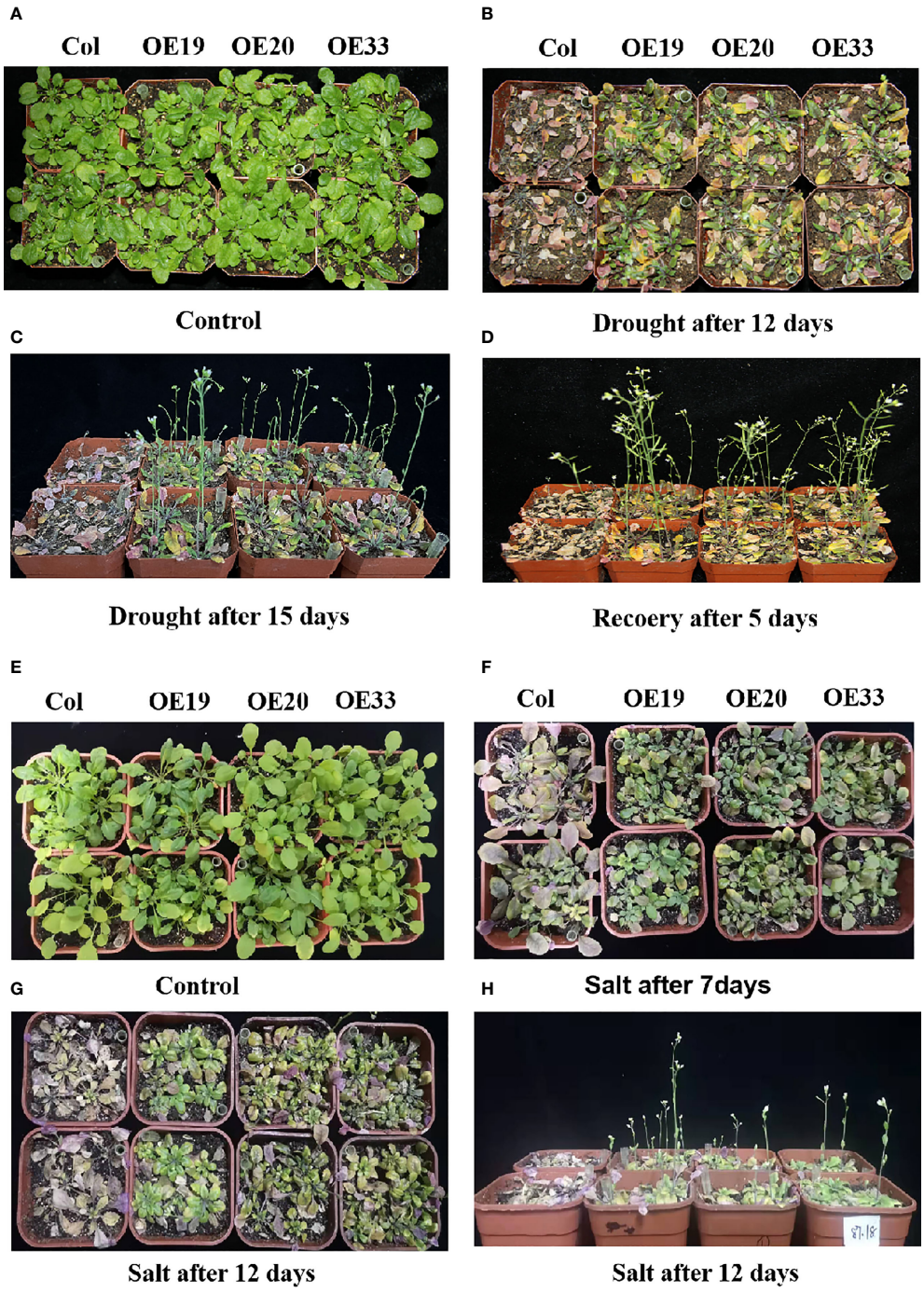


FIGURE 8
Phenotype of *MfERF053* transgenic overexpression *Arabidopsis* under drought NaCl and treatment in soil. **(A-D)**, The phenotype of wide type and transgenic of *Arabidopsis* under normal **(A)**, drought after 12 days **(B)**, drought after 15 days **(C)** and recover after 5 days **(D)**, respectively. **(E-H)**: The phenotype of wide type and transgenic of different strains of *Arabidopsis* under normal **(E)**, salt after 7 days **(F)**, salt after 12 days **(H)** and salt after 12 days **(I)**, respectively.

wild-type *Arabidopsis* was more severely stressed than the transgenic lines (Figure 8B). The wild type plants turned significantly purple, whereas some of the older leaves of the transgenic lines were still green or green-yellow (Figure 8B). When drought stress was continued for 15 d, the wild-type *Arabidopsis* plants withered and could not grow, while the transgenic *Arabidopsis* continued to grow and flowered, and the leaves remained green (Figure 8C). After rehydration for 5

days, the wild-type *Arabidopsis* plants did not recovered and dried completely, while the overexpression *Arabidopsis* plants grew well and the leaves became green and the inflorescences could develop into pods normally, indicating that the transgenic *Arabidopsis* plants over-expressing *MfERF053* have conferred drought resistance (Figure 8D).

Both the transgenic and the wild type *Arabidopsis* were treated with 300 mM NaCl in the soil (Figure 8E). After 7 d of treatment, the leaves of the wild-type plants turned purple with significantly reduced chlorophyll, while the transgenic *Arabidopsis* remained green with few leaves turning purple (Figure 8F). After continuous stress for 12 days (Figure 8G), the wild-type *Arabidopsis* withered and died, while the transgenic *Arabidopsis* remained green with some old leaves turning purple, and they could still flower normally and grow (Figure 8H), indicating that the transgenic *Arabidopsis* conferred salt resistance when compared with the wild-type.

Discussion

Among various environmental stress, drought is one of most serious stresses affecting the growth and development of plant. Drought stress triggers a series of responses from morphology to physiology, and to gene level. With the increase of drought stress, the antioxidant system of plants is destroyed, and the free radicals produced are greater than those cleared, resulting in excessive accumulation of ROS and membrane damage (Puyang et al., 2015; Nahar et al., 2017). In this study, the detailed information of physiological and transcriptome data of *M. falcata* under drought stress was provided. Under drought stress, the activities of SOD, POD and CAT increased (Figure 1), and the contents of proline and MDA increased, indicating that *M. falcata* has better ROS scavenging ability toward drought. It can be proven that the increase of these enzyme activities can eliminate stress-induced ROS and peroxides, inhibit plasma membrane peroxidation, and protect cells from damage (Zhang et al., 2004; Anjum et al., 2011; Koh et al., 2015; Quan et al., 2015; Xiong et al., 2022). The results of physiological indexes showed that *M. falcata* is capable of reducing the accumulation of harmful substances by regulating the activity of defense enzymes.

M. falcata is one of the candidate models to study abiotic stress response mechanism in legumes (Miao et al., 2015). The transcriptome analysis of drought provides a new biochemical and molecular mechanism for abiotic stress adaptation. Taking the genome of *M. sativa* cv 'Xinjiang Daye' as the reference genome, the tetraploid *M. falcata* transcriptome was sequenced, assembled and annotated, resulting in a total of 172,892 genes, with a comparison rate of more than 75% with the reference genome. The annotation of most genes is more similar to the annotation information of *M. truncatula*, indicating that the *M. falcata* transcriptome has good homologous sequence coverage, and it will also provide evidence for the expression of predicted genes in the genome of *M. truncatula* (Miao et al., 2015). Compared with the other transcriptome data for alfalfa with similar treatment (Luo et al., 2019), this study seems more complete. The number of genes obtained in each library in this study is higher than that produced in transcriptome of alfalfa under abiotic stress, which may be due to different samples used for sequencing. The sample for alfalfa is the roots, while the sample of this study is the whole seedlings. Therefore, the genes obtained for the transcriptome of alfalfa may be the genes specifically expressed in the roots, while in this study the obtained genes may be widely expressed in various parts.

In plants, some studies have screened stress resistance related genes through transcriptome sequencing (Gao et al., 2021; Li et al., 2022). In this study, we used *M. sativa* cv Xinjiang Daye as the reference genome, screened candidate genes through RNA-seq technology, and used expression patterns analysis and WGCNA analysis to mine hub *ERF* genes in response to dehydration of *M. falcata* (Figures 3, 4). The biological functions of these genes need to be further explored and verified, but at least qRT-PCR for these four candidate genes were consistent with the transcriptome data, which proves the reliability of WGCNA co-expression network analysis method. In conclusion, this strategy of screening functional genes related to drought stress is of great significance to the study of stress resistance in *M. falcata* as in other studies (Qin et al., 2020).

A drought responsive *MfERF053* gene in *M. falcata* was screened in this study by transcriptome sequencing (Figure 4). Previous studies have reported that *ERF* transcription factors act as both activators and repressors of transcriptional functions (Yant et al., 2010). MfERF053 was shown to have transcriptional activation activity based on transcriptional activity assays (Figure 6B), which is similar as that of the *SlERF3* gene (Pan et al., 2010). The activation activity of MfERF053 is consistent with its lack of an EAR (ERF-associated amphiphilic repression) inhibitory element, therefore, it is reasonable that MfERF053 is involved in different biological processes mainly in the form of activator.

Compared with the wild type, overexpression lines with higher expression level of *MfERF053* showed significant changes in root length and lateral roots as well as fresh weight (Figures 7, 8), indicating that overexpression of *MfERF053* had a significant effect on root growth in *Arabidopsis*. It has been shown that overexpression of the apple *MdERF11* and *MdERF106* genes (Han et al., 2020) can significantly improve the growth characteristics of plants to withstand abiotic stresses. Initial studies found that higher expression of *AtERF53* showed no significant difference in dehydration tolerance from wild type, it is speculated that the AtERF53 protein requires or may

require post-translational modifications controlled by another mechanism (Hsieh et al., 2013). In addition, other studies showed that the E3 ligase RGLG1 E3 can promote the degradation of *PP2CA* through an ABA dependent pathway, and the RING E3 ligase *RGLG2* interacts with *AtERF53* to negatively regulate drought stress response in *Arabidopsis* (Cheng et al., 2012; Belda-Palazon et al., 2019). *GmERF113* can enhance the drought resistance through activating the expression of *PR10-1* by binding to the GCC-box in soybean plants (Fang et al., 2022). Overexpression of the *GmERF75* gene in soybean hairy roots showed stronger growth than wild type under 100 μmol/L^{-1} ABA and 120 mM NaCl treatment, indicating that overexpression of *GmERF75* improved soybean tolerance to salinity and exogenous ABA (Zhao et al., 2019). All these results suggested that *ERF* genes have a conserved role in response to abiotic stresses in different plant species. In this study, overexpression of *MfERF053* improved the resistance of *Arabidopsis* to osmotic stress through the ABA transduction pathway. Nevertheless, the regulation mechanism of *MfERF053* on drought and salt resistance in both *M. falcata* and in transgenic *Arabidopsis* requires further investigation in the near future.

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

QL and WJ designed this experiment, performed the experiments, and drafted the manuscript. ZJ, WD, JS, ZQ analyzed experimental data, visualized and perform experiment. BZ, YW and YP revised the manuscript and directed the study. All authors contributed to the article and approved the submitted version.

Funding

This research was funded by National Nature Fund (31860675 and U1906201) and China Agriculture Research System of MOF and MARA (CARS-34).

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.2022.995754/full#supplementary-material

SUPPLEMENTARY TABLE 1
Primers used in this experiment.

SUPPLEMENTARY TABLE 2
Statistical information on transcriptome sequencing data.

SUPPLEMENTARY TABLE 3
Transcriptome alignment statistics.

SUPPLEMENTARY TABLE 4
Details of candidate genes.

SUPPLEMENTARY FIGURE 1
Phenotypic changes of *M. falcata* under mannitol treatment. **(A)**, Morphological response of *M. falcata* to different concentrations of mannitol. **(B)**, Effect of 400 mM mannitol drought stress at different times on leaf morphology of *M. falcata*.

SUPPLEMENTARY FIGURE 2
Principle component analysis of *M. falcata* at different times of 400 mM mannitol.

SUPPLEMENTARY FIGURE 3
Gene clustering map by WGCNA analysis. Left, Sample Hierarchical Clustering Tree. Right, Module level clustering diagram.

SUPPLEMENTARY FIGURE 4
Correlation analysis between qRNA-seq and RT-PCR data for the four candidate *MfERF* genes.

SUPPLEMENTARY FIGURE 5
Phenotype of *MfERF053* overexpression plants under normal plate growth condition. Col, the wide type control; Three over-expression lines: OE19, OE20 and OE33.

SUPPLEMENTARY FIGURE 6
Evaluation of overexpression plants *MfERF053* were treated with 300 mM mannitol. **(A)**, Overexpression of *Arabidopsis* 10-day phenotype under 300mM mannitol treatments. **(B)**, Root length of different plant lines. **(C)**, Lateral root number of different plant lines. **(D)**, Fresh weight of different plant lines.

SUPPLEMENTARY FIGURE 7
Evaluation of overexpression plants *MfERF053* were treated with different salt concentrations. **(A, E, I)**, Overexpression of *Arabidopsis* 10-day phenotype under 100 mM, 150 mM and 200 mM mannitol treatments, respectively. **(B, F, J)**, Root length of different plant lines. **(C, G, K)**, Lateral root number of different plant lines. **(D, H, L)**, Fresh weight of different plant lines.

SUPPLEMENTARY FIGURE 8

Evaluation of overexpression plants *MfERF053* were treated with different ABA concentrations. **(A, E, I)**, Overexpression of *Arabidopsis thaliana* 10-day phenotype under 100 μM, 150 μM and 200 μM ABA treatments, respectively. **(B, F, J)**, Root length of different plant lines. **(C, G, K)**, Lateral root number of different plant lines. **(D, H, L)**, Fresh weight of different plant lines.

References

Anjum, S. A., Wang, L., Farooq, M., Khan, I., and Xue, L. (2011). Methyl jasmonate-induced alteration in lipid peroxidation, antioxidative defence system and yield in soybean under drought. *J. Agron. Crop Sci.* 197 (4), 296–301. doi: 10.1111/j.1439-037X.2011.00468.x

Belda-Palazon, B., Julian, J., Coego, A., Wu, Q., Zhang, X., Batistic, O., et al. (2019). ABA inhibits myristoylation and induces shuttling of the *RGLG1* E3 ligase to promote nuclear degradation of *PP2CA*. *Plant J.* 98, 813–825. doi: 10.1111/tpj.14274

Cheng, M. C., Hsieh, E. J., Chen, J. H., Chen, H. Y., and Lin, T. P. (2012). Arabidopsis RGLG2, functioning as a RING E3 ligase, interacts with *AtERF53* and negatively regulates the plant drought stress response. *Plant Physiol.* 158, 363–375. doi: 10.1104/pp.111.189738

Chen, H. T., Zeng, Y., Yang, Y. Z., Huang, L. L., Tang, B. L., Zhang, H., et al. (2020). Allele-aware chromosome-level genome assembly and efficient transgene-free genome editing for the autotetraploid cultivated alfalfa. *Nat. Commun.* 11 (1), 2494–2505. doi: 10.1038/s41467-020-16338-x

Chen, S. F., Zhou, Y. Q., Chen, Y. R., and Gu, J. (2018). Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34 (17), 884–890. doi: 10.1093/bioinformatics/bty560

Dong-Keun, L., Suin, Y., Youn, S., and Hyunyong, K. (2016). Rice *OsERF71*-mediated root modification affects shoot drought tolerance. *Plant Signal Behav.* 12 (1), e1268311. doi: 10.1080/15592324.2016.1268311

Duan, M., Zhang, R., Zhu, F., Zhang, Z., Gou, L., Wen, J., et al. (2017). A lipid-anchored NAC transcription factor is translocated into the nucleus and activates glyoxalase I expression during drought stress. *Plant Cell* 29 (7), 1748–1772. doi: 10.1105/tpc.17.00044

Fang, X., Ma, J., Guo, F., Qi, D., Zhao, M. , Zhang, C., et al. (2022). The AP2/ERF *GmERF113* positively regulates the drought response by activating GmPR10-1 in soybean. *Int. J. Mol. Sci.* 23 (15), 8159–8181. doi: 10.3390/ijms23158159

Gao, G., Lv, Z., Zhang, G., Li, J., Zhang, J., and He, C. (2021). An ABA-flavonoid relationship contributes to the differences in drought resistance between different sea buckthorn subspecies. *Tree Physiol.* 41 (5), 744–755. doi: 10.1093/treephys/tpaa155

Gibbs, D., Conde, J. V., Berckhan, S., Mendiondo, G. M., Prasad, G., and Holdsworth, M. J. (2015). Group VII ethylene response factors coordinate oxygen and nitric oxide signal transduction and stress responses in plants. *Plant Physiol.* 169 (1), 23–31. doi: 10.1104/pp.15.00338

Gou, L., Zhuo, C., Lu, S., and Guo, Z. (2020). A universal stress protein from *Medicago falcata (MfUSP1)* confers multiple stress tolerance by regulating antioxidant defense and proline accumulation. *Environ. Exp. Bot.* 178, 104168–104179. doi: 10.1016/j.envexpbot.2020.104168

Gupta, A., Rico-Medina, A., and Cao-Delgado, A. I. (2020). The physiology of plant responses to drought. *Science* 368 (6488), 266–269. doi: 10.1126/science.aaz7614

Haake, V., Cook, D., Riechmann, J. L., Pineda, O., Thomashow, M. F., and Zhang, J. Z. (2002). Transcription factor *CBF4* is a regulator of drought adaptation in arabidopsis. *Plant Physiol.* 130 (2), 639–648. doi: 10.1104/pp.006478

Han, D., Han, J., Yang, G., Wang, S., Xu, T., and Li, W. (2020). An ERF transcription factor gene from *Malus* baccata l. borkh, *MbERF11*, affects cold and salt stress tolerance in *Arabidopsis*. *Forests* 11 (5), 514–529. doi: 10.3390/f11050514

Hsieh, E. J., Cheng, M. C., and Lin, T. P. (2013). Functional characterization of an abiotic stress-inducible transcription factor *AtERF53* in *Arabidopsis thaliana*. *Plant Mol. Biol.* 82 (3), 223–237. doi: 10.1007/s11103-013-0054-z

Jung, H., Chung, P. J., Park, S. H., Redillas, M., Kim, Y. S., Suh, J. W., et al. (2017). Overexpression of *OsERF48* causes regulation of *OsCML16*, a calmodulin-like protein gene that enhances root growth and drought tolerance. *Plant Biotechnol. J.* 15 (10), 1295–1308. doi: 10.1111/pbi.12716

Kang, Y., Han, Y., Torres-Jerez, I., Wang, M., Tang, Y., Monteros, M., et al. (2011). System responses to long-term drought and re-watering of two contrasting alfalfa varieties. *Plant J.* 68 (5), 871–889. doi: 10.1111/j.1365-313X.2011.04738.x

Kim, D., Landmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12 (4), 357–U121. doi: 10.1038/Nmeth.3317

Koh, J., Chen, G., Yoo, M. J., Zhu, N., Dufresne, D., Erickson, J. E., et al. (2015). Comparative proteomic analysis of *Brassica napus* in response to drought stress. *J. Proteome Res.* 14 (8), 3068–3081. doi: 10.1021/pr501323d

Langfelder, P., and Horvath, S. (2008). WGCNA: an r package for weighted correlation network analysis. *BMC Bioinf.* 9 (1), 559. doi: 10.1186/1471-2105-9-559

Li, J., Guo, X., Zhang, M., Wang, X., Zhao, Y., Yin, Z., et al. (2018). *OsERF71* confers drought tolerance *via* modulating ABA signaling and proline biosynthesis. *Plant Sci.* 270, 131–139. doi: 10.1016/j.plantsci.2018.01.017

Li, P., Lin, P., Zhao, Z., Li, Z., Liu, Y., Huang, C., et al. (2022). Gene co-expression analysis reveals transcriptome divergence between wild and cultivated sugarcane under drought stress. *Int. J. Mol. Sci.* 23 (1), 569–592. doi: 10.3390/ijms23010569

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

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

Luo, D., Zhou, Q., Wu, Y., Chai, X., Liu, W., Wang, Y., et al. (2019). Full-length transcript sequencing and comparative transcriptomic analysis to evaluate the contribution of osmotic and ionic stress components towards salinity tolerance in the roots of cultivated alfalfa (*Medicago sativa* L.). *BMC Plant Biol.* 19 (1), 32. doi: 10.1186/s12870-019-1630-4

Miao, Z., Xu, W., Li, D., Hu, X., Liu, J., Zhang, R., et al. (2015). Denovo transcriptome analysis of *Medicago falcata* reveals novel insights about the mechanisms underlying abiotic stress-responsive pathway. *BMC Genomics* 16, 818. doi: 10.1186/s12864-015-2019-x

Nahar, K., Hasanuzzaman, M., Alam, M. M., Rahman, A., Mahmud, J. A., Suzuki, T., et al. (2017). Insights into spermine-induced combined high temperature and drought tolerance in mung bean: osmoregulation and roles of antioxidant and glyoxalase system. *Protoplasma* 254 (1), 445–460. doi: 10.1007/s00709-016-0965-z

Okamuro, J. K., Caster, B., Villarroel, R., Van, M. M., and Jofuku, K. D. (1997). The AP2 domain of APETALA2 defines a large new family of DNA binding proteins in arabidopsis. *PNAS* 94 (13), 7076–7081. doi: 10.1073/pnas.94.13.7076

Pan, I. C., Li, C. W., Su, R. C., Cheng, C. P., Lin, C. S., and Chan, M. T. (2010). Ectopic expression of an EAR motif deletion mutant of *SlERF3* enhances tolerance to salt stress and ralstonia solanacearum in tomato. *Planta* 232 (5), 1075–1086. doi: 10.1007/s00425-010-1235-5

Pinhero, R. G., Rao, M. V., Paliyath, G., Murr, D. P., and Fletcher, R. A. (1997). Changes in activities of antioxidant enzymes and their relationship to genetic and paclobutrazol-induced chilling tolerance of maize seedlings. *Plant Physiol.* 114 (2), 695–704. doi: 10.1104/pp.114.2.695

Puyang, X., An, M., Han, L., and Zhang, X. (2015). Protective effect of spermidine on salt stress induced oxidative damage in two Kentucky bluegrass (*Poa pratensis* L.) cultivars. *Ecotoxicol. Environ. Saf.* 117, 96–106. doi: 10.1016/j.ecoenv.2015.03.023

Qiang, Z., Zhang, J., Gao, X., Tong, J., Xiao, L., Li, W., et al. (2010). The arabidopsis AP2/ERF transcription factor RAP2.6 participates in ABA, salt and osmotic stress responses. *Gene* 457 (1-2), 1–12. doi: 10.1016/j.gene.2010.02.011

Qin, T. Y., Chao, S., Zhen-Zhen, B., Wen-Jun, L., Peng-Cheng, L., Jun-Lian, Z., et al. (2020). Identification of drought-related co-expression modules and hub genes in potato roots based on WGCNA. *Acta Agron. Sin.* 46 (7), 19. doi: 10.3724/SP.J.1006.2020.94130

Quan, R., Hu, S., Zhang, Z., Zhang, H., Zhang, Z., and Huang, R. (2010). Overexpression of an ERF transcription factor *TSRF1* improves rice drought tolerance. *Plant Biotechnol. J.* 8 (4), 476–488. doi: 10.1111/j.1467-7652.2009.00492.x

Quan, W., Liu, X., Wang, H. , and Chan, Z. (2015). Comparative physiological and transcriptional analyses of two contrasting drought tolerant alfalfa varieties. *Front. Plant Sci.* 6 1256. doi: 10.3389/fpls.2015.01256

Qu, Y., Duan, M., Zhang, Z., Dong, J., and Wang, T. (2016). Overexpression of the *Medicago falcata* NAC transcription factor *MfNAC3* enhances cold tolerance in *Medicago truncatula*. *Environ. Exp. Bot.* 129, 67–76. doi: 10.1016/j.envexpbot.2015.12.012

Rao, M. J., Xu, Y., Tang, X., Huang, Y., Liu, J., Deng, X., et al. (2020). *CsCYT75B1*, a citrus CYTOCHROME P450 gene, is involved in accumulation of

antioxidant flavonoids and induces drought tolerance in transgenic arabidopsis. *Antioxid. (Basel)* 9 (2), 161–182. doi: 10.3390/antiox9020161

Ray, I. M., Han, Y. H., E, L., Meenach, C. D., Santantonio, N., Sledge, M. K., et al. (2015). Identification of quantitative trait loci for alfalfa forage biomass productivity during drought stress. *Crop Sci.* 55 (5), 2012–2033. doi: 10.2135/cropsci2014.12.0840

Smyth, G. K. (2010). edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26 (1), 139. doi: 10.1093/bioinformatics/btp616

Thoenes, E., Dixit, S., Pereira, A., Aharoni, A., Rooy, G. V., and Jetter, R. (2004). The SHINE clade of AP2 domain transcription factors activates wax biosynthesis, alters cuticle properties, and confers drought tolerance when overexpressed in *Arabidopsis*. *Plant Cell* 16 (9), 2463–2480. doi: 10.1105/TPC.104.022897

Waititu, J. K., Zhang, X., Chen, T., Zhang, C., Zhao, Y., and Wang, H. (2021). Transcriptome analysis of tolerant and susceptible maize genotypes reveals novel insights about the molecular mechanisms underlying drought responses in leaves. *Int. J. Mol. Sci.* 22 (13), 6980–7011. doi: 10.3390/ijms22136980

Wang, W. B., Kim, Y. H., Lee, H. S., Kim, K. Y., Deng, X. P., and Kwak, S. S. (2009). Analysis of antioxidant enzyme activity during germination of alfalfa under salt and drought stresses. *Plant Physiol. Bioch.* 47 (7), 570–577. doi: 10.1016/j.plaphy.2009.02.009

Wang, J. J., Yun, J. F., and Shi-Jie, L. V. (2008). Characteristic and utilizing values on the germplasm of *Medicago falcata*. *J. Inner Mongolia Agric. Univ. (Natural Sci. Edition)* 29 (1), 216–220. doi: 10.3901/JME.2008.09.177

Wei, T., Wang, Y., Xie, Z., Guo, D., Chen, C., Fan, Q., et al. (2019). Enhanced ROS scavenging and sugar accumulation contribute to drought tolerance of naturally occurring autotetraploids in poncirus trifoliata. *Plant Biotechnol. J.* 17 (7), 1394–1407. doi: 10.1111/pbi.13064

Wen, X., Geng, F., Cheng, Y., and Wang, J. (2021). Ectopic expression of *CsMYB30* from citrus sinensis enhances salt and drought tolerance by regulating wax synthesis in *Arabidopsis thaliana*. *Plant Physiol. Biochem.* 166, 777–788. doi: 10.1016/j.plaphy.2021.06.045

Xiong, Y., Fan, X. H., Wang, Q., Yin, Z. G., Sheng, X. W., Chen, J., et al. (2022). Genomic analysis of soybean PP2A-b family and its effects on drought and salt tolerance. *Front. Plant Sci.* 2 (12), 784038. doi: 10.3389/fpls.2021.784038

Yang, J., Li, Q., Du, W., Yao, Y., Shen, G., Jiang, W., et al. (2021). Genome-wide analysis of glycoside hydrolase family 35 genes and their potential roles in cell wall development in *Medicago truncatula*. *Plants (Basel)* 10 (8), 1639–1655. doi: 10.3390/plants10081639

Yant, L., Mathieu, J., Dinh, T. T., Ott, F., Lanz, C., Wollmann, H., et al. (2010). Orchestration of the floral transition and floral development in arabidopsis by the bifunctional transcription factor APETALA2. *Plant Cell* 22 (7), 2156–2170. doi: 10.1105/tpc.110.075606

Yue, X. Q., and Zhou, D. W. (2004). Good characteristics and its utilization of *Medicago falcata*. *Jilin J. Anim. Husb. Vet. Med.* 8, 26–28. doi: 10.3969/j.issn.1672-2078.2004.08.011

Yu, L. H., Wu, S. J., Peng, Y. S., Liu, R. N., Chen, X., Zhao, P., et al. (2016). *Arabidopsis EDT1/HDG11* improves drought and salt tolerance in cotton and poplar and increases cotton yield in the field. *Plant Biotechnol. J.* 14 (1), 72–84. doi: 10.1111/pbi.12358

Zhai, Y., Li, J. W., Li, X. W., Lei, T. T., and Wang, Q. Y. (2012). Isolation and characterization of a novel transcriptional repressor *GmERF6* from soybean. *Chem. Eng. J.* 57 (1), 624–629. doi: 10.1007/s10535-012-0146-7

Zhang, B., and Horvath, S. (2005). A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol.* 4 (1), 1–40. doi: 10.2202/1544-6115-1128

Zhang, Z., and Huang, R. (2010). Enhanced tolerance to freezing in tobacco and tomato overexpressing transcription factor *TERF2/LeERF2* is modulated by ethylene biosynthesis. *Plant Mol. Biol.* 73 (3), 241–249. doi: 10.1007/s11103-010-9609-4

Zhang, H., Huang, Z., Xie, B., Chen, Q., Tian, X., Zhang, X., et al. (2004). The ethylene, jasmonate, abscisic acid and NaCl-responsive tomato transcription factor *JERF1* modulates expression of GCC box-containing genes and salt tolerance in tobacco. *Planta* 220 (2), 262–270. doi: 10.1007/s00425-004-1347-x

Zhao, Q., Hu, R. S., Liu, D., Liu, X., Wang, J., Xiang, X. H., et al. (2020). The AP2 transcription factor *NtERF172* confers drought resistance by modifying *NtCAT*. *Plant Biotechnol. J.* 18 (12), 2444–2455. doi: 10.1111/pbi.13419

Zhao, M. J., Yin, L. J., Liu, Y., Ma, J., Zheng, J. C., Lan, J. H., et al. (2019). The ABA-induced soybean ERF transcription factor gene *GmERF75* plays a role in enhancing osmotic stress tolerance in *Arabidopsis* and soybean. *BMC Plant Biol.* 19 (1), 506. doi: 10.1186/s12870-019-2066-6

Zhuo, C. L., Wang, T., Lu, S. Y., Zhao, Y. Q., Li, X. G., and Guo, Z. F. (2013). A cold responsive galactinol synthase gene from *Medicago falcata (MfGolS1)* is induced by myo-inositol and confers multiple tolerances to abiotic stresses. *Physiol. Plantarum* 149, 1, 67–78. doi: 10.1111/ppl.12019

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 19 December 2022
DOI 10.3389/fpls.2022.1051165

OPEN ACCESS

EDITED BY
Wengang Xie,
Lanzhou University, China

REVIEWED BY
Lanting Zeng,
South China Botanical Garden (CAS), China
Gang Nie,
Sichuan Agricultural University, China

*CORRESPONDENCE
Bingzhe Fu
Fbzhe19@163.com
Shuxia Li
lishuxia620@163.com

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

RECEIVED 22 September 2022
ACCEPTED 02 December 2022
PUBLISHED 19 December 2022

CITATION
Wang J, Gao X, Wang X, Song W, Wang Q, Wang X, Li S and Fu B (2022) Exogenous melatonin ameliorates drought stress in *Agropyron mongolicum* by regulating flavonoid biosynthesis and carbohydrate metabolism. *Front. Plant Sci.* 13:1051165. doi: 10.3389/fpls.2022.1051165

Exogenous melatonin ameliorates drought stress in *Agropyron mongolicum* by regulating flavonoid biosynthesis and carbohydrate metabolism

Jing Wang[1], Xueqin Gao[1,2], Xing Wang[1], Wenxue Song[1], Qin Wang[1], Xucheng Wang[1], Shuxia Li[1,2]* and Bingzhe Fu[1,2,3]*

[1]School of Agriculture, Ningxia University, Yinchuan, Ningxia, China, [2]Ningxia Grassland and Animal Husbandry Engineering Technology Research Center, Yinchuan, Ningxia, China, [3]Key Laboratory for Model Innovation in Forage Production Efficiency, Ministry of Agriculture and Rural Affairs, Yinchuan, Ningxia, China

Drought is one of the most common abiotic stressors in plants. Melatonin (MT) is a high-efficiency and low-toxicity growth regulator that plays an important role in plant responses to drought stress. As a wild relative of wheat, *Agropyron mongolicum* has become an important species for the improvement of degraded grasslands and the replanting of sandy grasslands. However, the physiological and molecular mechanisms by which exogenous MT regulates drought stress in *A. mongolicum* remain unclear. To assess the effectiveness of MT intervention (100 mg·L^{-1}), polyethylene glycol 6000 was used to simulate drought stress, and its ameliorating effects on drought stress in *A. mongolicum* seedlings were investigated through physiology, transcriptomics, and metabolomics. Physiological analysis indicated that MT treatment increased the relative water content and chlorophyll content and decreased the relative conductivity of *A. mongolicum* seedlings. Additionally, MT decreased malondialdehyde (MDA) and reactive oxygen species (ROS) accumulation by enhancing antioxidant enzyme activities. The transcriptome and metabolite profiling analysis of *A. mongolicum* seedlings treated with and without MT under drought stress identified the presence of 13,466 differentially expressed genes (DEGs) and 271 differentially expressed metabolites (DEMs). The integrated analysis of transcriptomics and metabolomics showed that DEGs and DEMs participated in diverse biological processes, such as flavonoid biosynthesis and carbohydrate metabolism. Moreover, MT may be involved in regulating the correlation of DEGs and DEMs in flavonoid biosynthesis and carbohydrate metabolism during drought stress. In summary, this study revealed the physiological and molecular regulatory mechanisms of exogenous MT in alleviating drought stress in *A. mongolicum* seedlings, and it provides a reference for the development and utilization of MT and the genetic improvement of drought tolerance in plants from arid habitats.

KEYWORDS

***Agropyron mongolicum*, melatonin, drought stress, flavonoid biosynthesis, carbohydrate metabolism**

Introduction

Agropyron mongolicum is a diploid perennial herb of Poaceae. It has significant value for studying the origin, evolution, and genetic diversity of species, and it is an important grass in forage production and ecological construction (Che and Li, 2007). The ecological range of *A. mongolicum* is very broad, and it is mainly distributed in the desert steppe of central-western China. As a high-quality forage grass in grassland areas, the hay yield of *A. mongolicum* occupies an important position among the grasses of a typical steppe (Liu et al., 2013). In addition, as a wild relative species of wheat, *A. mongolicum* is not only an important resistance gene pool for wheat breeding but is also an important species for degraded grassland improvement and sandy grassland reseeding (Zhao et al., 2018).

With increasing global temperature, drought stress is becoming one of the major threats to the growth and development of plants (Rollins et al., 2013). Drought stress causes intricate physiological and biochemical reactions in plant cells by disturbing water balance and reducing plant water-use efficiency. Plants respond to drought stress mainly by increasing the content of antioxidants, low molecular weight osmolytes, and plant growth substances (Farooq et al., 2012). However, excessive drought in a short period or prolonged exposure to drought conditions can cause irreversible cell damage, eventually leading to cell death. Therefore, it is necessary to explore some alternative exogenous drought-resistant chemicals to enhance the intrinsic drought-resistant actions, thereby mitigating the adverse effects of drought on plant production.

Melatonin (MT), an indoleamine compound, was first discovered in animals (Lerner et al., 1958). In the 1990s, MT was detected in higher plants and then extensively used in the plant field (Hattori et al., 1995; Manchester et al., 2000; Ramakrishna et al., 2012). MT has multiple functions in the process of plant growth and development, including promoting photosynthesis (Zhao et al., 2015), participating in lateral root formation (Bajwa et al., 2014), delaying leaf senescence (Liang et al., 2015), regulating flower development and fruit ripening, and regulating the circadian rhythm (Sun et al., 2015), even if its content is extremely low. Additionally, MT is well-known for its unique antioxidant properties, and it acts as an important defender for plants against oxidative stress (Wang et al., 2018; Debnath et al., 2019). Many studies have demonstrated that exogenous MT can improve plant tolerance to abiotic stresses, such as ultraviolet radiation, low temperature, high temperature, drought, salinity, and alkali ions (Zhu et al., 2019; Cen et al., 2020; Guo et al., 2020; Yao et al., 2021). Recently, considerable research has been performed to decipher the functions of exogenous and endogenous melatonin in the drought stress mitigation of plants. These studies indicated that the pretreatment of melatonin through foliar spray, seedling treatment, and root dipping regulates physiological processes such as osmoregulation, germination, antioxidant mechanisms, photosynthesis, ion homeostasis, antisenescence primary and secondary metabolism, and hormonal cross-talks in plants under water deficit conditions (Cui et al., 2017; Li et al., 2018a). In addition, studies on the molecular mechanism of MT in alleviating drought stress in plants showed that MT enhances drought tolerance in plants mainly by regulating the MAPK pathway, photosynthetic system, flavonoid biosynthesis, ascorbate-glutathione cycle, hormone metabolism, and carbohydrate metabolism (Liang et al., 2018; Campos et al., 2019; Ibrahim et al., 2020; Moustafa et al., 2020; Hu et al., 2022). For example, exogenous MT application can alleviate the damage caused by drought stress on maize (Ren et al., 2021), cotton (Hu et al., 2022), loquat (Wang et al., 2021), and soybean (Cao et al., 2021) by regulating starch and sucrose metabolism. Furthermore, MT can enhance drought tolerance in citrus (Jafari and Shahsavar, 2021), grape (Deluc et al., 2009), and soybean (Cao et al., 2020) by promoting flavonoid biosynthesis. However, the regulatory mechanism of MT in *A. mongolicum* under drought stress has not yet been reported.

In this study, we analyzed the effects of 100 $mg{\cdot}L^{-1}$ MT on *A. mongolicum* under drought stress through physiology, transcriptomics, and metabolomics to provide a reference for the utilization of MT and the genetic improvement of drought tolerance in plants from arid habitats.

Materials and methods

Plant material and experimental design

Agropyron mongolicum 'Yanchi' is a new forage variety bred by Ningxia University after several years of selection and domestication. Seeds of *A. mongolicum* were disinfected with 10% sodium hypochlorite (*v/v*) for 10 min. Subsequently, the seeds were rinsed thoroughly with distilled water. The seeds were then placed in petri dishes (9 cm in diameter) with two layers of filter paper, and 50 seeds were placed in each petri dish after adding 5 mL of distilled water. Afterwards, the petri dishes were placed in an artificial climate incubator under the following conditions: temperature of 25°C/20°C (day/night), humidity of 55–65%, photoperiod of 14 h light/10 h dark, and light intensity of 1000 $\mu mol\ m^{-2}\ s^{-1}$. Water was replenished every day to maintain moisture, and the filter paper was changed every 3 days. After 12 days, the seedlings with the same morphology were transferred to Hoagland nutrient solution for hydroponics, and the nutrient solution was replaced every 5 days. When the seedlings were hydroponically cultivated for 28 days, the leaves were sprayed with distilled water or MT solution under pretreatment for 7 days; each pot was sprayed with 50 mL of

100 mg·L^{-1} MT or distilled water every day. After pre-spraying for 7 days, drought stress was simulated with 12% polyethylene glycol 6000 (PEG) solution. The concentrations of 100 mg·L^{-1} MT and 12% PEG were chosen based on a preliminary study. Thus, four treatments were included in this study: (i) control (CK), Hoagland nutrient solution plus foliar pre-spray for 7 days with distilled water; (ii) melatonin (M), Hoagland nutrient solution plus foliar pre-spray for 7 days with 100 mg·L^{-1} MT; (iii) drought stress (D), Hoagland nutrient solution containing 12% PEG plus foliar pre-spray for 7 days with distilled water; (iv) drought stress with melatonin (MD), Hoagland nutrient solution containing 12% PEG plus foliar pre-spray for 7 days with 100 mg·L^{-1} MT. Each treatment has three biological replicates for a total of 75 seedlings, which are uniformly distributed in the pots (20 × 14 × 13 cm). Water was replenished every 2 days, and the PEG concentration remained constant. *A. mongolicum* seedling leaves were collected on the 1st, 3rd, 5th, and 7th days after treatment for the physiological experiment. The leaves were sampled on the 7th day of treatment and immediately stored at −80°C for transcriptomic and metabolomic analysis.

Measurement of physiological indicators

Plant height was measured with a ruler. The number of wilted leaves in each treatment was counted to calculate the leaf wilt rate. The relative water content was estimated following Ahmad et al. (2019). The relative conductivity was estimated according to the methods of a previous study (Wu et al., 2017). Chlorophyll content was determined *via* spectrophotometry (Arnon, 1949). The malondialdehyde (MDA) content was determined using the thiobarbituric acid method (Puckette et al., 2007). Superoxide anion radical ($O_2^{\cdot-}$) content was determined by the p-aminobenzenesulfonic acid method, and hydrogen peroxide (H_2O_2) content was spectrophotometrically determined as described by Okuda et al. (1991). Superoxide dismutase (SOD) was determined using the nitroblue tetrazolium method (Niu et al., 2017), and peroxidase (POD) activity was measured using the methoxyphenol method (Wang et al., 2013). The proline (Pro) content was determined using the acid-ninhydrin method (Bates et al., 1973). Three biological replicates were performed for each of the above indicators.

RNA extraction and sequencing

The RNA samples were extracted from the leaves of *A. mongolicum* plants under different treatments. Each treatment was represented by three biological replicates of the leaf samples. Total RNA was extracted from leaf tissues using an RNAprep Pure Plant kit (Tiangen, Beijing, China). RNA quality was detected by a NanoPhotometer spectrophotometer (IMPLEN, CA, USA), Qubit 2.0 Fluorometer (Life Technologies, Carlsbad, CA, USA), and an Agilent Bioanalyzer 2100 system (Agilent Technologies, Santa Clara, CA, USA). The cDNA (~200 bp) was screened using AMPure XP beads. After amplification and purification, cDNA libraries were obtained and sequenced using the Illumina HiSeqTM 2000 system (Illumina, San Diego, CA, USA).

The raw data were filtered using fastp v 0.19.3, and all subsequent analyses were based on clean reads. Transcriptome assembly was performed using Trinity (v2.11.0). Gene expression levels were estimated using RSEM, and the fragments per kilobase of transcripts per million (FPKM) value of each gene was calculated based on the gene length. Differential expression analysis between the two groups was performed using DESeq2 v1.22.1. A $|\log_2\text{Fold Change}| \geq 1$ and FDR < 0.05 were used as the threshold for significant differential expression. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were performed based on a hypergeometric test.

Metabolites analysis by LC-MS/MS

The freeze-dried samples were ground into a powder and extracted with 70% aqueous methanol. After centrifugation at 10,000×g for 10 min, all supernatants were combined and filtered through a 0.22-mm pore size membrane and then processed for UPLC-ESI-MS/MS analysis (UPLC, SHIMADZU Nexera X2, https://www.shimadzu.com.cn/; MS, Applied Biosystems 4500 Q TRAP, https://www.thermofisher.cn/cn/zh/home/brands/applied-biosystems.html). Three biological replicates were set for each treatment in metabolites analysis. The effluent was alternatively connected to an ESI-triple quadrupole linear ion trap (Q TRAP)-MS. Linear ion trap (LIT) and triple quadrupole (QQQ) scans were acquired on a triple quadrupole-linear ion trap mass spectrometer equipped with an ESI Turbo Ion-Spray interface operating in positive ion mode and controlled by Analyst 1.6.3 software (AB Sciex). A scheduled multiple reaction monitoring method was used to quantify the metabolites.

Significantly regulated metabolites between groups were determined by VIP ≥ 1 and $\log_2|\text{FoldChange}| > 1$. The identified metabolites were annotated using the KEGG Compound database (http://www.kegg.jp/kegg/compound/). Annotated metabolites were then mapped to the KEGG Pathway database (http://www.kegg.jp/kegg/pathway.html). Pathways with significantly regulated metabolites were then input into metabolite set enrichment analysis (MSEA); their significance was determined by p-values from a hypergeometric test. Principal component analysis (PCA; base package), correlation analysis using R (base package; Hmisc), and plot heatmaps (ComplexHeatmap) were performed using R software.

Quantitative real-time RCR analyses

For qRT-PCR, RNA was extracted using a Trizol Total RNA Extraction Kit (Sangon Biotech, Shanghai, China) and reverse transcribed using an Evo M-MLV RT Mix Kit with gDNA Clean (Accurate Biotechnology, Hunan, China). The primers are listed in Supplementary Table 1. qRT-PCR was conducted using a BioEasy Master Mix (SYBR Green) Kit (Bioer, Hangzhou, China) and a C1000 TouchChihermal Cycler system (Bio-Rad). All tested transcripts were normalized to the reference gene *TLF*, and relative transcript levels were calculated according to the $2^{-\Delta\Delta Ct}$ method (Tian et al., 2016). Three biological and technical replications were performed.

Statistical analysis

The data were analyzed with Excel 2019 (Microsoft Inc., Redmond, USA) and SPSS 22 (SPSS Inc., Chicago, USA) using an analysis of variance (ANOVA), followed by Duncan's significant difference test at $P < 0.01$ or $P < 0.05$. The graphs were constructed using Origin 2020 (Electronic Arts Inc., San Francisco, CA, USA).

Results

Physiological analysis of MT-treated *A. mongolicum* seedlings under drought stress

To determine whether MT improved the drought tolerance of *A. mongolicum*, PEG and MT treatments were performed on the seedlings, and phenotypic and physiological traits were analyzed. The preliminary experiments for PEG concentration screening (0, 3, 6, 9, 12, and 15% PEG) showed that under 12% PEG, the changes of each index were larger, and the leaf wilt rate was close to 50%, indicating that the use of 12% PEG solution to simulate drought stress had a significant but not serious inhibitory effect on *A. mongolicum* seedlings. Therefore, this concentration was selected as the concentration of drought treatment in subsequent experiments (Supplementary Figures 1 and 2A–I). The preliminary experiments for MT concentration screening (0, 1, 10, 50, 100, 150, and 200 mg·L^{-1} MT) showed that 100 mg· L^{-1} MT treated *A. mongolicum* had a significantly higher plant height, dry aboveground weight, relative water content, and chlorophyll content compared with plants under drought stress; thus, this concentration was selected for the MT treatment in subsequent experiments (Supplementary Figures 3 and 4A–I).

Based on the optimal PEG and MT concentrations screened, we analyzed the effects of exogenous MT addition on the physiology of *A. mongolicum* seedlings under different drought stress times. MT treatment significantly alleviated the withering of seedling leaves under drought stress. After 7 days of drought stress, although the leaves of seedlings wilted in both the MD and D treatments, the degree of wilting in the MD treatment was lower than that in the D treatment. After 9 days of drought stress, the leaves under D treatment withered and died (Figure 1). On the 1st day of drought stress, all physiological indexes showed no significant difference between treatments. Compared with the control, drought stress significantly decreased the relative water content and chlorophyll content of leaves of *A. mongolicum* seedlings on the 7th day of drought stress. However, MT application dramatically mitigated this effect (Figures 2A, C). MT treatment significantly decreased the relative conductivity and MDA content under drought stress but had no obvious effect on the control plants (Figures 2B, D). In addition, drought treatment resulted in a significant increase in H_2O_2 and $O_2^{\cdot-}$ contents, while MT addition significantly decreased their content (Figures 2E, F). Drought treatment enhanced SOD and CAT activities, and there was no significant difference in them after MT addition (Figures 2G, H). The results showed that the *A. mongolicum* seedlings at 7 days directly reflected a turning point in the degree of drought stress. Therefore, we selected the *A. mongolicum* leaves of CK, M, D and MD treatments at 7 days (serious drought) to perform transcriptomic and metabolomic analysis.

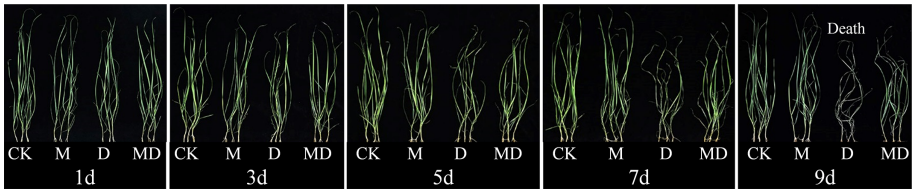


FIGURE 1
Effects of 100 mg·L^{-1} MT on the phenotype of *A. mongolicum* seedlings under different durations of drought stress.

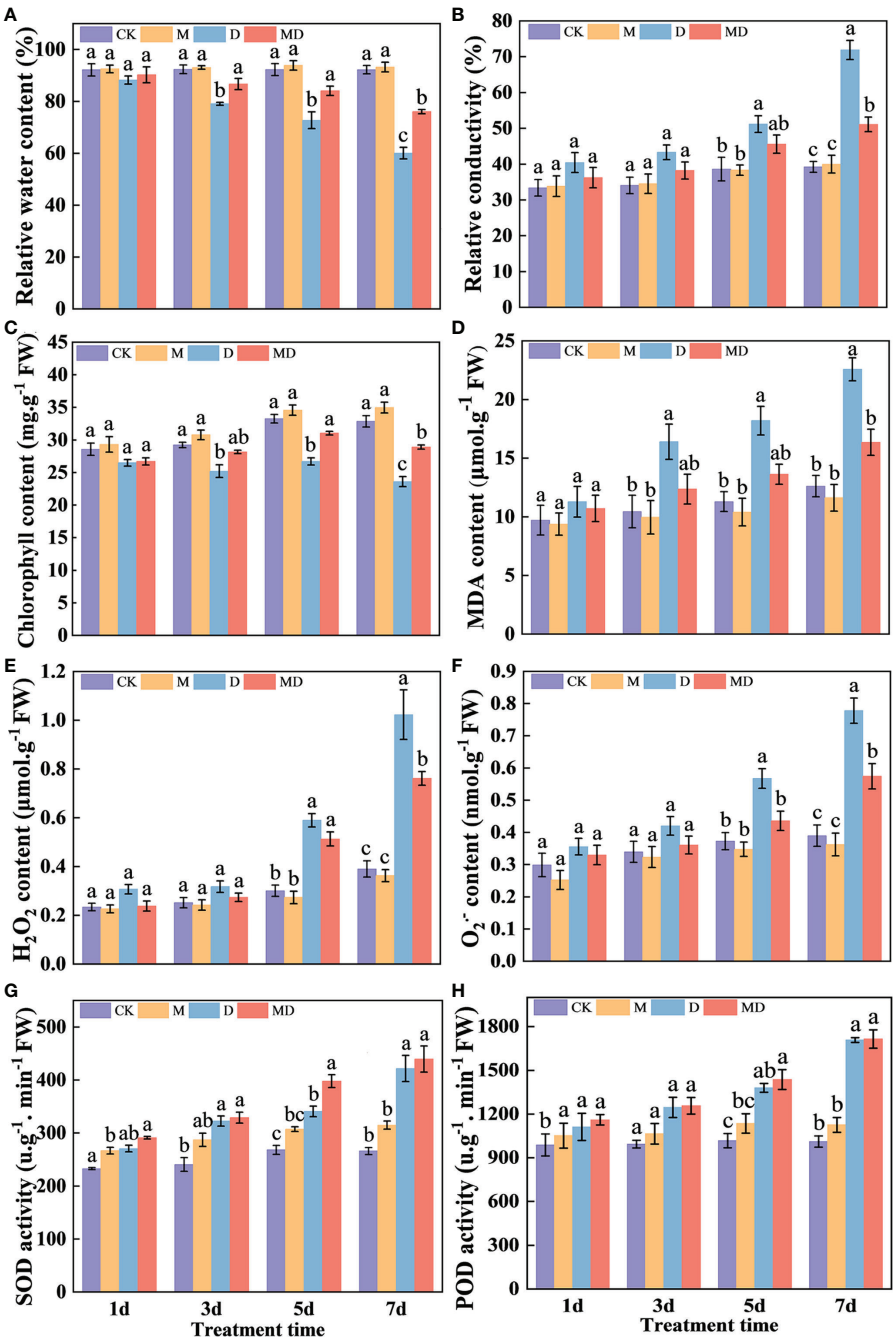


FIGURE 2
Effects of 100 mg·L^{-1} MT on the physiology of *A. mongolicum* seedlings under different durations of drought stress. **(A)** Relative water content, **(B)** relative conductivity, **(C)** chlorophyll content, **(D)** MDA content, **(E)** H_2O_2 content, **(F)** O_2^- content, **(G)** SOD activity, **(H)** POD activity. Mean values followed by different letters were significantly different by Duncan's test ($P < 0.01$).

Transcriptome analysis of MT-treated *A. mongolicum* seedlings under drought stress

To explore the molecular mechanism of MT regulating the drought resistance of *A. mongolicum* seedlings, transcriptome analysis was performed on *A. mongolicum* seedlings treated with and without MT under drought stress. After removing the low-quality reads, a total of 86.19 Gb clean data were obtained. The percentage of GC was 49.01–57.44%, and the Q30 bases of each sample were above 92.63%, indicating that the data quality and purity of transcriptome sequencing were high. A total of 149,051 genes were functionally annotated in at least one database (Supplementary Table 2). Moreover, 10,293 (5140 upregulated and 5153 downregulated) and 13,466 (6670 upregulated and 6796 downregulated) differentially expressed genes (DEGs) were identified in the comparison groups of "CK vs. D" and "D vs. MD," respectively (Supplementary Table 3). These results indicate that MT induced changes in the transcripts of *A. mongolicum* seedlings under drought stress.

DEG analysis of MT-treated *A. mongolicum* seedlings under drought stress

To further clarify the response to MT, the biological functions, GO and KEGG enrichment analyses of DEGs were performed in MT-treated *A. mongolicum* seedlings under drought stress (Supplementary Table 4). The GO terms of DEGs in "CK vs. D" and "D vs. MD" were significantly enriched in "fructosyltransferase activity," "sucrose 1F-fructosyltransferase activity," "beta-fructofuranosidase activity," "sucrose alpha-glucosidase activity," and "alpha-glucosidase activity" in molecular function classes (Figure 3A). In addition, the GO terms of DEGs in "CK vs. D" were mainly enriched in "cell killing," "killing of cells of other organism," and "disruption of cells of other organism" in biological process classes and were enriched in "photosystem I" and "senescence-associated vacuole" in cellular composition classes. The GO terms of DEGs in "D vs. MD" were mainly enriched in "negative regulation of translation" and "negative regulation of cellular amide metabolic process" in biological process classes and were enriched in "cortical endoplasmic reticulum" and "endoplasmic reticulum tubular network" in cellular composition classes (Figures 3A, B).

KEGG enrichment analysis showed that several pathways related to plant stress resistance or metabolism were enriched in "CK vs. D" and "D vs. MD," including galactose metabolism, starch and sucrose metabolism, phenylpropanoid biosynthesis, flavonoid biosynthesis, protein processing in endoplasmic reticulum, and alanine, aspartate, and glutamate metabolism (Figures 3C, D).

Metabolome analysis of MT-treated *A. mongolicum* seedlings under drought stress

To elucidate the specific effects on metabolites, we carried out broadly targeted metabolome assays. PCA showed that the three biological replicates for each treatment were clustered together and clearly divided into four groups, indicating significant metabolic differences between treatments. In addition, the metabolism of these four different treatments was significantly separated in the first component (PC1), except for CK and M treatments, indicating that the effect of MT on the metabolism of *A. mongolicum* under drought stress was obvious (Figure 4A). All identified metabolites are presented in Supplementary Table 5. The number of metabolites in "CK vs. M," "CK vs. D," "CK vs. MD," and "D vs. MD" was 162, 391, 314, and 271, respectively (Figure 4B and Supplementary Table 6). These differentially expressed metabolites (DEMs) were mainly classified into flavonoids (261), phenolic acids (160), lipids (142), alkaloids (110), amino acids and derivatives (88), organic acids (67), nucleotides and derivatives (56), lignans and coumarins (40), terpenoids (21), and others (103) (Figure 4C). Interestingly, among the 19 DEMs involved in "CK vs. M," "CK vs. D," "CK vs. MD," and "D vs. MD," flavones and flavonols were highly accumulated in "D vs. MD" but not in "CK vs. MD" and "CK vs. D." However, the DEMs of flavanones, flavonoid carbonoside, free fatty acids, gycerol ester, and phenolic acids were all significantly upregulated in "CK vs. D," but downregulated in "D vs. MD" (Figure 4D and Supplementary Table 7). MT induced changes in antioxidants in *A. mongolicum* seedlings under drought stress.

Integrated analysis of genes and metabolites related to flavonoid biosynthesis in MT-treated *A. mongolicum* under drought stress

Based on the integrated analysis of the transcriptome and metabolome, we focused on flavonoid biosynthesis, galactose metabolism, and starch and sucrose metabolism pathways (Supplementary Table 8).

To explore the impacts of MT on secondary metabolism in *A. mongolicum* under drought stress, the interaction of DEGs and DEMs related to flavonoid biosynthesis was analyzed (Figure 5 and Supplementary Table 9). We found that three DEGs encoding chalcone synthase (CHS), three DEGs encoding phlorizin synthase (PHS), and nine DEGs encoding shikimate

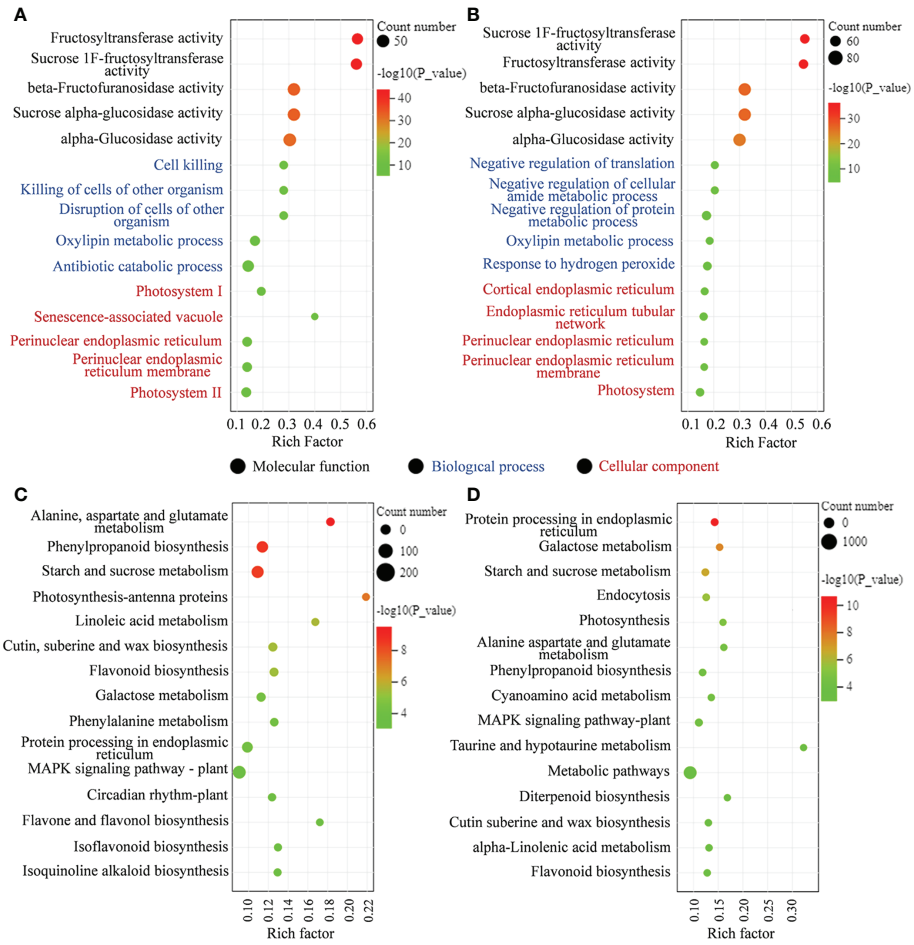


FIGURE 3
DEGs enriched on different GO terms and KEGG pathways. **(A)** GO terms of DEGs in "CK vs. D." **(B)** GO terms of DEGs in "D vs. MD." **(C)** KEGG pathway analysis of DEGs in "CK vs. D." **(D)** KEGG pathway analysis of DEGs in "D vs. MD.".

O-hydroxycinnamoyl transferase (HCT) were differentially expressed in "CK vs. D," and the expression patterns of these DEGs in "D vs. MD" were opposite to that in "CK vs. D." One gene encoding cinnamate-4-hydroxylase (C4H) was significantly upregulated in both "CK vs. D" and "D vs. MD," and the DEG encoding chalcone isomerase (CHI) was significantly downregulated. Furthermore, two DEGs encoding naringenin 7-O-methyltransferase (NOMT) were significantly upregulated in "D vs. MD," but they were not detected in "CK vs. D." Moreover, in the flavonoid biosynthesis pathway, sakuranetin and coumaroyl quinic acid accumulated in both "CK vs. D" and "D vs. MD," but phlorizin accumulated only in "D vs. MD."

As shown in Figures 6A–H, three metabolites, namely coumaroy shikimic acid, caffeoyl shikimic acid, and vitexin, in flavonoid biosynthesis were more abundant under D treatment compared to CK, and they were less abundant in MD treatment than in D treatment. The content of hesperetin-7-O-glucoside and luteolin in CK was significantly higher than those in M, D, and MD treatments. The content of coumaroyl quinic acid and sakuranetin was significantly higher in the D and MD treatments than in CK. Moreover, the phlorizin chalcone content in the CK and D treatments was significantly higher than that in the D and MD treatments. The DEGs and DEMs related to flavonoid biosynthesis in MT-treated *A. mongolicum* were considered to conjointly respond to drought stress.

Integrated analysis of genes and metabolites related to carbohydrate metabolism in MT-treated *A. mongolicum* under drought stress

Analyzing the interaction of DEGs and DEMs related to galactose metabolism enables us to better understand the complex mechanism of MT in the galactose metabolism of *A. mongolicum* under drought stress (Figure 7A and

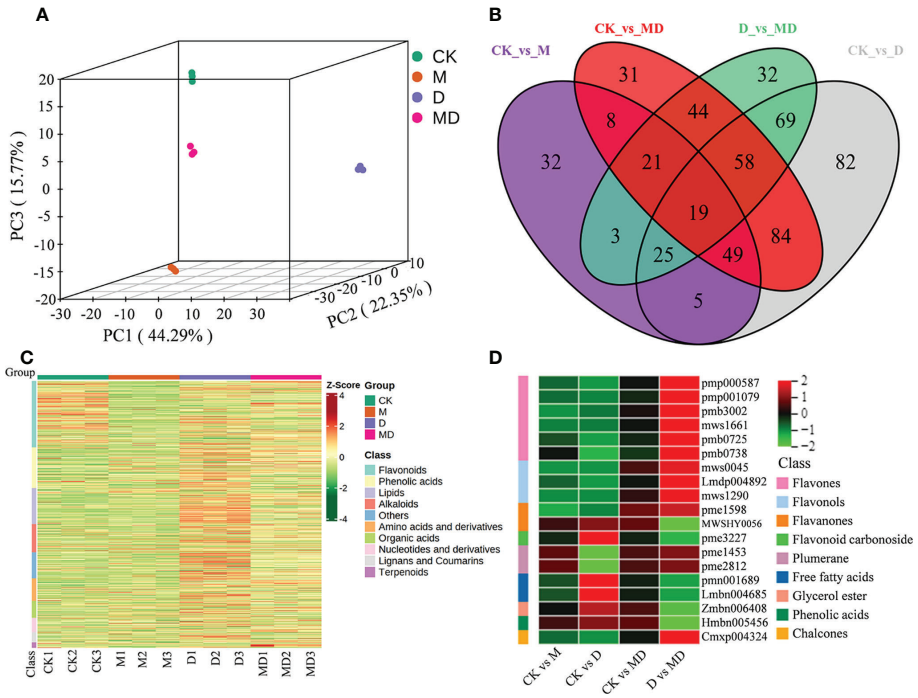


FIGURE 4
Metabolites analysis of *A. mongolicum* seedlings under drought stress. **(A)** PCA analysis of metabolites. **(B)** Venn diagram of DEMs. **(C)** Heat map visualization of metabolites. Red indicates high abundance, whereas relatively low-abundance metabolites are shown in green (color key scale at the right of the heat map). **(D)** Heat map analysis of common DEMs in "CK vs. MD," "CK vs. D," "CK vs. MD," and "D vs. MD.".

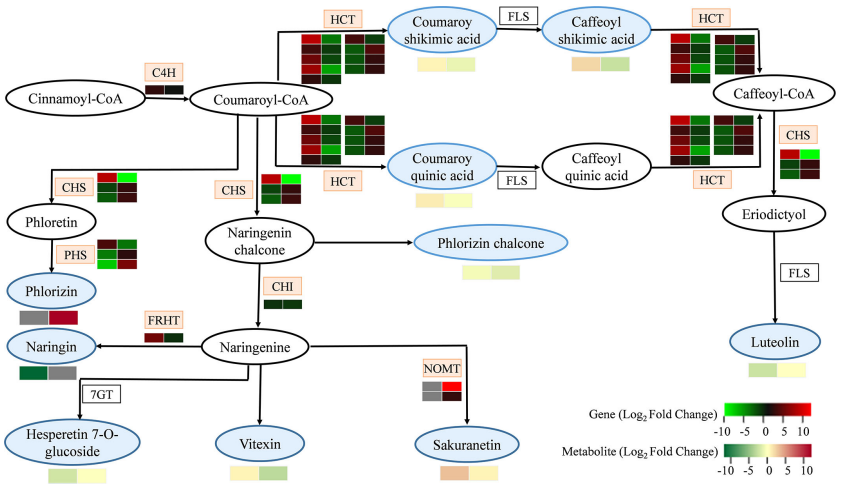


FIGURE 5
DEGs and DEMs involved in flavonoid biosynthesis in response to exogenous addition of MT under drought stress. The blue and pink patterns represent the metabolites and genes that changed after exogenous addition of MT under drought stress, respectively. The rectangle is divided into two equal parts: the left of the rectangle represents DEGs or DEMs in "CK vs. D"; the right of the rectangle represents DEGs or DEMs in "D vs. MD." The colors in the rectangle represent the genes or metabolites regulated after exogenous addition of MT under drought stress (red indicates upregulation; green indicates downregulation; gray indicates undetected).

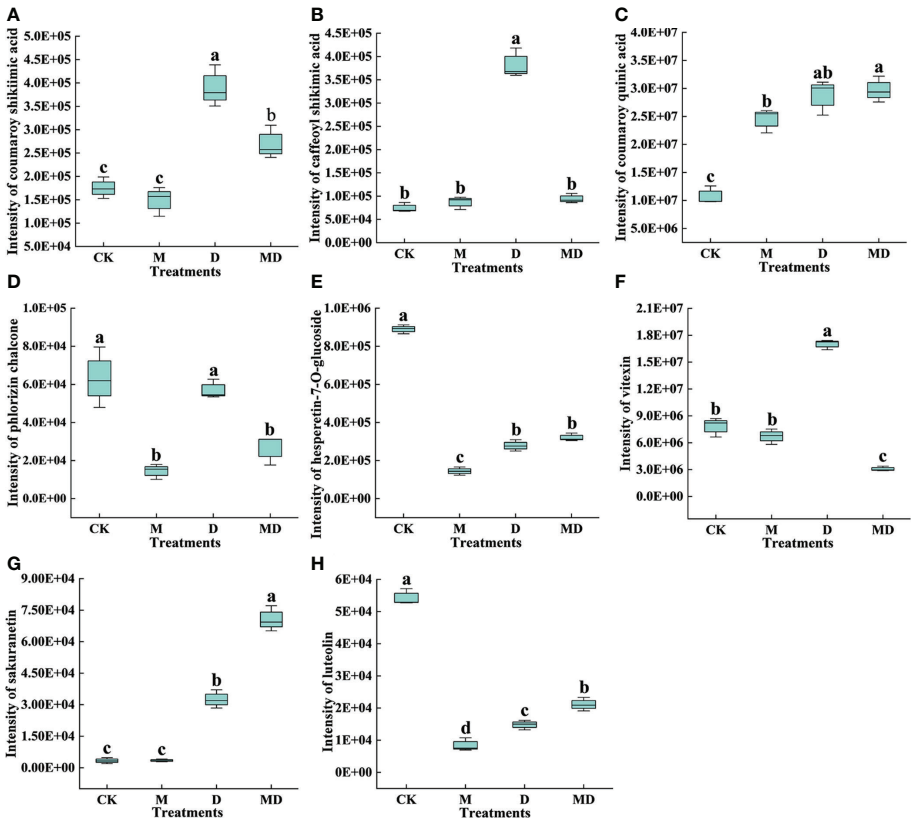


FIGURE 6
Effects of exogenous MT on DEMs in the flavonoid biosynthetic pathway in *A. mongolicum* leaves under drought stress. **(A)** Coumaroyl shikimate acid, **(B)** caffeoyl shikimic acid, **(C)** coumaroyl quinic acid, **(D)** phlorizin chalcone, **(E)** hesperetin-7-O-glucoside, **(F)** vitexin, **(G)** sakuranetin, **(H)** luteolin. Error bars represent standard errors of triplicate experiments. Mean values with different letters are significantly different at $P < 0.05$.

Supplementary Table 10). A DEG encoding raffinose synthase (RAFS) was upregulated in "CK vs. D," while three DEGs encoding RAFS were downregulated. Four DEGs encoding beta-fructofuranosidase (FRS) were upregulated in "CK vs. D," while two DEGs encoding FRS were downregulated. The expression patterns of 10 DEGs encoding RAFS and FRS were opposite in "CK vs. D" and "D vs. MD." The DEGs encoding stachyose synthetase (STS) were significantly upregulated in "CK vs. D" but were downregulated in "D vs. MD." A gene encoding alpha-galactosidase (GAS) was significantly downregulated in "CK vs. D" but was significantly upregulated in "D vs. MD." In addition, in the galactose metabolism pathway, the DEMs of galactinol, raffinose, stachyose, D-galactose, melibiose, D-glucose, and sucrose were significantly upregulated in "CK vs. D" but were downregulated in "D vs. MD." The DEGs and DEMs related to galactose metabolism in *A. mongolicum* conjointly responded to drought stress after exogenous MT addition.

Starch and sucrose metabolism plays an important role in plant growth and development. Analysis of the interaction of DEGs and DEMs related to this pathway can further clarify the regulatory mechanism of MT to alleviate drought stress in *A. mongolicum* (Figure 7B and Supplementary Table 11). The three DEGs encoding sucrose synthase (SUS) were significantly upregulated, and one DEG was significantly downregulated in "CK vs. D." The expression pattern of these DEGs in "D vs. MD" was opposite that in "CK vs. D." The four DEGs encoding trehalose 6-phosphate synthase/phosphatase (TPS/P) and one DEG encoding trehalose 6-phosphate phosphatase (TPP) were significantly downregulated in "CK vs. D" but were upregulated in "D vs. MD." One DEG encoding alpha, alpha-trehalase (TRE) was significantly upregulated in "CK vs. D" but was downregulated in "D vs. MD." Among starch and sucrose metabolic pathways, the DEMs of sucrose, trehalose 6-phosphate, trehalose, and D-glucose were significantly upregulated in

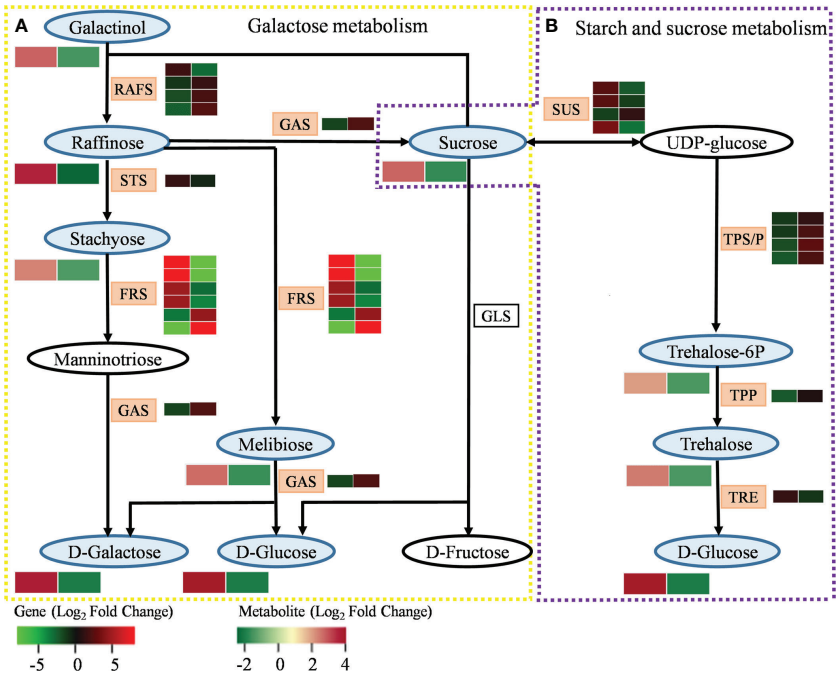


FIGURE 7
DEGs and DEMs involved in carbohydrate metabolism in response to exogenous addition of MT under drought stress. **(A)** Galactose metabolism. **(B)** Starch and sucrose metabolism. The blue and pink patterns represent the metabolites and genes that changed after exogenous addition of MT under drought stress, respectively. The rectangle is divided into two equal parts: the left of the rectangle represents DEGs or DEMs in "CK vs. D"; the right of the rectangle represents DEGs or DEMs in "D vs. MD." The colors in the rectangle represent the genes or metabolites regulated after exogenous addition of MT under drought stress (red indicates upregulation; green indicates downregulation; gray indicates undetected).

"CK vs. D" but were significantly downregulated in "D vs. MD." The DEGs and DEMs related to the starch and sucrose metabolism of *A. mongolicum* conjointly responded to drought stress after exogenous MT addition.

As shown in Figures 8A–I, based on the content of DEMs related to carbohydrate metabolism, nine metabolites (galactinol, raffinose, stachyose, D-galactose, melibiose, D-glucose, sucrose, trehalose 6-phosphate, and D-trehalose) were more abundant in the D treatment compared to the CK treatment, and they were less abundant in the MD treatment when compared to the D treatment.

Correlation analysis of DEGs and DEMs in flavonoid metabolism and carbohydrate metabolism in MT-treated *A. mongolicum* seedlings under drought stress

To examine the relationship between DEGs and DEMs in MT-treated *A. mongolicum* seedlings under drought stress, a correlation analysis of DEGs and DEMs involved in flavonoid and carbohydrate metabolism was conducted (|Pearson correlation coefficient| > 0.8, P value < 0.05) (Figure 9 and Supplementary Tables 12, 13). In flavonoid metabolism, caffeoyl shikimic acid and coumaroyl shikimate were strongly positively correlated with genes encoding flavanone 7-O-glucoside 2"-O-beta-L-rhamnosyltransferase (FRHT), HCT, PHS, and CHS. Meanwhile, sakuranetin and phlorizin were strongly positively correlated with the *C4H* and *NOMT* genes. In carbohydrate metabolism, lactobiose, stachyose, galactinol, D-trehalose, dulcitol, raffinose, D-galactose, inositol, D-glucose, trehalose 6-phosphate, D-sucrose, D-mannose, and melibiose were strongly positively correlated with genes encoding FRS, SUS, and TRE.

qRT-PCR validation

To confirm the accuracy and reproducibility of the RNA-seq data, nine DEGs were selected from flavonoid biosynthesis, galactose metabolism, and starch and sucrose metabolism for qRT-PCR. RNA-seq (FPKM) and qRT-PCR analyses were in concordance (Supplementary Figure 5), showing the same

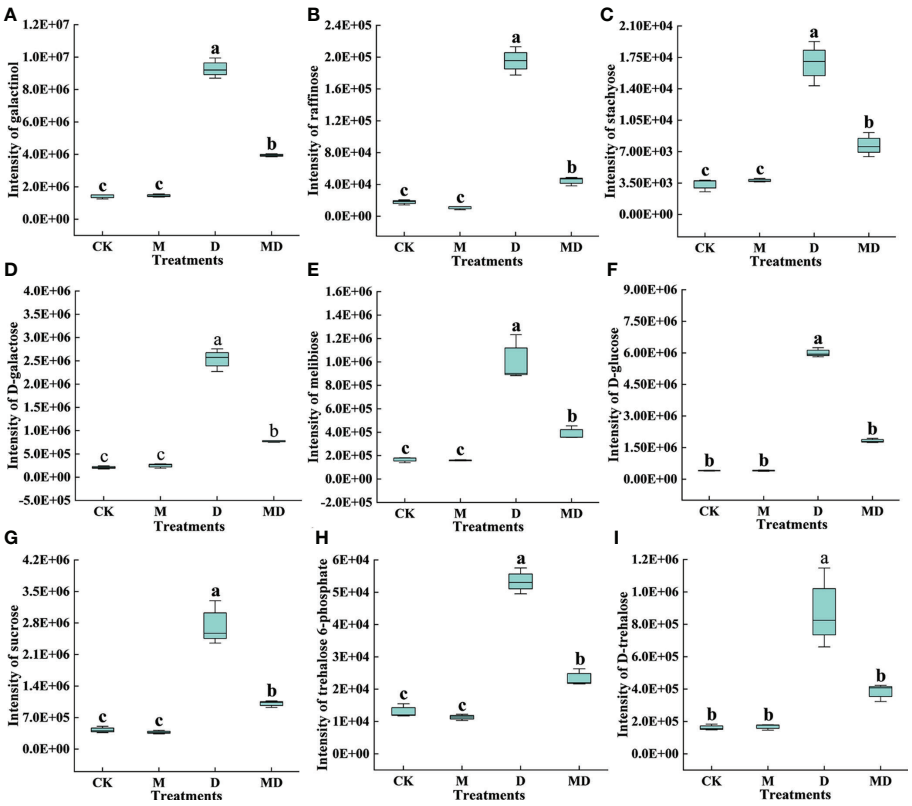


FIGURE 8
Effects of exogenous MT on DEMs in the carbohydrate metabolism pathway in *A. mongolicum* leaves under drought stress. **(A)** Galactinol, **(B)** raffinose, **(C)** stachyose, **(D)** D-galactose, **(E)** melibiose, **(F)** D-glucose, **(G)** sucrose, **(H)** trehalose 6-phosphate, **(I)** D-trehalose. Error bars represent standard errors of triplicate experiments. Mean values with different letters are significantly different at $P < 0.05$.

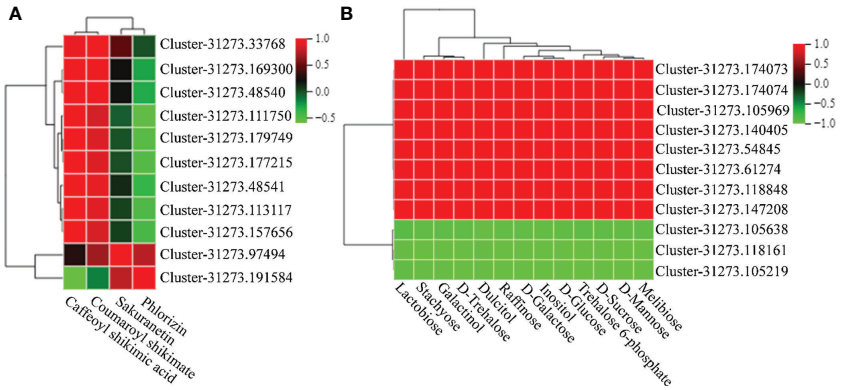


FIGURE 9
Correlation cluster heatmap of genes and metabolites. **(A)** Correlation cluster heatmap of DEGs or DEMs in flavonoid biosynthetic pathways. **(B)** Correlation cluster heatmap of DEGs or DEMs in carbohydrate metabolism pathways. The color in the box represents the correlation of genes and metabolites (red indicates positive correlation; green indicates negative correlation).

general expression trends. These results verified the reproducibility and reliability of the RNA-seq data.

Discussion

Physiological mechanism by which MT alleviates drought stress in *A. mongolicum* seedlings

As one of the leading abiotic stressors, drought detrimentally affects plant morphology and physiology, thereby inhibiting the growth and development of plants (Liu et al., 2015). Our results show that drought stress delayed the growth and development of *A. mongolicum* seedlings by reducing the relative water content and chlorophyll content, accumulating ROS and MDA, which causes membrane lipid peroxidation, and ultimately disrupting the structure and function of the cell membrane. MT is known as a broad-spectrum antioxidant that scavenges ROS and increases antioxidant enzyme activity under drought stress (Reiter et al., 2016). Herein, we found that MT application increased the relative water content and chlorophyll content, decreased the accumulation of ROS and MDA, and further increased the activities of antioxidant enzymes in *A. mongolicum* seedlings exposed to drought stress (Figures 2A–F). These results suggest that drought stress is alleviated by MT and those of its metabolites that are free endogenous radical scavengers. Similarly, Antoniou et al. (2017) and Cui et al. (2017) reported that exogenous MT addition significantly facilitated the growth and development of alfalfa and wheat and alleviated the accumulation of intracellular ROS and MDA induced by drought stress by enhancing its ability to scavenge free radicals. Therefore, the exogenous addition of an appropriate concentration of MT can alleviate the growth inhibition of *A. mongolicum* seedlings under drought stress by affecting physiological effects.

MT improved flavonoid biosynthesis in *A. mongolicum* seedlings under drought stress

Flavonoid biosynthesis is one of the most widely studied metabolic pathways (Mierziak et al., 2014). Many studies have shown that flavonoid metabolism is involved in plant response to drought stress (Yuan et al., 2012). Flavonoids are important secondary metabolites that can eliminate ROS accumulation through synergistic action with other stress response factors, such as MT and ABA, and ultimately improve the adaptability of plants to arid environments (Naghizadeh et al., 2019; Hossain et al., 2020; Yang et al., 2021b; Zou et al., 2021).

In the flavonoid biosynthesis pathway, the *C4H* gene controls the synthesis of p-coumaric acid from trans-cinnamic acid, which ultimately produces pigments (flavonoids) and plant defense compounds (furanocoumarins, isoflavonoids, and norlignans) (Suzuki et al., 2002; Suzuki et al., 2004). Drought stress caused an increase in the expression of the *C4H* gene in the leaves of *A. mongolicum* seedlings, and exogenous MT addition further enhanced the expression of the *C4H* gene under drought stress (Figure 5). Previous studies have also shown that C4Hs can enhance plant resistance to drought stress by increasing phenolic synthesis (Wang et al., 2017). Together, CHS and CHI constitute rate-limiting enzymes for flavonoid biosynthesis (Braune et al., 2016). Coumaroyl-CoA is catalyzed by CHS to form chalcone (Tan et al., 2015). Naringenin chalcone is an intermediate in flavonol biosynthesis and is converted to naringenin by CHI (Muir et al., 2001). In many plants, some external stimuli, such as adversity stress, induce the rapid response and expression of *CHS* and *CHI* (Muir et al., 2001). In this study, one *CHI* was significantly downregulated in both "CK vs. D" and "D vs. MD." Two *CHS* genes were significantly downregulated in "CK vs. D" but were upregulated in "D vs. MD," which indicated that MT induced the upregulation of two *CHS* genes and indirectly increased the flavonoid content, such as phloridzin and luteolin, and improved the antioxidant capacity of *A. mongolicum* (Figure 5). Previous studies have shown the complementary effects of phloridzin in dihydrochalcone compounds, which act as functional antioxidants to reduce oxidative stress in plants (Venisse et al., 2001). Natural flavonoids, such as luteolin, have the effect of scavenging free radicals (Wang et al., 2013) and can enhance the enzymatic activity of SOD, CAT, and other antioxidant enzymes (Lim et al., 2007; Choi et al., 2008). Song et al. (2022) showed that MT treatment greatly enhanced the stress tolerance of pigeon pea by promoting luteolin biosynthesis.

NOMT is a key enzyme in phytoalexin biosynthesis that catalyzes the methylation of naringenin to sakuratin (Tamogami et al., 1997). Rakwal et al. (2000) showed that NOMT was not found in the healthy tissues of rice or in the control treatment in our study. However, MT treatment significantly upregulated the expression of two *NOMT* genes under drought stress, thus enhancing the drought resistance of *A. mongolicum* seedlings by regulating the downstream metabolite sakuratin. In addition, HCT, a key enzyme in flavonoid biosynthesis, is required for coumaroyl-CoA to generate caffeoyl-CoA (Sun et al., 2018). *HCT* overexpression in pears leads to flavonoid accumulation that enhances their tolerance to drought stress (Yang et al., 2021a). In our study, six genes regulating HCT were significantly upregulated in "CK vs. D" but were significantly downregulated in "D vs. MD." This was similar to the content of two intermediate metabolites of coumaroyl shikimate acid and caffeoyl shikimic acid in the process of p-coumaroyl-CoA producing caffeoyl-CoA.

However, the content of some compounds, such as coumaroyl shikimate acid and caffeoyl shikimic acid, decreased after exogenous MT addition under drought stress, which may be due to the reduced accumulation of some flavonoids induced by MT, resulting in more energy conversion to cope with drought stress (Zou et al., 2021).

MT enhanced carbohydrate metabolism in *A. mongolicum* seedlings under drought stress

Carbohydrate metabolism is one of the most important metabolic processes in plants. It plays an important role in the complex regulatory networks of hormones, reactive oxygen species (ROS) production and scavenging, energy storage, and signal transduction throughout the plant life cycle (Bolouri et al., 2010). As a plant bioregulator, exogenous MT can change the content of metabolites in the process of carbohydrate metabolism by regulating related genes and activating the activities of related enzymes, thereby enhancing the tolerance of plants to drought stress (Arnao et al., 2021).

In carbohydrate metabolism, RAFS is a key enzyme that catalyzes the galactinol production of raffinose. Studies have shown that raffinose content in plants increases significantly when subjected to adversity stress (Egert et al., 2013). In this study, a gene encoding RAFS was upregulated under drought stress, increasing the raffinose content (Figure 7). Stachyose is mainly used as an antioxidant to scavenge ROS accumulated under stress conditions. In the present study, a gene encoding STS was significantly upregulated under drought stress, resulting in a massive accumulation of stachyose in the leaves of *A. mongolicum*. In addition, four genes encoding FRS were significantly upregulated under drought stress, causing an increase in melibiose content. Khan et al. (2021) showed that the expression level of FRS increased after drought stress in citrus trees. However, after exogenous MT addition under drought stress, *RAFS STS* and four *FRSs* were downregulated, resulting in a decrease in the content of downstream metabolites raffinose, stachyose, and melibiose. In addition, GAS, as an exoglycosidase that catalyzes the hydrolysis of α-galactosidic bonds, is mainly involved in the hydrolysis of melibiose, raffinose, and stachyose. Enhanced GAS activity reduces melibiose, raffinose, and stachyose (Hu et al., 2016). The results of this study showed that after exogenous MT addition under drought stress, a gene encoding GAS was significantly upregulated, decreasing the content of upstream metabolites, such as melibiose, raffinose, and stachyose. Trehalose, as a non-reducing disaccharide, is synthesized by TPS and TPP (Holmström et al., 1996). The accumulation of trehalose in plant cells can improve drought tolerance (Romero et al., 1997). In this study, trehalose was significantly accumulated under drought stress, and then, it was decomposed into glucose under the action of TRE. However, under drought stress, the trehalose content decreased after exogenous MT addition, and the downregulated expression of a gene encoding TRE resulted in a decrease in the content of its downstream metabolite glucose, which may indicate that MT addition alleviated drought stress by reducing trehalose degradation.

SUS is a key enzyme that regulates sucrose synthesis (Liu et al., 2016). Drought stress typically induces increases in plant SUS activity, which in turn increases sucrose accumulation and promotes the activity of pathways related to carbohydrate metabolism, ultimately alleviating the damage caused by drought (Yang et al., 2001). This study showed that the expression levels of the three *SUS* genes and sucrose content increased under drought stress. However, after exogenous MT addition, the expression levels of the three *SUS* genes decreased, causing a decrease in sucrose content, which indicates that MT application might accelerate the decomposition of sucrose, resulting in a reduction in sucrose content (Li et al., 2018b). Wang et al. (2021) also showed that two *SUS* genes were downregulated after exogenous MT addition under drought stress. Therefore, we speculate that the decrease in carbohydrate content after exogenous MT addition under drought stress may be due to MT enhancing carbohydrate conversion and transport capacity under drought stress, and exogenous MT addition leads to a decrease in energy production by reducing carbohydrate metabolism, thus slowing plant growth and development to adapt to the stimulation of an arid environment.

Conclusion

This study provides new insights into the mitigating effects of MT against drought stress in *A. mongolicum*. Based on the analysis of the physiology, transcriptome, and metabolome of *A. mongolicum* seedlings under drought stress by exogenous MT addition, the regulatory mechanism of MT in *A. mongolicum* seedlings under drought stress was revealed. Physiological analysis showed that MT alleviated drought stress by increasing chlorophyll content and enhancing antioxidant defense activity to reduce cell membrane damage and ROS accumulation. The integrated transcriptomic and metabolomic analysis showed that MT responds to drought stress mainly by regulating DEGs and DEMs in flavonoid biosynthesis and carbohydrate metabolism pathways. Our findings provide a basis for a further comprehensive and systematic analysis of the molecular mechanism by which MT alleviates drought stress in *A. mongolicum*.

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

JW, BF, and SL designed the experiments and wrote the manuscript. All authors performed the experiment. JW, XG, and XW analyzed the data and prepared the figures. BF and SL provided ideas and revised the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Natural Science Foundation of China (32101426), Ningxia Natural Science Foundation (2022AAC03058), and Ningxia Hui Autonomous Region Agricultural Breeding Special Project (2019NYYZ0403).

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.2022.1051165/full#supplementary-material

References

Ahmad, S., Kamran, M., Ding, R., Meng, X., Wang, H., Ahmad, I., et al. (2019). Exogenous melatonin confers drought stress by promoting plant growth, photosynthetic capacity and antioxidant defense system of maize seedlings. *PeerJ* 7, e7793. doi: 10.7717/peerj.7793

Antoniou, C., Chatzimichail, G., Xenofontos, R., Pavlou, J. J., Panagiotou, E., Christou, A., et al. (2017). Melatonin systemically ameliorates drought stress-induced damage in *Medicago sativa* plants by modulating nitro-oxidative homeostasis and proline metabolism. *J. Pineal Res.* 62, np–n/a. doi: 10.1111/jpi.12401

Arnao, M. B., Hernández-Ruiz, J., Cano, A., and Reiter, R. J. (2021). Melatonin and carbohydrate metabolism in plant cells. *Plants* 10, 1917. doi: 10.3390/plants10091917

Arnon, D. I. (1949). Copper enzymes in isolated chloroplasts. polyphenoloxidase in *Beta vulgaris*. *Plant Physiol.* 24, 1–15. doi: 10.1104/pp.24.1.1

Bajwa, V. S., Shukla, M. R., Sherif, S. M., Murch, S. J., and Saxena, P. K. (2014). Role of melatonin in alleviating cold stress in *Arabidopsis thaliana*. *J. Pineal Res.* 56, 238–245. doi: 10.1111/jpi.12115

Bates, L. S., Waldren, R. P., and Teare, I. D. (1973). Rapid determination of free proline for water-stress studies. *Plant Soil.* 39, 205–207. doi: 10.1007/BF00018060

Bolouri, M. R., Le Roy, K., Xiang, L., Rolland, F., and Van den Ende, W. (2010). Sugar signalling and antioxidant network connections in plant cells. *FEBS J.* 277, 2022–2037. doi: 10.1111/j.1742-4658.2010.07633.x

Braune, A., Engst, W., Elsinghorst, P. W., Furtmann, N., Bajorath, J., Gütschow, M., et al. (2016). Chalcone isomerase from eubacterium ramulus catalyzes the ring contraction of flavanonols. *J. Bacteriol.* 198, 2965–2974. doi: 10.1128/JB.00490-16

Campos, C. N., Ávila, R. G., de Souza, K. R. D., Azevedo, L. M., and Alves, J. D. (2019). Melatonin reduces oxidative stress and promotes drought tolerance in young *Coffea arabica* l. plants. *Agr. Water Manage.* 211, 37–47. doi: 10.1016/j.agwat.2018.09.025

Cao, L., Jin, X., Zhang, Y., Zhang, M., and Wang, Y. (2020). Transcriptomic and metabolomic profiling of melatonin treated soybean (*Glycine max* l.) under drought stress during grain filling period through regulation of secondary metabolite biosynthesis pathways. *PloS One* 15, e239701. doi: 10.1371/journal.pone.0239701

Cao, L., Qin, B., and Zhang, Y. X. (2021). Exogenous application of melatonin may contribute to enhancement of soybean drought tolerance *via* its effects on glucose metabolism. *Biotechnol. Biotec. Eq.* 35, 964–976. doi: 10.1080/13102818.2021.1941254

Cen, H., Wang, T., Liu, H., Tian, D., and Zhang, Y. (2020). Melatonin application improves salt tolerance of alfalfa (*Medicago sativa* l.) by enhancing antioxidant capacity. *Plants* 9, 220. doi: 10.3390/plants9020220

Che, Y. H., and Li, L. H. (2007). Genetic diversity of prolamines in *Agropyron mongolicum* keng indigenous to northern China. *Genet. Resour. Crop Ev.* 54, 1145–1151. doi: 10.1007/s10722-006-9006-7

Choi, B., Lim, D., Lee, J., Gao, S. S., Kwon, D. Y., Kim, B. R., et al. (2008). Luteolin suppresses cisplatin-induced apoptosis in auditory cells: possible mediation through induction of heme oxygenase-1 expression. *J. Med. Food.* 11, 230–236. doi: 10.1089/jmf.2007.591

Cui, G., Zhao, X., Liu, S., Sun, F., Zhang, C., and Xi, Y. (2017). Beneficial effects of melatonin in overcoming drought stress in wheat seedlings. *Plant Physiol. Biochem.* 118, 138–149. doi: 10.1016/j.plaphy.2017.06.014

Debnath, B., Islam, W., Li, M., Sun, Y., Lu, X., Mitra, S., et al. (2019). Melatonin mediates enhancement of stress tolerance in plants. *Int. J. Mol. Sci.* 20, 1040. doi: 10.3390/ijms20051040

Deluc, L. G., Quilici, D. R., Decendit, A., Grimplet, J., Wheatley, M. D., Schlauch, K. A., et al. (2009). Water deficit alters differentially metabolic pathways affecting important flavor and quality traits in grape berries of *Cabernet sauvignon* and *Chardonnay*. *BMC Genomics* 10, 212. doi: 10.1186/1471-2164-10-212

Egert, A., Keller, F., and Peters, S. (2013). Abiotic stress-induced accumulation of raffinose in *Arabidopsis* leaves is mediated by a single raffinose synthase (RS5, At5g40390). *BMC Plant Biol.* 13, 218. doi: 10.1186/1471-2229-13-218

Farooq, M., Hussain, M., Wahid, A., and Siddique, K. H. M. (2012). *Drought stress in plants: An overview* (Berlin, Heidelberg: Springer Berlin Heidelberg), pp1–p33.

Guo, Y. Y., Li, H. J., Zhao, C. F., Xue, J. Q., and Zhang, R. H. (2020). Exogenous melatonin improves drought tolerance in maize seedlings by regulating photosynthesis and the ascorbate-glutathione cycle. *Russ. J. Plant Physiol.* 67, 809–821. doi: 10.1134/S1021443720050064

Hattori, A., Migitaka, H., Iigo, M., Itoh, M., Yamamoto, K., Hara, M., et al. (1995). Identification of melatonin in plants and its effects on plasma melatonin levels and binding to melatonin receptors in vertebrates. *Biochem. Mol. Biol. Int.* 35, 627.

Holmström, K., Mäntylä, E., Welin, B., Mandal, A., Palva, E. T., Tunnela, O. E., et al. (1996). Drought tolerance in tobacco. *Nature* 379, 683–684. doi: 10.1038/379683a0

Hossain, M. S., Li, J., Sikdar, A., Hasanuzzaman, M., Uzizerimana, F., Muhammad, I., et al. (2020). Exogenous melatonin modulates the physiological and biochemical mechanisms of drought tolerance in tartary buckwheat (*Fagopyrum tataricum* (L.) gaertn). *Molecules* 25, 2828. doi: 10.3390/molecules25122828

Hu, Z., Fan, J., Xie, Y., Amombo, E., Liu, A., Gitau, M. M., et al. (2016). Comparative photosynthetic and metabolic analyses reveal mechanism of improved cold stress tolerance in bermudagrass by exogenous melatonin. *Plant Physiol. Biochem.* 100, 94–104. doi: 10.1016/j.plaphy.2016.01.008

Hu, W., Zhang, J., Wu, Z., Loka, D. A., Zhao, W., Chen, B., et al. (2022). Effects of single and combined exogenous application of abscisic acid and melatonin on cotton carbohydrate metabolism and yield under drought stress. *Ind. Crop Prod.* 176, 114302. doi: 10.1016/j.indcrop.2021.114302

Ibrahim, M. F. M., Elbar, O. H. A., Farag, R., Hikal, M., El-Kelish, A., El-Yazied, A. A., et al. (2020). Melatonin counteracts drought induced oxidative damage and stimulates growth, productivity and fruit quality properties of tomato plants. *Plants* 9, 1276. doi: 10.3390/plants9101276

Jafari, M., and Shahsavar, A. (2021). The effect of foliar application of melatonin on changes in secondary metabolite contents in two citrus species under drought stress conditions. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.692735

Khan, F. S., Gan, Z., Li, E., Ren, M., Hu, C., and Zhang, J. (2021). Transcriptomic and physiological analysis reveals interplay between salicylic acid and drought stress in citrus tree floral initiation. *Planta* 255, 24. doi: 10.1007/s00425-021-03801-2

Lerner, A. B., Case, J. D., Takahashi, Y., Lee, T. H., and Mori, W. (1958). Isolation of melatonin, the pineal gland factor that lightens melanocytes1. *J. Am. Cheml. Soc* 80, 2587. doi: 10.1021/ja01543a060

Liang, B., Ma, C., Zhang, Z., Wei, Z., Gao, T., Zhao, Q., et al. (2018). Long-term exogenous application of melatonin improves nutrient uptake fluxes in apple plants under moderate drought stress. *Environ. Exp. Bot.* 155, 650–661. doi: 10.1016/j.envexpbot.2018.08.016

Liang, C., Zheng, G., Li, W., Wang, Y., Hu, B., Wang, H., et al. (2015). Melatonin delays leaf senescence and enhances salt stress tolerance in rice. *J. Pineal Res.* 59, 91–101. doi: 10.1111/jpi.12243

Li, X., Brestic, M., Tan, D. X., Zivcak, M., Zhu, X., Liu, S., et al. (2018a). Melatonin alleviates low PSI-limited carbon assimilation under elevated CO_2 and enhances the cold tolerance of offspring in chlorophyll b-deficient mutant wheat. *J. Pineal Res.* 64, n/a–n/a. doi: 10.1111/jpi.12453

Li, J. J., Zeng, L., Cheng, Y., Lu, G. Y., Fu, G. P., Ma, H. Q., et al. (2018b). Exogenous melatonin alleviates damage from drought stress in *brassica napus* l. (Rapeseed) seedlings. *Acta Physiol. Plant* 40, 1–11. doi: 10.1007/s11738-017-2601-8

Lim, J. H., Park, H., Choi, J., Lee, I., and Choi, H. J. (2007). Isoorientin induces Nrf2 pathway-driven antioxidant response through phosphatidylinositol 3-kinase signaling. *Arch. Pharm. Res.* 30, 1590–1598. doi: 10.1007/BF02977329

Liu, H., Wang, P., Hu, Y., Gong, G., Zhao, G., Song, J., et al. (2013). Construction of an RNAi expression vector and transformation into penicillium chrysogenum. *Ann. Microbiol.* 64, 113–120. doi: 10.1007/s13213-013-0639-4

Liu, J., Wang, W., Wang, L., and Sun, Y. (2015). Exogenous melatonin improves seedling health index and drought tolerance in tomato. *Plant Growth Regul.* 77, 317–326. doi: 10.1007/s10725-015-0066-6

Liu, J., Zhang, R., Sun, Y., Liu, Z., Jin, W., and Sun, Y. (2016). The beneficial effects of exogenous melatonin on tomato fruit properties. *Sci. Hortic.* 207, 14–20. doi: 10.1016/j.scienta.2016.05.003

Manchester, L. C., Tan, D., Reiter, R. J., Park, W., Monis, K., and Qi, W. (2000). High levels of melatonin in the seeds of edible plants: Possible function in germ tissue protection. *Life Sci.* 67, 3023–3029. doi: 10.1016/S0024-3205(00)00896-1

Mierziak, J., Kostyn, K., and Kulma, A. (2014). Flavonoids as important molecules of plant interactions with the environment. *Molecules* 19, 16240–16265. doi: 10.3390/molecules191016240

Moustafa, F. M., Mahmoud, A., Arnao, M. B., Sheteiwy, M. S., Dafea, M., Soltan, M., et al. (2020). Melatonin-induced water stress tolerance in plants: Recent advances. *Antioxidants* 9, 809. doi: 10.3390/antiox9090809

Muir, S. R., Collins, G. J., Robinson, S., Hughes, S. G., Bovy, A. G., Vos, C. H. R. D., et al. (2001). Overexpression of petunia chalcone isomerase in tomato results in fruit containing increased levels of flavonols. *Nat. Biotechnol.* 19, 470–474. doi: 10.1038/88150

Naghizadeh, M., Kabiri, R., Hatami, A., Oloumi, H., Nasibi, F., and Tahmasei, Z. (2019). Exogenous application of melatonin mitigates the adverse effects of drought stress on morpho-physiological traits and secondary metabolites in Moldavian balm (*Dracocephalum moldavica*). *Physiol. Mol. Biol. Pla.* 25, 881–894. doi: 10.1007/s12298-019-00674-4

Niu, K., Ma, X., Liang, G., Ma, H., Jia, Z., Liu, W., et al. (2017). 5-aminolevulinic acid modulates antioxidant defense systems and mitigates drought-induced damage in Kentucky bluegrass seedlings. *Protoplasma* 254, 2083–2094. doi: 10.1007/s00709-017-1101-4

Okuda, T. T., Matsuda, Y. Y., Yamanaka, A. A., and Sagisaka, S. S. (1991). Abrupt increase in the level of hydrogen peroxide in leaves of winter wheat is caused by cold treatment 1. *Plant Physiol.* 97, 1265–1267. doi: 10.1104/pp.97.3.1265

Puckette, M. C., Weng, H., and Mahalingam, R. (2007). Physiological and biochemical responses to acute ozone-induced oxidative stress in *Medicago truncatula*. *Plant Physiol. Biochem.* 45, 70–79. doi: 10.1016/j.plaphy.2006.12.004

Rakwal, R., Agrawal, G. K., Yonekura, M., and Kodama, O. (2000). Naringenin 7-o-methyltransferase involved in the biosynthesis of the flavanone phytoalexin sakuranetin from rice (*Oryza sativa* l.). *Plant Sci.* 155, 213–221. doi: 10.1016/S0168-9452(00)00223-5

Ramakrishna, A., Giridhar, P., Sankar, K. U., and Ravishankar, G. A. (2012). Melatonin and serotonin profiles in beans of coffea species. *J. Pineal Res.* 52, 470–476. doi: 10.1111/j.1600-079X.2011.00964.x

Reiter, R. J., Mayo, C., Tan, D. X., Sainz, R., Alatorre-Jimenez, M., and Qin, L. (2016). Melatonin as an antioxidant: Under promises but over delivers. *J. Pineal Res.* 61, 253–278. doi: 10.1111/jpi.12360

Ren, J., Yang, X., Ma, C., Wang, Y., and Zhao, J. (2021). Melatonin enhances drought stress tolerance in maize through coordinated regulation of carbon and nitrogen assimilation. *Plant Physiol. Biochem.* 167, 958–969. doi: 10.1016/j.plaphy.2021.09.007

Rollins, J. A., Habte, E., Templer, S. E., Colby, T., Schmidt, J., and Korff, M. (2013). Leaf proteome alterations in the context of physiological and morphological responses to drought and heat stress in barley (*Hordeum vulgare* l.). *J. Exp. Bot.* 64, 3201–3212. doi: 10.1093/jxb/ert158

Romero, C., Belles, J. M., Vaya, J. L., Serrano, R., and Culianez-Macia, F. A. U. P. (1997). Expression of the yeast trehalose-6-phosphate synthase gene in transgenic tobacco plants: Pleiotropic phenotypes include drought tolerance. *Planta* 201, 293–297. doi: 10.1007/s004250050069

Song, Z., Yang, Q., Dong, B., Li, N., Wang, M., Du, T., et al. (2022). Melatonin enhances plant stress tolerance by promoting flavonoid enrichment, focusing on luteolin for salt stress. *J. Exp. Bot.* 73, 5992–6008. doi: 10.1093/jxb/erac276

Sun, C., Yang, C., and Tzen, J. T. C. (2018). Molecular identification and characterization of hydroxycinnamoyl transferase in tea plants (*Camellia sinensis* l.). *Int. J. Mol. Sci.* 19, 3938. doi: 10.3390/ijms19123938

Sun, Q., Zhang, N., Wang, J., Zhang, H., Li, D., Shi, J., et al. (2015). Melatonin promotes ripening and improves quality of tomato fruit during postharvest life. *J. Exp. Bot.* 66, 657–668. doi: 10.1093/jxb/eru332

Suzuki, S., Nakatsubo, T., Umezawa, T., and Shimada, M. (2002). First *in vitro* norlignan formation with asparagus officinalis enzyme preparation. *Chem. Commun.* 1088-1089. doi: 10.1039/b200217e

Suzuki, S., Yamamura, M., Shimada, M., and Umezawa, T. (2004). A heartwood norlignan, (E)-hinokiresinol, is formed from 4-coumaryl 4-coumarate by a cryptomeria japonica enzyme preparation. *Chem. Commun.*, 2838–2839. doi: 10.1039/b409686j

Tamogami, S., Rakwal, R., and Kodama, O. (1997). Phytoalexin production by amino acid conjugates of jasmonic acid through induction of naringenin-7-O-methyltransferase, a key enzyme on phytoalexin biosynthesis in rice (*Oryza sativa* l.). *FEBS Lett.* 401, 239–242. doi: 10.1016/S0014-5793(96)01482-2

Tan, Y., Xiao, L., Zhao, J., Sun, Y., and Bai, L. (2015). Molecular and functional characterization of the ecdysone receptor isoform-a from the cotton mirid bug, *Apolygus lucorum* (Meyer-dür). *Gene* 574, 88–94. doi: 10.1016/j.gene.2015.07.085

Tian, Q., Wang, S., Du, J., Wu, Z., Li, X., and Han, B. (2016). Reference genes for quantitative real-time PCR analysis and quantitative expression of P5CS in *Agropyron mongolicum* under drought stress. *J. Integr. Agr.* 15, 2097–2104. doi: 10.1016/S2095-3119(15)61238-2

Venisse, J., Gullner, G., and Brisset, M. (2001). Evidence for the involvement of an oxidative stress in the initiation of infection of pear by erwinia amylovora. *Plant Physiol.* 125, 2164–2172. doi: 10.1104/pp.125.4.2164

Wang, F., Chen, L., Chen, S., Chen, H., and Liu, Y. (2021). Characterization of two closely related citrus cultivars using UPLC-ESI-MS/MS-based widely targeted metabolomics. *PloS One* 16, e254759. doi: 10.1371/journal.pone.0254759

Wang, Y., Reiter, R. J., and Chan, Z. (2018). Phytomelatonin: A universal abiotic stress regulator. *J. Exp. Bot.* 69, 963–974. doi: 10.1093/jxb/erx473

Wang, P., Sun, X., Li, C., Wei, Z., Liang, D., and Ma, F. (2013). Long-term exogenous application of melatonin delays drought-induced leaf senescence in apple. *J. Pineal Res.* 54, 292–302. doi: 10.1111/jpi.12017

Wang, A., Zhu, M., Luo, Y., Liu, Y., Li, R., Kou, M., et al. (2017). A sweet potato cinnamate 4-hydroxylase gene, IbC4H, increases phenolics content and enhances drought tolerance in tobacco. *Acta Physiol. Plant* 39, 1–12. doi: 10.1007/s11738-017-2551-1

Wu, W., Zhang, Q., Ervin, E. H., Yang, Z., and Zhang, X. (2017). Physiological mechanism of enhancing salt stress tolerance of perennial ryegrass by 24-epibrassinolide. *Front. Plant Sci.* 8, 1017. doi: 10.3389/fpls.2017.01017

Yang, S., Bai, M., Hao, G., Zhang, X., Guo, H., and Fu, B. (2021a). Transcriptome survey and expression analysis reveals the adaptive mechanism of 'Yulu xiang' pear in response to long-term drought stress. *PloS One* 16, e246070. doi: 10.1371/journal.pone.0246070

Yang, W., Li, N., Fan, Y., Dong, B., Song, Z., Cao, H., et al. (2021b). Transcriptome analysis reveals abscisic acid enhancing drought resistance by regulating genes related to flavonoid metabolism in pigeon pea. *Environ. Exp. Bot.* 191, 104627. doi: 10.1016/j.envexpbot.2021.104627

Yang, J., Zhang, J., Wang, Z., and Zhu, Q. (2001). Activities of starch hydrolytic enzymes and sucrose-phosphate synthase in the stems of rice subjected to water stress during grain filling. *J. Exp. Bot.* 52, 2169–2179. doi: 10.1093/jexbot/52.364.2169

Yao, J. W., Ma, Z., Ma, Y. Q., Zhu, Y., Lei, M. Q., Hao, C. Y., et al. (2021). Role of melatonin in UV-b signaling pathway and UV-b stress resistance in *Arabidopsis thaliana*. *Plant Cell. Environ.* 44, 114–129. doi: 10.1111/pce.13879

Yuan, Y., Liu, Y., Wu, C., Chen, S., Wang, Z., Yang, Z., et al. (2012). Water deficit affected flavonoid accumulation by regulating hormone metabolism in *Scutellaria baicalensis georgi* roots. *PloS One* 7, e42946. doi: 10.1371/journal.pone.0042946

Zhao, Y., Gao, C., Shi, F., Yun, L., Jia, Y., and Wen, J. (2018). Transcriptomic and proteomic analyses of drought responsive genes and proteins in *Agropyron mongolicum* keng. *Curr. Plant Biol.* 14, 19–29. doi: 10.1016/j.cpb.2018.09.005

Zhao, H., Xu, L., Su, T., Jiang, Y., Hu, L., and Ma, F. (2015). Melatonin regulates carbohydrate metabolism and defenses against pseudomonas syringae pv. tomato DC3000 infection in *Arabidopsis thaliana*. *J. Pineal Res.* 59, 109–119. doi: 10.1111/jpi.12245

Zhu, Y., Gao, H., Lu, M., Hao, C., Pu, Z., Guo, M., et al. (2019). Melatonin-nitric oxide crosstalk and their roles in the redox network in plants. *Int. J. Mol. Sci.* 20, 6200. doi: 10.3390/ijms20246200

Zou, J., Yu, H., Yu, Q., Jin, X., Cao, L., Wang, M., et al. (2021). Physiological and UPLC-MS/MS widely targeted metabolites mechanisms of alleviation of drought stress-induced soybean growth inhibition by melatonin. *Ind. Crop Prod.* 163, 113323. doi: 10.1016/j.indcrop.2021.113323

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 09 January 2023
DOI 10.3389/fpls.2022.1095359

OPEN ACCESS

EDITED BY
Wengang Xie,
Lanzhou University, China

REVIEWED BY
Zongyu Zhang,
Lanzhou University, China
Victor Manuel Rodriguez,
Biological Mission of Galicia (CSIC), Spain

*CORRESPONDENCE
Marty J. Faville
marty.faville@agresearch.co.nz

†PRESENT ADDRESS
Sofie M. Pearson,
The University of Queensland, Warwick, QLD, Australia
Poppy P. Miller,
Plant &Food Research, Te Puke, New Zealand
Catherine M. McKenzie,
Plant &Food Research, Te Puke, New Zealand

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

RECEIVED 11 November 2022
ACCEPTED 09 December 2022
PUBLISHED 09 January 2023

CITATION
Pearson SM, Griffiths AG, Maclean P, Larking AC, Hong SW, Jauregui R, Miller PP, McKenzie CM, Lockhart PJ, Tate JA, Ford JL and Faville MJ (2023) Outlier analyses and genome-wide association study identify *glgC* and *ERD6-like 4* as candidate genes for foliar water-soluble carbohydrate accumulation in *Trifolium repens*. *Front. Plant Sci.* 13:1095359. doi: 10.3389/fpls.2022.1095359

Outlier analyses and genome-wide association study identify *glgC* and *ERD6-like 4* as candidate genes for foliar water-soluble carbohydrate accumulation in *Trifolium repens*

Sofie M. Pearson[1,2†], Andrew G. Griffiths[2], Paul Maclean[2], Anna C. Larking[2], S. Won Hong[2], Ruy Jauregui[2], Poppy Miller[2†], Catherine M. McKenzie[2†], Peter J. Lockhart[1], Jennifer A. Tate[1], John L. Ford[3] and Marty J. Faville[2*]

[1]School of Natural Sciences, Massey University, Palmerston North, New Zealand, [2]Resilient Agriculture, AgResearch Grasslands, Palmerston North, New Zealand, [3]Grasslands, PGG Wrightson Seeds Limited, Palmerston North, New Zealand

Increasing water-soluble carbohydrate (WSC) content in white clover is important for improving nutritional quality and reducing environmental impacts from pastoral agriculture. Elucidation of genes responsible for foliar WSC variation would enhance genetic improvement by enabling molecular breeding approaches. The aim of the present study was to identify single nucleotide polymorphisms (SNPs) associated with variation in foliar WSC in white clover. A set of 935 white clover individuals, randomly sampled from five breeding pools selectively bred for divergent (low or high) WSC content, were assessed with 14,743 genotyping-by-sequencing SNPs, using three outlier detection methods: PCAdapt, BayeScan and KGD-F_{ST}. These analyses identified 33 SNPs as discriminating between high and low WSC populations and putatively under selection. One SNP was located in the intron of *ERD6-like 4*, a gene coding for a sugar transporter located on the vacuole membrane. A genome-wide association study using a subset of 605 white clover individuals and 5,757 SNPs, identified a further 12 SNPs, one of which was associated with a starch biosynthesis gene, glucose-1-phosphate adenylyltransferase, *glgC*. Our results provide insight into genomic regions underlying WSC accumulation in white clover, identify candidate genomic regions for further functional validation studies, and reveal valuable information for marker-assisted or genomic selection in white clover.

KEYWORDS

genome-wide association study, genotyping-by-sequencing, outlier detection, white clover, water-soluble carbohydrate

1 Introduction

White clover (*Trifolium repens* L.) is sown in temperate pastures globally, as it provides high quality forage for ruminants and is a source of bioavailable nitrogen, fixed through symbiosis with soil *Rhizobium* bacteria (Ulyatt, 1997). It is a recent (15 – 28,000 years ago) allotetraploid that resulted from the hybridization of two diploid *Trifolium* species, *T. occidentale* and *T. pallescens* (Ellison et al., 2006; Griffiths et al., 2019). Retention of the combined genomes as *T. occidentale* and *T. pallescens*-derived subgenomes likely underpins the broad adaptation and phenotypic plasticity of this agronomically successful species (Griffiths et al., 2019). White clover foliage has a high concentration of crude protein but a relatively low concentration of water-soluble carbohydrate (WSC) (Cosgrove et al., 2009). Foliar WSC is important because it provides readily available energy to the rumen microbiome, which improves the efficiency of protein utilisation by the animal (Woodfield et al., 2001). Higher levels of WSC available for consumption by ruminant microbes enables a shift in the partitioning of digested nitrogen, with less excreted as urea and more utilised for animal growth and production (Easton et al., 2009). Breeding for increased foliar WSC in pasture species, including white clover, can therefore have beneficial effects for the environment as less nitrogen is lost *via* urine and dung to nitrous oxide emission and nitrate leaching (Luo et al., 2015; Selbie et al., 2015). In addition to improved nutritional quality and positive environmental outcomes, WSC has been found to be important in conferring cold tolerance and drought resistance in plants due to its role in osmotic adjustment (Kerepesi and Galiba, 2000; Dalmannsdóttir et al., 2001; Livingston et al., 2009).

Several aspects of WSC composition and variation in white clover plants have been studied, including seasonal and diurnal foliar WSC variation (Michell, 1973; Ruckle et al., 2018; Kagan et al., 2020). While research has addressed the genetic control of WSC accumulation in stolons (Inostroza et al., 2018), little is known about the genetic mechanisms underlying foliar WSC accumulation in white clover. Improved understanding of the genes that influence foliar WSC accumulation would support the development and application of molecular breeding tools, such as marker-assisted and genomic selection, that could be used by breeders to accelerate genetic improvement of this trait.

Two complementary approaches may be used to detect genomic loci linked to trait phenotype variation, as a means to identify candidate genes. Outlier analysis can be used to identify single nucleotide polymorphisms (SNPs) that differentiate populations with divergent phenotypes. This approach most commonly involves F_{ST}-based tests (Antao et al., 2008; Whitlock and Lotterhos, 2015) or principal component analyses (PCA) (Luu et al., 2017) to identify differentiating loci that are distinct from those under neutral selection. Genome-wide association studies (GWAS) offer a second approach, utilising SNP markers across the entire genome with the goal of associating specific variants with phenotypic variation, as measured in a population or a panel of diverse individuals. This method can be used in both model and non-model organisms and has successfully identified genes underlying traits in forage species (Arojju et al., 2016; Sakiroglu and Brummer, 2017; Biazzi et al., 2017). Both outlier approaches and GWAS require a preliminary assessment of population structure to avoid false positive associations (Sul et al., 2018).

Selective breeding to create experimental white clover populations with divergent levels of foliar WSC was previously undertaken in two New Zealand breeding programmes, with five discrete breeding pools. Three were described by Widdup et al. (2010) and conducted over four cycles of divergent recurrent selection, and the remaining two pools were part of another programme (Mr. John Ford, pers. comm.) in which selection took place over six cycles. Both programmes included pools in different leaf size classes (large leaf and small leaf). The divergently-selected populations in these pools represent a valuable genetic resource for investigating the genetic basis of WSC accumulation, including the relationship between leaf size and WSC levels (Woodfield et al., 2001).

Genotyping-by-sequencing (GBS), enables highly efficient and cost-effective simultaneous SNP discovery and genotyping (Elshire et al., 2011; Poland et al., 2012b) and has been implemented in numerous forage species (Biazzi et al., 2017; Sakiroglu and Brummer, 2017; Faville et al., 2018; Guo et al., 2018), including white clover (Wright et al., 2017; Griffiths et al., 2019). In our study, we applied GBS in the five white clover pools to support investigation of genomic regions and loci under selection for WSC accumulation. Analyses were based on genome-wide GBS-derived SNP data from individuals within the breeding pools, targeting three generational time points within each pool. The overall aim was to identify SNPs associated with foliar WSC accumulation, that may subsequently be developed and used to support gene discovery and molecular breeding approaches in white clover populations to aid breeding for increased WSC accumulation in white clover.

There were three principal objectives: (1) to confirm that foliar WSC phenotypes were significantly and directionally different amongst the populations used in the study, ensuring that subsequent genetic studies were performed on truly divergent phenotypes; (2) to establish whether selective breeding had altered WSC independently of changing leaf area in those populations; and (3) apply outlier detection and GWAS approaches to identify SNPs associated with foliar WSC accumulation.

Abbreviations: DF, Discriminant Function; DM, Dry Matter; EMMs, estimated marginal means; HMW, High Molecular Weight; *K*, number of genetic clusters; KGD, Kinship using Genotyping-by-sequencing with Depth adjustment; K_{PC}, number of principal components; LMW, Low Molecular Weight; PC, Principal Components; PCA, Principal Component Analysis; SSS, Soluble Sugars and Starch; WSC, Water-Soluble Carbohydrate.

2 Materials and methods

2.1 Plant material

The plant material used in this study was bred and supplied by Grasslands Innovation Ltd. Selective breeding for foliar levels of water-soluble carbohydrate (WSC) in white clover was completed previously in two breeding programmes, one consisting of four cycles of recurrent selection in three breeding pools (Widdup et al., 2010) and the other over six cycles in two pools (Mr John Ford, pers. comm.) (Figure 1). In all breeding pools, divergent selection was undertaken at each cycle to create populations with either low or high levels of foliar WSC, so that at each generation there is both a low and a high WSC population. In the programme described by Widdup et al. (2010) there were 24 populations generated (3 breeding pools × 4 cycles × low/high WSC) and in the Ford programme there were also 24 populations (2 breeding pools × 6 cycles × low/high WSC). Adding the five parental generations, a total of 53 populations were available for evaluation. Of these, 25 were chosen for phenotyping and genotypic analyses (Figure 1). These were the parental, middle and end generation populations within each pool. The middle generation was cycle 2 or cycle 3 and the end generation cycle 4 or 6, for the Widdup and Ford pools, respectively. Seed from all populations was acquired from the Margot Forde Germplasm Centre (Palmerston North, New Zealand).

Nomenclature for population names is: W = Widdup, F = Ford; NZ = New Zealand, US = United States of America; LL = large leaf, SL = small leaf; Low = low water-soluble carbohydrate (WSC), High = high WSC; P = parental generation, Mid = middle generation and End = end generation. The 25 populations were: WNZLL-Low-End, WNZLL-Low-Mid, WNZLL-Parent, WNZLL-High-Mid, WNZLL-High-End, WNZSL-Low-End, WNZSL-Low-Mid, WNZSL-Parent, WNZSL-High-Mid, WNZSL-High-End, WUSLL-Low-End, WUSLL-Low-Mid, WUSLL-Parent, WUSLL-High-Mid, WUSLL-High-End, FNZLL-Low-End, FNZLL-Low-Mid, FNZLL-Parent, FNZLL-High-Mid, FNZLL-High-End, FNZSL-Low-End, FNZSL-Low-Mid, FNZSL-Parent, FNZSL-High-Mid, and FNZSL-High-End.

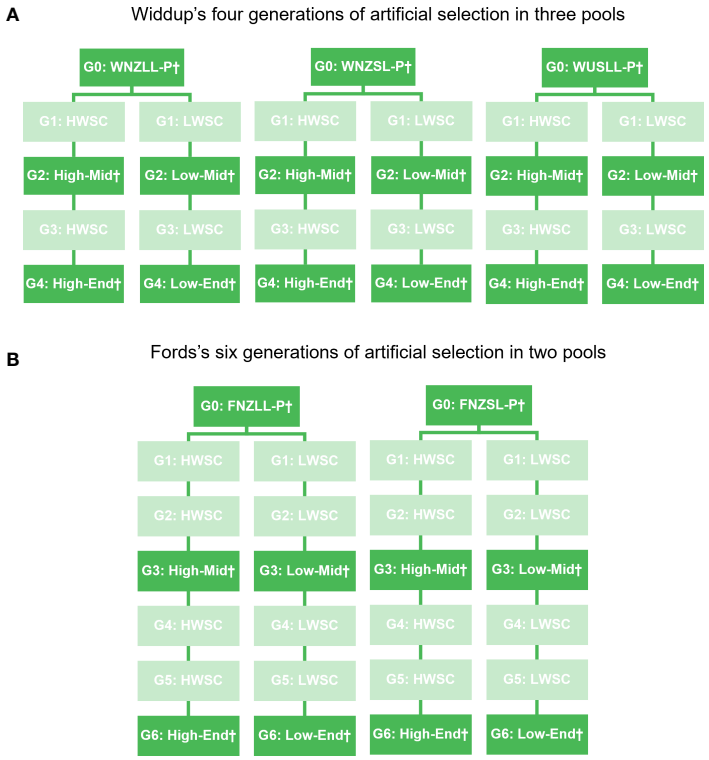


FIGURE 1
Schematic representation of white clover populations from the **(A)** Widdup et al. (2010) and **(B)** Ford breeding programmes. G, generation; W, Widdup; F, Ford; NZ, New Zealand/Aotearoa; US, United States of America; LL, large leaf; SL, small leaf; P, Parent generation; HWSC and High, high water-soluble carbohydrate (WSC); LWSC and Low, low WSC; Mid, Middle generation; and End, End generation. † denotes populations used in phenotyping and genotyping studies.

2.2 Population establishment and experimental design

Approximately 100 seeds from each of the 25 populations were germinated, planted into propagation trays and grown under standard greenhouse conditions for two months. A total of 900 plants from the 25 populations (180 plants per breeding pool) were then randomly selected and transplanted into 2L pots for phenotyping. In each breeding pool, 60 plants per parent population (Parent) were randomly selected, along with 30 plants each from the middle generation low and high WSC populations (Low-Mid, High-Mid) and 30 plants each from the end generation low and high WSC populations (Low-End, High-End). Potted plants were kept in the greenhouse to establish for two weeks before being placed outside at Palmerston North, New Zealand (40.38°S, 175.61°E), in autumn, late May 2017, in a randomised Latin square design with three replicate blocks. The block was set up on a 9 × 9 m concrete pad and plants were set at 30 cm centre-to-centre spacing.

2.3 Plant phenotyping

2.3.1 Water-soluble carbohydrate phenotyping using near infra-red reflectance spectroscopy

Leaves from the 900 plants were sampled over three consecutive days in November 2017. One replicate block was harvested per day, between 8:00 and 10:00 am to minimise diurnal variation in WSC levels and to be consistent with the methodology originally used for breeding these divergent WSC selections (Widdup et al., 2010). Thirty fully expanded, healthy leaf laminae were removed per plant to constitute a single plant sample, snap frozen in liquid nitrogen, and stored at -20°C before being freeze-dried and milled. Samples from two of the three replicate blocks (n = 600) were analysed by near infra-red spectroscopy (NIRS) at the Massey University Nutrition Laboratory, (Palmerston North, New Zealand) for nutritive quality attributes including WSC concentration (g kg^{-1} dry matter). Two NIRS calibrations were used initially to determine WSC: (a) total soluble sugars and starch (SSS) (Corson et al., 1999), the calibration originally used in the Widdup et al. (2010) and Ford breeding programmes; (b) WSC-NIRS; the sum of high molecular weight (HMW)- and low molecular weight (LMW)-WSC fractions estimated using a calibration developed from perennial ryegrass (Cosgrove et al., 2009).

2.3.2 Leaf area phenotyping

Leaf area was assessed for each of the five pools. A total of 450 of the 900 plants were sampled from every second column, across all three blocks (150 plants per block). Four leaves per plant (first leaves from the stolon tip with fully opened laminae) were collected, glued to 1 mm graph paper and scanned. Scanned images were converted to binary in ImageJ (Schindelin et al., 2012), a global calibration curve was set to convert from pixel distance to actual distance (cm), and leaf area (cm^2) was estimated. The mean of four leaves was used as a proxy for average plant leaf area.

2.3.3 Phenotype statistical analyses and correlation investigation

Estimated Marginal Means for the WSC phenotype, using SSS, and leaf area were calculated for each population using a linear mixed model in the R package "*emmeans*" v 1.4.2 (Lenth et al., 2020) that accounted for treatment and spatial variation (**Equation 1**).

$$y \sim \mathrm{TrtC} + (1 \mid \mathrm{Block}) + (1 \mid \mathrm{Block:Column}) + (1 \mid \mathrm{Block:Row})$$

Where: y is the phenotypic trait, and *TrtC* is the interaction between Pool, Generation (i.e., Parent, Mid and End) and *Trt* (i.e., None (Parent; no selection), high WSC and low WSC).

Leaf area data were transformed by square root such that residuals conformed to constant variance and normality (Supplementary Figure 1). SSS required no transformation as residual assumptions were met. Leaf area data were back-transformed to derive population-fitted values on the original cm^2 scale. The Parent population was used as the baseline for comparisons between populations within each pool. Significant differences in these comparisons were investigated using "*emmeans*" at α = 0.05. The Low-End population was then used as the baseline for the SSS dataset so differences between the Low-End and High-End populations for each pool could be determined.

Three hundred plants common to the sets used for evaluation of WSC and leaf area were used for phenotypic correlation analysis (Pearson correlation). This dataset was then split further into five datasets corresponding to the five pools (WNZLL, WNZSL, WUSLL, FNZLL and FNZSL) within which correlations between the two variables were investigated. Prior to correlation analysis, leaf area and SSS data were evaluated for normality using the Shapiro-Wilk Normality Test implemented in R (Royston, 1995; R Core Team, 2019). Data that did not follow a normal distribution were transformed using Box-Cox Transformation analysis (Venables and Ripley, 2002). Models were created for each pool using R statistical software, and adjusted coefficient of determination (r^2) values of the lines of best fit were used to initially compare the predictive abilities of leaf area for SSS. Regression analysis amongst populations was then conducted using linear mixed models constructed for each pool. Models were constructed with SSS as the dependent variable and the interaction between leaf area and population was used as the independent variable. The Parent population for each pool was used as the baseline for

comparison. Adjusted r^2 was reported for both types of linear models (SSS ~ leaf area and SSS ~ 0 + Population * leaf area) to avoid false inflation of r^2 values. Residuals were checked for normality and assumptions were met; thus, no data transformation was required. The R scripts used to produce estimated phenotype means and perform correlation and regression analyses can be found at https://github.com/SofiePearson/White_Clover_WSC_Outlier_Detection_GWAS.

2.4 Plant genotyping

2.4.1 Genotyping-by-sequencing library preparation and sequencing

Genomic DNA was extracted from 1,536 white clover individuals (approximately 60 individuals from each of the 25 populations) using the freeze-dried tissue protocol described in Anderson et al. (2018). A subset of 1,175 individuals, comprising 47 from each of the 25 populations, were chosen for genotyping-by-sequencing (GBS). Each GBS library consisted of 94 samples, plus one negative control (water) and one positive control DNA from an inbred white clover, "S9" (Griffiths et al., 2019). Thirteen GBS libraries were created in total, and any spare wells filled with duplicated DNA chosen at random from all populations, including duplication of samples within a library and duplication of samples among libraries. GBS libraries were constructed following Poland et al. (2012b) using restriction enzymes *PstI* and *MspI*, with some modifications (see Supplementary Methods). Each library was then sequenced in parallel on two lanes of a flow cell on an Illumina HiSeq 2500 (Illumina, San Diego, CA, USA) at Invermay Agricultural Centre (AgResearch, Mosgiel, New Zealand).

2.4.2 SNP calling, filtering and genotyping-by-sequencing library quality control

Raw data FASTQ files containing sequence reads were processed for SNP identification using Trait Analysis by aSSociation, Evolution and Linkage (TASSEL) v 5.0 (Glaubitz et al., 2014), using default parameters except minor allele frequency was set to 0.01. An AgResearch white clover genome assembly (version 5) was used as the reference genome (Griffiths et al., 2019). Raw sequence data from 1,222 samples, 13 positive controls and 13 negative samples were analysed together. Sequence reads were first trimmed to 64 bp and identical reads were grouped into sequence tags. The sequence tags were then aligned to the reference genome using Burrows-Wheeler Alignment (BWA) tool (Li and Durbin, 2009).

After SNP calling, all filtering was performed using VCFtools v 0.1.16 (Danecek et al., 2011). To provide a SNP marker dataset for downstream analyses, the marker set was restricted to high quality SNPs: biallelic SNPs with read depth range of 5 – 150, missing genotype data ≤ 20% per SNP, and minor allele frequency ≥ 0.03. Samples with a large proportion of missing data were removed from the dataset and negative control samples were removed after checking that they did not contain unduly high levels of sequence data.

2.5 Analysis of population genetic structure and variation

Genetic structure in the dataset was explored by Discriminant Analysis of Principal Components (DAPC) (Jombart et al., 2010), implemented in "*adegenet*" v 2.1.1 (Jombart, 2008) for R software v 3.6.1 (R Core Team, 2019). The analysis used 14,743 SNPs from 1,113 individuals in 24 populations (WNZSL-Parent population excluded due to high sample missing data). *K*-means clustering was used to detect the number of clusters amongst the samples, without prior assumptions of assignment based on pool or population. *K*-means was run sequentially from 1 to 40 genetic clusters (*K*), with 800 principal components retained, accounting for approximately 90% of the total genetic variation. The optimal clustering solution corresponded with the lowest Bayesian Information Criterion (BIC). Individual assignment from the *a priori* grouping to the *K*-means determined clusters was visualised and compared using the R package "*ade4*" v 1.7-15 (Chessel et al., 2004).

DAPC analysis was implemented using the cross-validation "*xvalDapc()*" function, and was run with 100 replicates from 1 – 50 principal components (PCs) with individuals grouped based on the *K*-means determined number of clusters. This cross-validation method determined 6 PCs should be used, hence the final "*dacp()*" was run with "n.pca = 6" and 6 discriminant functions (DFs) retained. A scatter plot of individuals grouped by *K*-means on DFs was created using "*adegenet*" (Jombart, 2008).

Genetic variation within and among populations was assessed by Analysis of Molecular Variance (AMOVA) implemented with "*poppr*" v 2.9.1 (Kamvar et al., 2014) in R. This was conducted for all 24 populations, as well as among genetic clusters identified using the most supported DAPC *K*-value ($K = 11$) with 9,999 permutations (Excoffier et al., 1992; Kamvar et al., 2014). A matrix of pairwise genetic differentiation between all population pairs from two possible population structures (the *a priori* of $K = 24$ and the DAPC-determined grouping of $K = 11$) was also computed using the R package "*hierfstat*" v 0.04-22 (Goudet and Jombart, 2015) using the fixation index F_{ST} (Weir and Cockerham, 1984). The R scripts used to perform population genetic structure analyses can be found at https://github.com/SofiePearson/White_Clover_WSC_Outlier_Detection_GWAS.

2.6 Detection of loci under selection

Three approaches were used to analyse the 14,743 SNP dataset for loci under divergent selection: PCAdapt (Luu et al., 2017), BayeScan (Foll and Gaggiotti, 2008) and an F_{ST} outlier detection approach in the package Kinship using Genotyping-by-sequencing with Depth adjustment (KGD, available from https://github.com/AgResearch/KGD.git) (Dodds et al., 2015). Missing data for SNPs were not imputed.

2.6.1 PCAdapt

Individuals from the DAPC analysis were split into five datasets using VCFtools (Danecek et al., 2011) to enable identification of outlier loci differentiating high and low WSC populations within pools. These datasets consisted of five VCF files with individuals split into their respective pools but with the Parent populations removed, resulting in a total of 935 individuals retained for outlier detection analyses. All VCF files were converted into PLINK format (BED, BIM and FAM files) using PLINK v 1.9 (Purcell et al., 2007). The R package "*pcadapt*" v 4.0.3 was used to detect loci driving variation on the principal components (Luu et al., 2017). The K_{PC} value (number of principal components) with the best fit to the data was determined using the scree test (Cattell, 1966) and interpretation of score plots with K_{PC} values higher than those determined by the scree test method. To determine the number of principal components (K_{PC}) separating high and low WSC populations, each of the five pools was analysed in PCAdapt separately. Scree plots were produced (Supplementary Figure 2) and visually assessed for optimal K_{PC} value (Cattell, 1966), with components retained from the steep portion of the curve, prior to inflection into a flat line. Outlier SNPs based on Mahalanobis distance (Luu et al., 2017) at the optimal K_{PC} value were identified, after correcting for false positives using the Bonferroni correction in each pool. SNPs were visualised on Manhattan plots using "*qqman*" v 0.1.4 (Turner, 2018), with outliers shown as exceeding Bonferroni false discovery thresholds, of $\alpha = 0.01$ and $\alpha = 0.05$.

2.6.2 BayeScan

The five VCF files were converted into BayeScan format in R using "*vcfR*" v 1.8.0, "*adegenet*" v 2.1.1, and "*hierfstat*" v 0.04-22 (Jombart, 2008; Goudet and Jombart, 2015; Knaus and Grünwald, 2017). Population structure analysis identified little genetic differentiation between generations within the low WSC and high WSC divergent selections, respectively. Therefore, for this analysis within each pool the two high WSC populations were merged and two low WSC populations were merged, resulting in five population pairs to be tested. The analysis was conducted separately for each pool so that SNPs under divergent selection could be traced back to each pool. BayeScan was run with default parameters (20 pilot runs with 5,000 iterations, followed by a burn-in of 50,000 iterations, and prior odds for the neutral model was 10). The *q*-value for each locus was calculated and a false discovery rate of $\alpha = 0.05$ was used to determine significant outlier loci that had positive alpha values.

2.6.3 KGD-F_{ST}

The filtered VCF file was converted to a Reference Alternative file using the KGD *vcf2ra_ro_ao.py* python script, and a separate file containing individual and population information was constructed. The "*Fst.GBS.pairwise()*" function was used to calculate approximate mean F_{ST} for each SNP between each population pair, accounting for GBS read depth (Dodds et al., 2015). The five population pairs were tested and SNPs with F_{ST} values greater than 0.3 and present in more than two pools at that threshold were called as outlier SNPs. Manhattan plots were created as described above. Outlier SNPs from all three analyses (PCAdapt, BayeScan and KGD-F_{ST}) were visualised in a Venn diagram using the R package "*VennDiagram*" v 1.6.20 (Chen and Boutros, 2011). The scripts used to perform the outlier detection analyses can be found at https://github.com/SofiePearson/White_Clover_WSC_Outlier_Detection_GWAS.

2.7 Genome-wide association

A mixed-linear model implemented in the R package "*rrBLUP*" (Endelman, 2011) was used for an association analysis on a subset of individuals ($n = 605$) possessing both genotypic and phenotypic information (Supplementary Figure 3). Markers for this analysis were filtered to retain those with ≤ 50% missing data before the "*A.mat()*" function was used to impute missing values using the EM algorithm designed for GBS markers (Poland et al., 2012a). This resulted in 5,757 SNPs used in the analysis. Population structure and family relatedness was accounted for with "n.PC = 2" and a kinship matrix calculated by rrBLUP from the genotypic data. To account for multiple testing, a Bonferroni correction was applied and markers passing the threshold at an α of 0.05 were considered statistically significant (Bonferroni, 1936). Manhattan and Quantile-Quantile (Q-Q) plots were created for each phenotypic trait: leaf area, soluble sugars and starch (SSS), water-soluble carbohydrate (WSC) and other nutritional attributes including ash, crude protein, neutral detergent fibre, acid detergent fibre and lipid content. The R script used to perform GWAS can be found at https://github.com/SofiePearson/White_Clover_WSC_Outlier_Detection_GWAS.

2.8 Changes in genotypes due to selection over time

As a complement to the methods described above, changes in genotype frequencies from generation to generation were evaluated. Each outlier SNP detected by ≥2 outlier analyses, in each population, was assessed in this way. Genotype proportions for each SNP were extracted using VCFtools v 0.1.16 *–extract-FORMAT-info* GT and patterns were investigated (Danecek et al., 2011).

2.9 Identification of candidate genes

The physical positions of outlier SNPs identified by ≥2 outlier analyses were used to locate potential candidate genes. For outlier SNPs located to introns or exons, the host gene was recorded as the best candidate. SNPs identified in coding regions of a gene were investigated further to determine if they were likely to affect protein function or structure. Geneious Prime v 2019.1.1 (http://www.geneious.com/) was used to determine the position of the SNP in the protein and whether there was a synonymous or non-synonymous change. For outlier SNPs that occurred outside of genes, a maximum distance of 10 Kbp either side of each SNP were recorded. When considering potential candidate genes, the upstream and downstream regulatory elements of the gene, including the promoter region, were also considered. If intergenic outlier SNPs were located less than 1 Kbp away from the start codon of a gene, they were also classified as putatively in linkage disequilibrium with the gene due to the proximity to a promoter. White clover genome annotations (Griffiths et al., 2019), BLAST (Johnson et al., 2008), UniProt (The Uniprot Consortium, 2018) and STRING v 11.0 (Szklarczyk et al., 2019) were used to identify genes and their functions.

3 Results

3.1 Phenotypic variation

3.1.1 Water-soluble carbohydrate phenotyping

Water-soluble carbohydrate (WSC) content was measured using two NIRS calibrations (soluble sugars and starch, SSS; and WSC-NIRS), which were found to be highly correlated within the sample set ($r^2 = 0.92$, $p < 0.0001$), and therefore subsequent analyses focused principally on SSS alone. Population fitted values for SSS are presented in Figure 2. Comparison of population fitted values within each pool showed an overall trend for SSS to increase by selection cycle for high WSC selections and, conversely, decrease by selection cycle for low

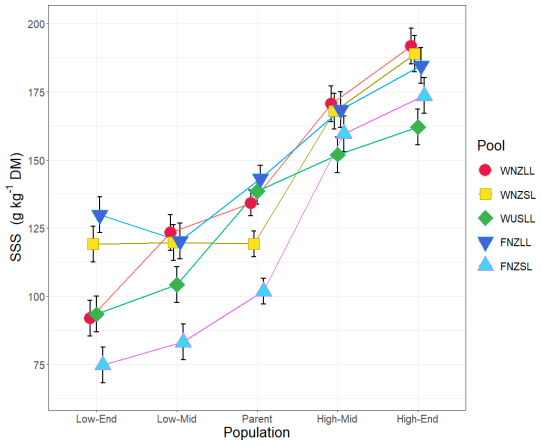


FIGURE 2
Population fitted values (adjustment for treatment, block, row and column effects) and standard error, for soluble sugars and starch (SSS) as a measure of water-soluble carbohydrate (WSC). Populations are grouped by pool as indicated by colour and symbol combinations. The *x*-axis indicates a timeline by generation, where: Low, low WSC; High, high WSC; End, End generation; Mid, Middle generation; Parent, Parent generation. Population means for SSS are based on $n = 20 - 40$.

WSC selections (Table 1). When averaged across all pools, there were significant differences ($p < 0.01$) for all comparisons, with the low WSC populations lower for SSS than the Parent, and the high WSC populations higher for SSS content than the Parent (Table 1). Within individual pools, high and low WSC populations did not always differ significantly from the Parent population, but in all pools there were significant ($p < 0.05$) differences between the Low-End and High-End populations (mean difference of 78.3 g kg^{-1} DM) (Table 1). These results confirm that breeding for divergent WSC in the five pools was successful, with a mean 76.9% difference in SSS between the Low-End and High-End populations (Supplementary Table 1).

3.1.2 Leaf area phenotyping

Population fitted values for mean leaf area (cm^2) are presented in Supplementary Table 2. When compared against Parent population values, there was a trend for leaf area to increase or decrease with WSC selection, but there were only four instances where that change was statistically significant ($p < 0.05$) (Table 1). The FNZLL (-5.4 cm^2, $p < 0.01$) and WUSLL (-4.4 cm^2, $p = 0.02$) pools showed a significant decrease in leaf area from the Parent to Low-End population. In the WNZSL pool there was a significant increase in leaf area in both the High-Mid and High-End populations relative to the Parent population (+5.9 cm^2, $p < 0.01$ and +5.5 cm^2, $p < 0.01$, respectively).

TABLE 1 Estimated phenotype means for each divergently-selected population compared to the parental mean after adjusting for treatment, block, row and column effects.

	SSS (g kg^{-1} DM)			Leaf area (cm^2)		
Comparison	Difference	SE	*p*-value	Difference	SE	*p*-value
WNZLL-Low-End – WNZLL-Parent	-42.2	7.77	<0.01**	-0.83	0.16	1
WNZLL-Low-Mid – WNZLL-Parent	-10.9	7.78	0.97	0.6	0.16	1
WNZLL-High-Mid – WNZLL-Parent	36.4	7.77	<0.01**	2.74	0.16	0.37
WNZLL-High-End – WNZLL-Parent	57.5	7.77	<0.01**	2.92	0.16	0.27
WNZSL-Low-End – WNZSL-Parent	-0.098	7.79	1	0.56	0.16	1
WNZSL-Low-Mid – WNZSL-Parent	0.44	7.78	1	1.81	0.16	0.85
WNZSL-High-Mid – WNZSL-Parent	48.7	7.76	<0.01**	5.89	0.16	<0.01**
WNZSL-High-End – WNZSL-Parent	69.6	7.78	<0.01**	5.48	0.16	<0.01**
WUSLL-Low-End – WUSLL-Parent	-45.1	7.8	<0.01**	-4.41	0.17	0.02*
WUSLL-Low-Mid – WUSLL-Parent	-34.2	7.77	<0.01**	-1.65	0.16	0.99
WUSLL-High-Mid – WUSLL-Parent	13.4	7.76	0.82	1.14	0.16	1
WUSLL-High-End – WUSLL-Parent	23.6	7.76	0.054	-0.87	0.16	1
FNZLL-Low-End – FNZLL-Parent	-13.4	7.77	0.82	-5.37	0.16	<0.01**
FNZLL-Low-Mid – FNZLL-Parent	-22.8	7.79	0.07	-3.49	0.16	0.15
FNZLL-High-Mid – FNZLL-Parent	25.2	7.78	0.029*	0.72	0.16	1
FNZLL-High-End – FNZLL-Parent	41.4	7.79	<0.01**	1.19	0.16	1
FNZSL-Low-End – FNZSL-Parent	-27.1	7.76	0.012*	-1.5	0.16	0.94
FNZSL-Low-Mid – FNZSL-Parent	-18.7	7.77	0.29	0	0.16	1
FNZSL-High-Mid – FNZSL-Parent	57.8	7.78	<0.01**	-0.75	0.16	1
FNZSL-High-End – FNZSL-Parent	71.6	7.77	<0.01**	-1.77	0.16	0.8
Low-End – Parent	-25.6	3.47	<0.01**	-2.18	0.07	<0.01**
Low-Mid – Parent	-17.2	3.49	<0.01**	-0.43	0.07	1
High-Mid – Parent	36.7	3.46	<0.01**	1.95	0.07	0.01**
High-End – Parent	52.7	3.48	<0.01**	1.35	0.07	0.24
WNZLL-High-End – WNZLL-Low-End	99.6	8.99	<0.01**			
WNZSL-High-End – WNZSL-Low-End	69.7	8.98	<0.01**			
WUSLL-High-End – WUSLL-Low-End	68.7	8.98	<0.01**			
FNZLL-High-End – FNZLL-Low-End	54.8	8.98	<0.01**			
FNZSL-High-End – FNZSL-Low-End	98.7	8.98	<0.01**			

Low, low water-soluble carbohydrate (WSC), High, high WSC; Parent, Parent generation; Mid, Middle generation; End, End generation; W, Widdup; F, Ford; NZ, New Zealand/Aotearoa; US, United States of America; LL, large leaf; SL, small leaf and SE, standard error.
Significance codes: ** ≤ 0.01, * = 0.01 – 0.05, no symbol ≥ 0.05 at $\alpha = 0.05$.
Soluble sugars and starch (SSS; grams per kilogram dry matter, g kg^{-1} DM) and leaf area (cm^2) population fitted values compared to the Parent population are presented in the "Difference" column. For SSS, the High-End population fitted values are compared to the Low-End population fitted values for each pool and are also presented in the "Difference" column. Standard error and *p*-values are shown for each comparison. *p*-values are adjusted for multiple comparisons calculated using the R package "*emmeans*".

3.1.3 Correlation and regression analysis between water-soluble carbohydrate and leaf area

Correlation analysis was used to measure the strength of relationship between SSS and leaf area for each pool and for the combined pool dataset (Supplementary Figure 4). Weak to moderate positive linear relationships between SSS and leaf area were observed in all pools (Supplementary Figure 4), significant at $p < 0.05$ except for FNZSL ($p = 0.73$). Pearson's coefficients of determination for WNZLL ($r^2 = 0.13$), WNZSL (0.33), WUSLL (0.26), FNZLL (0.09), FNZSL (-0.015) and the combined dataset (0.14), indicated between 1.5 – 33% of the observed SSS phenotypic variation was accounted for by leaf area in each of the pools. These low r^2 values demonstrate that the basic linear model (SSS the dependent variable and leaf area the independent variable) provided a poor to average fit to the data.

The data were then split into populations for each pool and regression analysis was used to test if leaf area was significantly predictive of SSS for each pool at the population level. Linear models were constructed with SSS as the dependent variable and the interaction between leaf area and population was used as the independent variable. Including populations in the linear model increased the adjusted r^2 values up to 94 – 97% (Supplementary Table 3). This indicated that splitting the data into populations for each pool and analysing separately provided a better fit to the data than combining all population data points within each pool. The majority of slope coefficients for each pool (Supplementary Figure 5 and Supplementary Table 3) were gradual and not significant ($p > 0.05$), with the exception of WUSLL-Parent (slope = 0.37, $p = 0.004$), FNZLL-Low-End (slope = 0.44, $p = 0.037$) and FNZLL-High-End (slope = 0.21, $p = 0.04$). Because all populations, except WUSLL-Parent, had non-significant p-values ($p > 0.05$) for both the intercept and slope, we were unable to reject the null hypothesis, allowing the conclusion that there was no relationship between SSS and leaf area for all but one population (WUSLL-Parent).

3.2 Genotyping

3.2.1 DNA isolation, genotyping-by-sequencing library evaluation and single nucleotide polymorphism filtering

High molecular weight (> 15 Kbp) genomic DNA and free from RNA contamination, was isolated successfully from 1,536 plants. For GBS library construction, 47 individuals per population with DNA concentration > 10 ng μL^{-1} were selected (n = 1,175 total). Bioanalyzer evaluation of the 13 pooled GBS libraries, prior to sequencing, showed that small adapter dimers present in the pre-size selection libraries (88 bp) were removed successfully post-size selection and that library fragment sizes were limited to the targeted 193 – 313 bp range.

A total of 191,484 SNPs were called initially across all samples. After filtering for depth, multiallelic loci, missing and minor allele frequency, a total of 14,743 SNPs were retained for 1,113 samples. A total of 109 samples were removed across all populations, most of which were population WNZSL-Parent, due to a high proportion of missing data (> 80%). A further 15 samples from amongst other populations were removed due to high missing data (> 80%). Positive control samples, a single genotype repeated in all 13 GBS libraries, were at first retained to check for consistency across GBS libraries using a principal component analysis (PCA). Duplicated GBS data from 47 individuals were also removed as they were derived from duplicated technical replicates included for quality control and were not required for subsequent analyses.

3.2.2 Single nucleotide polymorphism distribution and density

The number of SNPs found on each pseudomolecule and the SNP density across the white clover reference genome was investigated using the 14,743 SNPs from 1,113 samples identified above. Pseudomolecules were assigned to their relevant subgenomes, where pseudomolecules 1 to 8 belong to the *T. occidentale*-derived (Tr_{To}) subgenome, and pseudomolecules 9 to 16 belong to the *T. pallescens*-derived (Tr_{Tp}) subgenome (Griffiths et al., 2019). The number of SNPs per pseudomolecule was strongly correlated to pseudomolecule length ($r^2 = 0.96$, Supplementary Figure 6), with higher numbers of SNPs on the longer pseudomolecules, demonstrating that the SNPs are evenly distributed across the genome. A mean of 16.1 (± 1.6 standard deviation) SNPs per Mbp was found across all pseudomolecules. The lowest SNP density was found on pseudomolecules 14, 10 and 6 with 12.5, 13.9 and 14.5 SNPs per Mbp, respectively. The highest densities were on pseudomolecules 5, 3 and 4 with 19.1, 18.0 and 17.7 SNPs per Mbp, respectively.

3.3 Population structure

3.3.1 Discriminant analysis of principal components

Discriminant analysis of principal components (DAPC) was conducted to determine the number of clusters described by the data and to validate the pre-defined genetic clusters (i.e., populations within pools). The lowest Bayesian information criterion (BIC) value from the "*find.clusters()*" function corresponded to $K = 11$ (Supplementary Figure 7) which was therefore selected as the number of clusters described by the data. The assignment of individuals to the 11 clusters was compared with the *a priori* population grouping (Figure 3). In the following, the names of the 11 DAPC clusters are italicised and the 24 a *priori* population names are non-italicised. Individuals in both high WSC populations within a pool tended to group together in a single cluster, e.g., WNZLL-

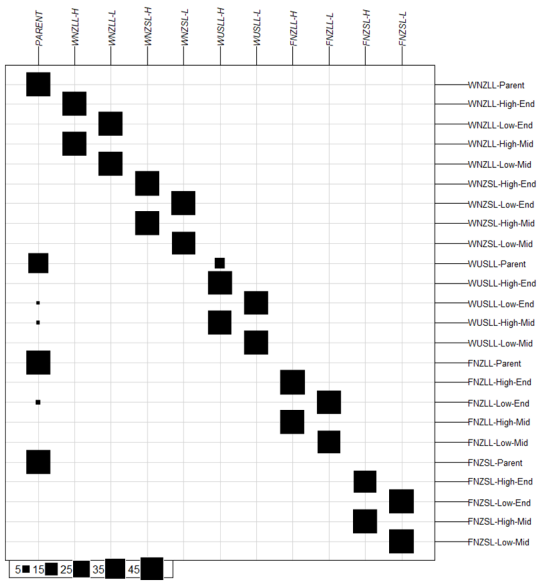


FIGURE 3
Group assignment based on *K*-means clustering prior to discriminant analysis of principal components for 24 populations. Original populations are positioned horizontally, and *K*-means determined clusters are positioned vertically. Size of black boxes represent the number of individuals assigned to the *K*-means determined cluster from the original population, with the scale presented in the bottom left-hand corner. High, high water-soluble carbohydrate (WSC); Low, low WSC; Parent, Parent generation; E, End generation; and M, Middle generation.

High-Mid and WNZLL-High-End comprised the *WNZLL-H* cluster, as did individuals from low WSC *a priori* populations e.g., WNZLL-Low-Mid and WNZLL-Low-End comprised the *WNZLL-L* cluster. Individuals from all the Parent populations grouped in a single cluster (*PARENT*), except for nine individuals from the WUSLL-Parent population that grouped with *WUSLL-H*. One to two samples from the FNZLL-Low-End, WUSLL-High-Mid and WUSLL-Low-End also grouped with the *PARENT* cluster.

The "*xvalDapc()*" function determined the optimal number of PCs to retain was 43 (root mean square error = 0.0033). However, both the root mean square error and mean successful assignment plateaued at 6 PCs (root mean square error = 0.0107 and mean successful assignment = 0.9928) with very little change thereafter (Supplementary Figure 8), therefore DAPC was run with 6 PCs retained. A scatter plot showing the 11 clusters inferred by *K*-means and the two axes representing the first two discriminant functions (DFs) of the DAPC analysis (Figure 4). The first DF showed a general separation of high WSC and low WSC populations with *High* clusters centred to the right of the plot, *Low* clusters centred to the left, and the *PARENT* plants clustering in the middle of the plot. The *WNZLL-H* and *FNZSL-L* clusters were clearly isolated from the bulk of the clusters. With respect to their counterpart populations (*WNZLL-L* and *FNZSL-H*), separation occurred on both the first and second DF. For *WNZSL* and *WUSLL*, Low and High population clusters showed very little separation on the first two DF and in fact clear separation was not observed on any DF (data not presented). The *FNZLL-L* and *FNZLL-H* clusters only showed clear separation on the fifth DF.

3.3.2 Pairwise F_{ST}

Pairwise fixation index (F_{ST}) was calculated both for the *a priori* population grouping (K = 24) and for the grouping identified above (K = 11). For biallelic marker systems Wright (1978) suggests that F_{ST} values from 0 – 0.05 indicate little differentiation, 0.05 – 0.15 moderate differentiation, 0.15 – 0.25 great differentiation, and values above 0.25 indicate high differentiation (Balloux and Lugon-Moulin, 2002; Hartl and Clark, 2007). For the *a priori* population grouping (K = 24), the WUSLL-Parent population showed moderate genetic differentiation from the NZ Parent populations (0.06 – 0.08), while F_{ST} values were low among the three NZ Parent populations (0.03 – 0.04). Pairwise comparisons of the high WSC populations (Mid and End) within each pool and of the low WSC populations (Mid and End) within each pool showed low to moderate differentiation, with F_{ST} values ranging from 0.03 – 0.09. Divergent low and high WSC pairs within a pool

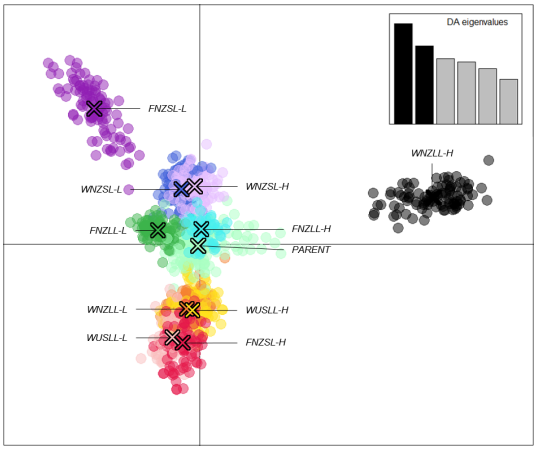


FIGURE 4
Discriminant analysis of principal components (DAPC) scatter plot of 1,113 individuals using 14,743 SNPs based on 11 assigned genetic clusters. Six principal components (Supplementary Figure 8) and six discriminant functions (DFs) were retained for analyses to describe the relationship between the genetic clusters. The scatter plot shows the first two DF from the DAPC analysis (X and Y axis explaining 24.6% and 19.3% of genetic variance, respectively) with the scree plot of eigenvalues of the linear discriminant analysis (LDA) shown in the inset. Populations are labelled and colour coded at K = 11 as determined from the *K*-means clustering algorithm. Each dot represents a single individual and the centre of each cluster, as determined by a minimum spanning tree based on the squared distances between populations, is indicated by a cross.

(e.g., WNZLL-Low-Mid compared with WNZLL-High-Mid) showed greater differentiation as indicated by an F_{ST} range of 0.12 – 0.22 (Supplementary Table 4). Pairwise F_{ST} values based on the 11 *K*-means determined clusters showed moderate genetic differentiation between *High* and *Low* clusters within each pool, ranging between 0.13 – 0.19. Very low genetic differentiation was observed between the *PARENT* cluster and all the other clusters (0.07 – 0.10) (Supplementary Table 5).

3.3.3 Analysis of molecular variance

A total of 342 loci with missing values less than 5% were used for the analysis of molecular variance (AMOVA). The AMOVA for $K = 24$ revealed that most genetic variation was partitioned within populations (80.7%), and the remainder partitioned among populations (19.3%). A hierarchical AMOVA for the 11 genetic clusters determined by *K*-means revealed that 15.4% of the variance was distributed among clusters, and only 5.4% was distributed among populations within clusters. Approximately similar variance was found within populations or clusters ($K = 24$: 80.7%, $p < 0.001$; $K = 11$: 79.2%, $p < 0.001$), indicating that genetic variation is mainly distributed within populations (Supplementary Table 6). Each of the DAPC, pairwise F_{ST} and AMOVA results indicate that, genetically, the two high WSC populations (High-Mid and High-End) for each pool can be grouped together and the two low WSC populations (Low-Mid and Low-End) for each pool can be grouped together. Population groupings for subsequent outlier locus detection were based on this $K = 10$ grouping (Parent populations were excluded).

3.4 Outlier loci detection

PCAdapt was used to identify SNPs corresponding to PCs that differentiate low and high WSC populations, within each pool. Cattell's scree test (Supplementary Figure 2) and interpretation of score plots (Supplementary Figure 9), determined that the K_{PC} value, the number of PCs to investigate, in all four pools was $K_{PC} = 1$. The first PC captures the distinction between high and low WSC populations in all pools (Figure 5). Therefore, to identify SNPs related to WSC, we focused on the SNPs associated with PC1 only. The number of SNPs used for outlier detection in PCAdapt ranged from 10,976 to 11,479 per pool, with a mean of 11,133. The genome-wide significance thresholds were determined for each pool at Bonferroni false discovery thresholds of $\alpha = 0.01$ and $\alpha = 0.05$, using their respective total number of SNPs (Supplementary Table 7). The average *p*-value threshold used for outlier detection at $\alpha = 0.05$ was 4.49e-06, which is 5.34 on the log scale. Any SNPs associated with PC1 with $-\log_{10}(p\text{-values})$ larger than 5.34 were retained from each pool and identified as putative outliers. To reduce the risk of detecting SNPs associated with population structure, SNPs identified as outliers were subjected to another criterion: they had to be present as outliers in two or more pools. Of 643 total outlier SNPs detected using PCAdapt, 36 were found in common amongst two or more pools based on PC1. A total of 329 outliers were detected using BayeScan at $\alpha = 0.05$, with 27 common to two or more pools (Supplementary Table 8). All outliers identified by BayeScan exhibited positive alpha values, indicating that only SNPs putatively under directional selection were detected. The KGD-F_{ST} method detected the largest number of outliers of the three methods, with 1,188 in total and 229 in common amongst two or more pools (Supplementary Table 8). The strongest candidates for selection were 33 SNPs found using two or more of the outlier detection methods (Supplementary Figure 10). The two F_{ST} based methods (BayeScan and KGD-F_{ST}) had the most SNPs in common ($n = 22$) whereas PCAdapt and BayeScan only had two SNPs in common and PCAdapt and KGD-F_{ST} had 13 SNPs in common.

Of the 33 candidate SNPs, five were found in exons, 15 were in introns and the remainder were either intergenic or in promoter regions. One of the exon-located SNPs exhibited a synonymous mutation, while the other four had non-synonymous mutations leading to a change in amino acid (Supplementary Table 9). Seven of the SNP-associated genes had unknown functions, while the remaining 23 had putative functions (Supplementary Table 9). One SNP, 16_32428574, was identified by BayeScan and KGD-F_{ST} in the WNZLL and FNZSL pools and found in the intron of *ERD6-like 4*, a gene coding for a sugar transporter located on the vacuole membrane.

To assess the potential that SNP genotypes changed due to random genetic drift, generational changes in genotype and allele frequencies of the 33 candidates were investigated. The vast majority of the 33 SNPs identified as outliers demonstrated a complete sweep where fixation of the reference allele occurred in the high WSC populations and the alternate allele became fixed in the low WSC populations, within the first two to three generations (Supplementary Table 10). This was observed for all 33 SNPs in two or more pools. There were instances where fixation was not achieved but genotype frequencies showed an apparent directional shift across generations. For example, in both the WUSLL and FNZLL pools at SNP 2_6673787 (Supplementary Table 10) allele frequencies were fixed for the reference allele in the low WSC populations but there was a transition from Low-End and Low-Mid to the Parent populations and then to the High-Mid and the High-End populations whereby the alternate allele increased in frequency over successive generations.

3.5 Genome-wide association study

A genome-wide association study (GWAS) was carried out using 24 white clover populations that had both genotype and WSC

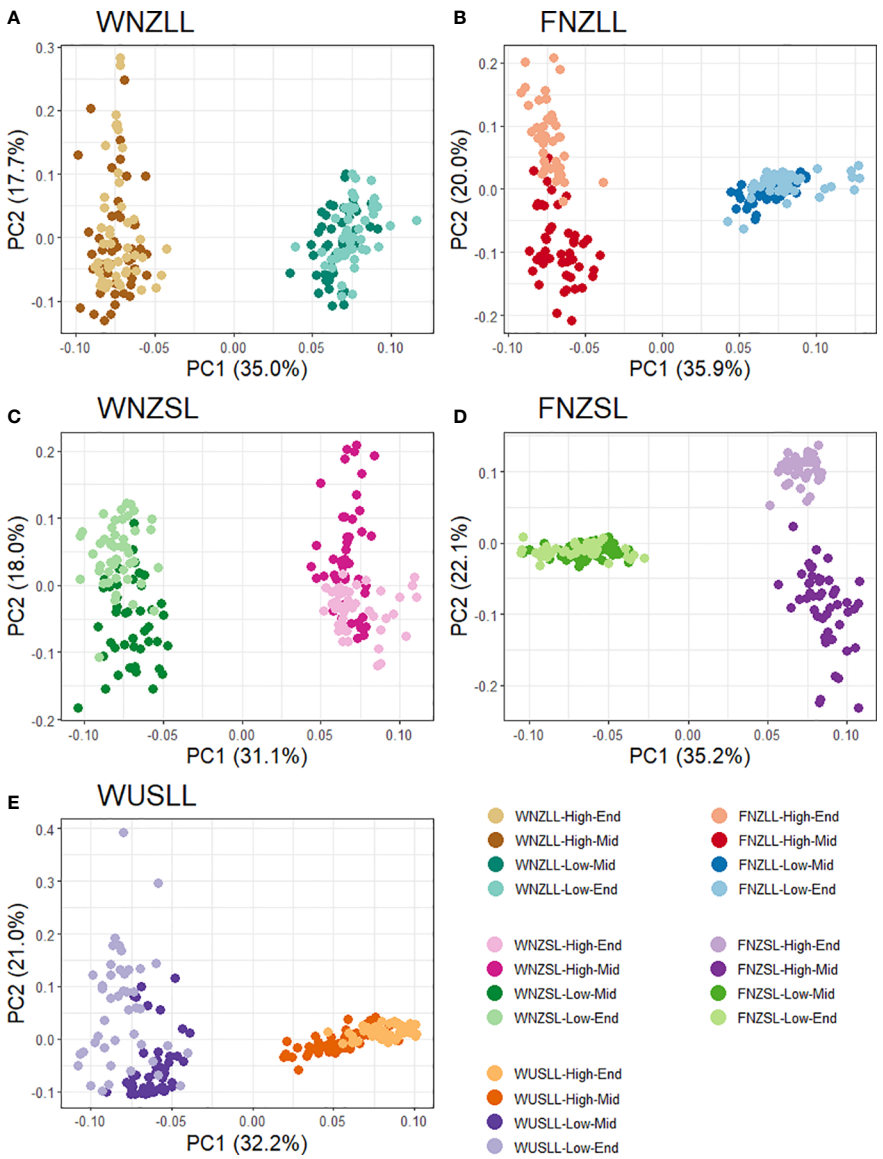


FIGURE 5
Score plots from PCAdapt analysis using the first two principal components (PC) for all five pools. Each dot represents an individual and the colour corresponds to individuals from the same population. Each pool has four populations as the Parent populations were excluded from the analysis. A total of 188 individuals were used from the WNZLL pool **(A)**, 186 from WNZSL **(B)**, 195 from WUSLL **(C)**, 182 from FNZLL **(D)** and 184 from FNZSL **(E)**, for a combined total of 935. Population information is displayed in the key in the bottom right corner.

phenotype data. No SNPs were found to be significantly associated with SSS after correction for multiple testing (Figure 6 and Supplementary Figure 11). However, ten SNPs on pseudomolecules 1, 3, 4, 6, 8, 9 and 11 were ranked highly for the SSS trait, and close to the false discovery threshold with -$\log_{10}$(*p*-values) > 3. By the same criterion two additional SNPs on pseudomolecules 1 and 5 were ranked highly for the WSC-NIRS trait but not SSS (Table 2). On the white clover reference genome, one of the SNP markers on pseudomolecule 1 located to the coding region of a *VPS35B* (Vacuolar protein sorting-associated protein 35B) gene, the second located to the intron of a *CBP* (chlorophyll a-b binding protein) gene, and the third located to the coding region of a La-related protein 7-like gene. The two markers on pseudomolecule 3 located to the intron of a transcription regulating protein (PKS-NRPS hybrid synthetase CHGG_01239-like), and the marker on pseudomolecule 4 was located 305bp

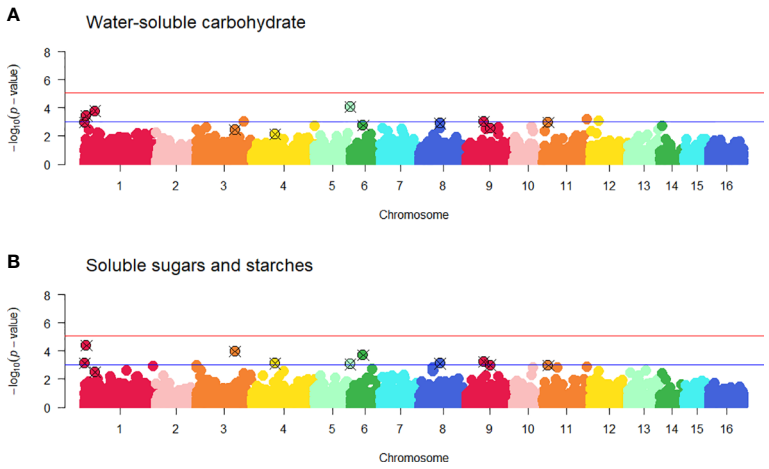


FIGURE 6
Manhattan plots from the genome-wide association study (GWAS) of water-soluble carbohydrate, WSC-NIRS **(A)** and soluble sugars and starch, SSS **(B)** using 5,757 SNP markers and 605 individuals. -$\log_{10}$(p-values) are plotted against physical map position of SNPs with subgenomes of corresponding chromosomes (i.e., pseudomolecules) similarly coloured (Tr_{To} 1 – 8 and Tr_{Tp} 9 – 16). Significant loci lie above the false discovery rate thresholds as denoted by the red (α = 0.01) and blue (-$\log_{10}$(p-value) > 3) solid lines. Twelve SNPs with -$\log_{10}$(p-values) > 3 identified for the WSC-NIRS and SSS traits are highlighted. Quantile-Quantile plots for each trait are presented in Supplementary Figure 11.

upstream of the start codon of a gene with unknown function. The SNP on pseudomolecule 5 was located 855 bp before a *glgC* (glucose-1-phosphate adenylyltransferase) gene, the SNP on pseudomolecule 6 was intergenic, the SNP on pseudomolecule 8 located to the intron of *COG8* (conserved oligomeric Golgi complex subunit 8), the SNPs on pseudomolecule 9 were located in the coding region of *UPL6* (E3 ubiquitin-protein ligase) gene and on the intron of a mixed-amyrin synthase gene. The SNP on pseudomolecule 11 located to the coding region of a clustered mitochondria protein gene. Two SNPs located in coding regions conferred non-synonymous mutations, as both altered the first base of a codon and subsequently the encoded amino acid. The SNP located in *VPS35B* caused an isoleucine to valine change, while the SNP in *UPL6* caused a glutamine to glutamic acid change. The SNP in the coding region for the clustered mitochondrial protein gene and the SNP for the La-related protein 7-like gene both conferred synonymous mutations. The SNP on pseudomolecule 5 was located near a gene (*glgC*) that may be considered a prime candidate for a role in WSC accumulation, based on its inferred function. Significant SNPs for other phenotypic traits are found in Supplementary Table 11 and visualised in Supplementary Figure 12.

4 Discussion

This study substantially increases our knowledge of genomic regions and candidate genes underpinning WSC accumulation in white clover leaves, by utilising GBS and phenotypic data from a population resource in which divergent recurrent selection for foliar WSC was undertaken over multiple generations.

4.1 The association between selection for WSC and change in leaf size

Phenotypic analysis of foliar WSC in five breeding pools confirmed breeding progress, showing that populations with divergent WSC phenotypes had been generated in each of the breeding pools. For the purpose of characterising the genetic control of foliar WSC *per se*, it is crucial to account for any effects of leaf size. Woodfield et al. (2001) reported that large-leaved white clover types tended to have higher levels of WSC than medium- and small-leaved cultivars, highlighting a potential positive correlation between the two traits and suggesting that gains in foliar WSC could be achieved by breeding for increased leaf size. However, an inverse relationship exists between leaf size and stolon density, a trait that influences vegetative persistence (Charlton and Stewart, 1999) meaning that breeding for larger leaves to increase WSC content could negatively impact sward persistence (Jahufer et al., 1999; Woodfield and Clark, 2009). In the current study, only weak correlations between WSC and leaf size were observed, in some pools, and once population was accounted for in the linear models there was little relationship between the two variables. The association between leaf area and WSC has not been researched extensively in legumes, and the outcomes are not

TABLE 2 Genome position and gene annotation for SNPs with large -$\log_{10}$(*p*-values) identified from a genome-wide association study (GWAS).

Pseudomolecule	SNP position (bp)	-$\log_{10}$ (*p*-value) for WSC-NIRS	-$\log_{10}$ (*p*-value) for SSS	Genomic region	Gene model identifier and Gene annotation (Ann)	Potential function of gene and codon change
1 (Tr_{To} – 1)	102,715	2.93	3.14	Intron	chr1.jg10.t1Ann: Chlorophyll a-b binding protein (*Trifolium pratense*)	Photosynthesis, light harvesting in photosystem I.
1 (Tr_{To} – 1)	2,338,028	3.51	4.52	Exon	chr1.jg302.t1Ann: VPS35B Vacuolar protein sorting-associated protein 35B (*Arabidopsis thaliana*)	Protein storage and vacuole biogenesis. Retarding the senescence of leaves. ATC to GTC changes isoleucine to valine.
1 (Tr_{To} – 1)	13,746,102	3.99	2.70	Exon	chr1.jg1888.t1Ann: La-related protein 7-like, partial (*Trifolium pratense*)	RNA processing GTG to GTC no change
3 (Tr_{To} – 3)	51,874,121 51,874,123	2.39	3.97	Intron	chr3.jg7956.t1Ann: PKS-NRPS hybrid synthetase CHGG_01239-like (*Cicer arietinum*)	Positive regulation of transcription from RNA polymerase II promoter in response to iron ion starvation.
4 (Tr_{To} – 4)	30,841,647	2.38	3.44	Promoter	305 bp from start codon of chr4.jg4214.t1Ann: Protein with unknown function/hypothetical protein MTR_6g072130 *(Medicago truncatula*)	
5 (Tr_{To} – 5)	47,903,593	3.83	2.85	Promoter	855 bp upstream from start codon of chr5.jg7106.t1Ann: *glgC* Glucose-1-phosphate adenylyltransferase small subunit 1, chloroplastic (*Vicia faba*)	Starch biosynthesis and glycan biosynthesis.
6 (Tr_{To} – 6)	16,865,077	3.01	3.94	Intergenic		
8 (Tr_{To} – 8)	27,636,527	2.79	3.01	Intron	chr8.jg3785.t1Ann: *COG8* conserved oligomeric Golgi complex subunit 8 (*Medicago truncatula*)	Intra-Golgi vesicle-mediated transport and protein transport.
9 (Tr_{Tp} – 1)	23,070,656	2.99	3.19	Exon	chr9.jg3440.t1Ann: UPL6 E3 ubiquitin-protein ligase UPL6 (*A. thaliana*)	Protein post-translational modifications.Response to water deficit and cold stress. CAA to GAA changes glutamine to glutamic acid.
9 (Tr_{Tp} – 1)	31,736,793	2.65	3.18	Intron	chr9.jg4697.t1Ann: Mixed-amyrin synthase (*Pisum sativum)*	Pentacyclic triterpenoid biosynthetic process, alpha- and beta-amyrin synthase activity.
11 (Tr_{Tp} – 3)	7,689,073	3.29	3.36	Exon	chr11.jg1201.t1Ann: Clustered mitochondria protein (*Cicer arietinum*)	Translational initiation ACT to ACC no change

A total of 605 individuals were used for the GWAS with a mean of 25 individuals per population ($n = 24$). A total of 122 individuals were used from the WNZLL pool, 83 from WNZSL, 136 from WUSLL, 127 from FNZLL and 137 from FNZSL.
bp, base pairs; WSC-NIRS, water-soluble carbohydrate by NIRS; SSS, Soluble sugars and starch by NIRS; Tr_{To}, white clover Trifolium occidentale-derived subgenome; Tr_{Tp}, white clover T. pallescens-derived subgenome.

consistent. For example, in lentil, a weak negative relationship between WSC and leaf area was identified (Tahir et al., 2019), whereas either positive (Malinowski et al., 1998; Woodfield et al., 2001) or no association (Ruckle et al., 2018) has been identified in white clover. Similarly, Ruckle et al. (2017) reported little to no correlation between biomass and starch concentration in red clover. The current study suggests that white clover plants with small leaves can accumulate WSC to levels comparable with large-leaved plants, as demonstrated in Figure 2, and that increased WSC content can be achieved without changing leaf size classes. We conclude that, although there was variation in leaf size in some of the breeding pools, changes in leaf WSC did not result from indirect selection for leaf size.

4.2 Population genetic structure

Assessment of population structure, which reflects relatedness among samples, is required in outlier detection and GWAS analyses as it can be a confounding factor - population structure can result in strong signals being assigned to SNPs that are not associated with the trait of interest (Sul

et al., 2018), producing false positive associations. To investigate population structure and partitioning of genetic variation for the outlier detection methodologies, three analyses were used.

K-means clustering determined that the SNP dataset from the 24 populations described 11 genetic clusters in which the two high WSC populations (High-Mid and High-End) within each pool coalesced to a single cluster, as did the two low WSC populations (Low-Mid and Low-End). By contrast, the Parent populations for all pools formed a single cluster. AMOVA on this reduced genetic cluster set showed greater variation among clusters (15.4%) than among populations within clusters (5.4%), indicating the populations within clusters were very similar. This is supported by F_{ST} values (F_{ST} 0.03 – 0.09) representing low genetic differentiation (Wright, 1978) between populations within clusters and higher values (F_{ST} 0.12 – 0.22) between Low WSC and High WSC clusters, indicating selection was successful in driving genetic separation between the high and low WSC populations. Selection within the first 2 – 3 generations (Parent to Mid), therefore, produced the greatest genetic change, with less change occurring in the next 2 – 3 generations (Mid to End). This pattern is reflected in both WSC phenotypes (Figure 2) and genotype frequencies of outlier SNPs. A mean difference of 36 g SSS kg^{-1} DM occurred between Parent and High-Mid populations, compared with 16 g SSS kg^{-1} DM between High-Mid and High-End populations and, similarly a -17.24 g SSS kg^{-1} DM differential occurred between Parent and Low-Mid populations compared with -8.34 g SSS kg^{-1} DM between Low-Mid and Low-End populations (Table 1). From a heterozygous state in the Parent population most outlier SNPs exhibited rapid fixation for one allele in the Mid populations, and no further changes were observed between the Mid and End populations (Supplementary Table 10). Thus, selection for high WSC white clover individuals can be achieved in a short time frame of 2 – 3 generations and for the purposes of identifying SNPs under selection for divergent foliar WSC, the population grouping that showed the most genetic differentiation (K = 11) should be used. One SNP associated with movement of sugars showed this pattern (Supplementary Table 10) and is discussed in greater detail below.

DAPC placed all Parent populations into a single cluster, suggesting a lack of population structure at the parental source material level. However, pairwise F_{ST} analysis showed material from the Widdup US large leaf (WUSLL) pool was slightly more genetically distinct from the New Zealand (NZ) cultivars (F_{ST} 0.06 – 0.08), than the NZ Parent populations were among themselves (F_{ST} 0.03 – 0.04) (Supplementary Table 4). It has been demonstrated that white clover has extremely large effective population sizes worldwide and exhibits negligible population structure on continental and global scales (George et al., 2006; Olsen et al., 2007; Kooyers and Olsen, 2012; Kooyers and Olsen, 2013), with pairwise F_{ST} values < 0.03 in previous studies (Wright et al., 2017; Inostroza et al., 2018). It is, therefore, not surprising that while F_{ST} values were slightly higher between different countries of origin, these values were still very low. Furthermore, cultivars used in the WUSLL-Parent population included a mixture of US (e.g., 'Tillman' and 'SRVR'; (Widdup et al., 2015)) and NZ material (e.g., 'Grasslands Huia' and 'Ranger'; (Williams, 1983; Caradus et al., 1995; Widdup et al., 2010)). As the pedigrees of the Parent populations had cultivars in common (Supplementary Table 12), it may be expected that the Parent populations should not show a great level of genetic differentiation. All three methods suggest minimal population structure was observed in the 24 white clover populations. However, divergent selection has created a structure that differentiates high and low WSC populations and within the divergent lines there was low genetic variation.

Analysis of molecular variance (AMOVA) revealed that the genetic variation within each of the 24 white clover populations accounted for a mean 81% of the total variation, whereas 19% of the variation occurred among populations (Supplementary Table 6). These results align with previous AMOVA studies where 19 – 24% of genetic variation partitioned among white clover populations (Gustine and Huff, 1999; Khanlou et al., 2011; Collins et al., 2012). The high levels of intra-population diversity, as observed in the current study, may be attributed not only to white clover's obligate outcrossing (Annicchiarico and Piano, 1995), which leads to genetically diverse populations with less genetic differentiation among populations (Hamrick and Godt, 1996; Nybom, 2004), but also other factors including a very recent human-associated range expansion (Zeven, 1991).

4.3 Outlier detection methodologies

After assessing population structure, three genome scan methods identified SNPs associated significantly with WSC levels that may also be linked to genes influencing this trait in white clover. The number of outlier SNPs differentiating high and low WSC populations varied among the methods, with few significant SNPs in common (Supplementary Figure 10). Unsurprisingly, the strongest overlap occurred between the two F_{ST}-derived methods: BayeScan and KGD-F_{ST}. However, BayeScan F_{ST} values were higher than those estimated by KGD-F_{ST}, for all pairwise comparisons. For example, in the WNZLL pool, the minimum F_{ST} value determined for a SNP locus by BayeScan was 0.21 with a maximum of 0.69 and mean of 0.23. In the equivalent KGD-F_{ST} analysis the minimum F_{ST} value was 0.0, maximum of 0.99, and the mean was 0.04. Both sets of F_{ST} values fitted χ^2 distributions (data not presented), but the mean BayeScan F_{ST} values were higher. BayeScan is relatively robust against confounding demographic processes, but strong selection, hierarchical structure, population bottlenecks and recent migration can impact this method and artificially inflate F_{ST} (Hermisson, 2009; Narum and Hess, 2011; Lotterhos and Whitlock, 2014; Lotterhos and Whitlock, 2015). Inflation of F_{ST} values due to population structure and a similar evolutionary

history has been documented (Excoffier et al., 2009; Eckert et al., 2010b). Given the contrasting outcomes from BayeScan and KGD-F_{ST}, it is likely that the inflated minimum F_{ST} values were due to analysing populations that are closely related. It is also not unusual for minimum F_{ST} values to sit around 0.2 when strong selection is occurring (Narum and Hess, 2011). For example, Reinert et al. (2019) presented F_{ST} values from a BayeScan analysis in barley that never dropped below 0.1. One of the benefits to BayeScan is that it runs two models for each locus: a neutral and a selection model. The posterior probability for both models is calculated, and the posterior odds are used as evidence for one model or the other. Log_{10}(posterior odds) ≤ 2 indicates decisive evidence for selection and corresponds to large positive alpha values. Therefore, not only do the SNPs need to have high F_{ST} values, but also large alpha values to be indicative of adaptive selection. Although the inflated F_{ST} values technically indicate all the SNPs are under selection, as F_{ST} values greater than 0.25 indicate high genetic differentiation (Wright, 1978), the type of selection they are under can be determined from the alpha value. BayeScan calculates *q*-values for each locus, which is a test statistic directly related to the false discovery rate (FDR) and should be used to make decisions (Foll and Gaggiotti, 2008). Therefore, the F_{ST} values calculated by BayeScan in this study may be disregarded and focus should instead be placed on the alpha values of SNPs above the FDR threshold as indicated by the *q*-value, the approach implemented in the current study.

Most SNPs identified as outliers indicated a complete sweep, that is, fixation of one allele occurred in the high WSC populations and the alternate allele was fixed in the low WSC populations. This was often achieved within few generations as the Mid populations often showed fixation or near fixation for one allele. Because these SNPs exhibit clear changes in allele frequencies, multiple detection methods were able to detect these as outliers. BayeScan, however, appeared unable to detect outliers due to subtle changes in allele frequency such as occurs for an incomplete sweep. This was demonstrated in multiple pools, with SNPs 2_6673787, 4_9733285 and 12_3437942, as examples. These SNPs each showed gradual increases in reference allele homozygotes in the low WSC populations and alternate allele homozygotes in the high WSC populations, moving directionally from Mid to End populations. KGD-F_{ST} and PCAdapt detected these SNPs as outliers but BayeScan did not. In support of this observation, Narum & Hess (2011) showed that, under simulated strong selection, BayeScan could correctly identify all markers under selection. However, under weak selection, only two of five markers were correctly identified. In the current study, we aimed to maximise detection of SNPs under strong selection (e.g., complete sweep), hence we compared populations that exhibited the largest differences in allele frequencies. However, to detect both types of selection (strong and weak), multiple detection methodologies need to be utilised.

4.4 An outlier SNP associated with a candidate gene for WSC accumulation

Of the 33 SNPs common to more than one analysis method, one SNP was identified as located within a gene of biological significance, indicating a potential functional association. A white clover homologue of early responsive to dehydration (*ERD*) monosaccharide vacuole transporter, *ERD6-like 4* (At1g19450), was physically associated with SNP 16_32428574, detected in both the WNZLL and FNZSL pools. *ERD6*-like transporters are involved in energy-independent sugar efflux from the vacuole (Wei et al., 2014) and are induced by dehydration and cold treatment (Kiyosue et al., 1998). Endler et al. (2006) identified *ERD6-like 4* as being highly homologous to the sugar beet (*Beta vulgaris*) hexose transporter *U43629* (Chiou and Bush, 1996). This sugar transporter has been hypothesised to catalyse facilitated diffusion of glucose across the vacuole membrane (Chiou and Bush, 1996). In plants it is uncommon for monosaccharide transporters to function as facilitators (Büttner, 2007; Taiz and Zeiger, 2010). However, in the last decade Sugars Will Eventually be Exported Transporters (*SWEET*) proteins were discovered (Chen et al., 2010), which function as energy-independent uniporters of sucrose and glucose at the plasma membrane (Chen et al., 2010; Chen et al., 2012), and of fructose across the tonoplast membrane (Chardon et al., 2013). This evidence suggests a role for *ERD6-like 4* in WSC accumulation. It is possible that *ERD6* homologs are involved in movement of sugars out of the vacuole during circumstances where carbohydrate reallocation is important (Büttner, 2007). Although there is compelling evidence to suggest that *ERD6-like 4* plays a role in WSC accumulation, possibly in the form of osmotic adjustment, further experiments need to be done to test an hypothesis that *ERD6-like 4* protein expression leads to differences in WSC content in white clover leaves.

4.5 Putative candidate genes identified from GWAS

GWAS failed to identify SNPs significantly associated with variation in leaf WSC levels, after accounting for population structure. The false discovery rate applied was controlled by Bonferroni's multiple testing correction method, which has been suggested to be too stringent (Wang et al., 2005; Hirschhorn and Daly, 2005) and may exclude real associations (Yang et al., 2010). Other studies have used a conservative *p*-value ($p < 0.001$) approach to reduce Type I error, whereby SNPs that are the highest ranking and have $-\log_{10}(p\text{-value}) \geq 3.0$ are reported (Kang et al., 2015; Sakiroglu and Brummer, 2017; Biazzi et al., 2017). Applying this approach in the current study, ten SNPs in the SSS plot exceeded the $\log_{10}(p\text{-values})$ 3.0 threshold and two

SNPs in the WSC plot exceeded the $\log_{10}(p\text{-values})$ 3.0 threshold. Seven candidate genes in close proximity to these highly ranked SNPs were investigated further for their potential association with genes.

SNP 5_47903593 is associated with *glgC*, encoding the small subunit of ADP-glucose pyrophosphorylase, which is involved in starch biosynthesis and in turn may affect glucose and sucrose concentrations (Burgess et al., 1997). Zhang et al. (2015) also found the large subunit of *glgC* (Glucose-1-phosphate adenylyltransferase) to be associated with drought resistance in *Medicago sativa*. This suggests that *glgC* may also play a role in osmotic adjustment or osmoprotection, affecting the capacity of cells to accumulate solutes and lowering water potential during periods of osmotic stress. Low molecular weight sugars are often involved in osmotic adjustment which adds to the confidence of associating variation in *glgC* with WSC accumulation in white clover.

Two SNPs (3_51874121 and 3_51874123) occur in genes associated with regulation of transcription, while SNP 11_7689073 is associated with translation initiation. Another SNP (9_23070656) was associated with the *UPL6* gene, which is involved in protein post-translational modification, mediating the addition of ubiquitin groups to target proteins for subsequent proteasomal degradation. Increased protein degradation due to environmental stress has been observed in plants as a way to mobilise nitrogen or eliminate damaged proteins (Eckert et al., 2010a). This may affect the ratio of protein to sugar concentrations in the cells as other studies have found evidence of induced protein turnover (enhancement of ubiquitin ligase) in response to water deficit and cold stress (Kiyosue et al., 1996; Abernethy and McManus, 1999; Schulze et al., 2003; Kim and Kim, 2013; Patel et al., 2015). A sixth SNP (1_2338028) occurred in the exon of *VPS35B*, which plays a role in vesicular protein sorting (Yamazaki et al., 2008) and is involved in plant growth and leaf senescence. Similarly, SNP 8_27636527 was associated with *COG8* which is involved in protein transport. Finally, SNP 1_102715 is associated with the chlorophyll a-b binding protein 7 gene, which functions as a light receptor that captures and delivers excitation energy to photosystem I. Expression of these genes will need to be investigated further to identify their role in WSC accumulation in white clover.

Inostroza et al. (2018) investigated WSC in white clover and identified four candidate genes associated with stolon WSC degradation rate, including a prolyl 4-hydroxylase alpha-like protein, a putative RING-finger E3 ubiquitin ligase, plant invertase/pectin methylesterase inhibitor and peptide/nitrate transporter. None of these genes were in common with the candidates identified in the current study, but there are important differences between the two investigations. Firstly, Inostroza et al. (2018) examined the stolon, typically a carbon sink, whereas leaves, as investigated in the current study, are carbon sources and the control of WSC in both tissues are likely to be under different mechanisms (Ballicora et al., 2004). Secondly, Inostroza et al. (2018) examined degradation of WSC whereas in the current study WSC accumulation was the focus. Finally, Inostroza et al. (2018) completed GWAS using a dataset of a greater SNP density (16.6 SNPs per Mbp) than the current study (6 SNPs per Mbp), therefore all genes important for WSC accumulation may not have been detected. The genetic control of foliar WSC accumulation requires further elucidation and should be investigated further.

5 Conclusions

The genetic control of foliar WSC accumulation in white clover was examined for the first time, by examining genetic changes in five breeding pools subject to divergent selection. Breeding for divergent foliar WSC was successful in all pools, with significant differences in WSC between low and high WSC populations achieved at the conclusion of the recurrent selection programmes, and these differences were not attributable to changes in leaf area. Outlier analyses identified GBS SNP markers that differentiate low and high WSC populations and, from these and from GWAS, two strong candidate genes were identified: *ERD6-like 4* and *glgC*. SNPs associated with a range of other candidates were also identified, which are involved in numerous aspects of plant development, membrane transport, post-translational processing, cell division and pathogen response. The clear phenotypic separation of the high and low WSC populations provides a robust platform for further investigation of foliar WSC accumulation in white clover, using transcriptomics and proteomics.

Code availability

The code can be found publicly available on GitHub. https://github.com/SofiePearson/White_Clover_WSC_Outlier_Detection_GWAS.

Data availability statement

The original contributions presented in the study are publicly available. The DNA sequence data have been deposited in the NCBI SRA database with links from BioProject accession number PRJNA915069. The phenotype data used in the study are provided in Supplementary Tables 13 and 14.

Author contributions

SP contributed to experimental design, analysed phenotypic data, conducted laboratory work, undertook SNP filtering and genetic analyses, and contributed to the writing of the manuscript. MF and AG conceived and designed the study and contributed to the writing of the manuscript. CM developed the experimental design and PPM and PM completed statistical analyses of data. AL and SH completed laboratory work supporting GBS, and RJ performed raw sequence processing and SNP calling. JF bred the white clover populations used in the experiment. JT and PL contributed to writing the final version of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

The project and SP's studentship were supported from the "Genomics for Production and Security in a Biological Economy" programme, funded by the New Zealand Ministry for Business, Innovation and Employment (C10X1306). Additional funding was provided by the Marsden Fund, Royal Society Te Apārangi "Improved modelling in evolutionary transcriptomics and proteomics will advance understanding of plant adaptation" (MAU1707).

Acknowledgments

We thank Craig Anderson for DNA extraction guidance, Rachel Tan for GBS library preparation guidance and Rachael Ashby for running KGD. A special thank you to all those who assisted with sample collecting: Craig Anderson, Emmy Bethel, Grace Ehoche, Emma Griffiths, Zulfi Jahufer, Yulia Morozova, Narsaa Na, Jana Schmidt, Chaewon Song and Prue Taylor. We are especially grateful to Prue Taylor for her help with sample processing and nursery assistance. Thank you to Andrew Faram and Steven Odering for plant care advice and nursery assistance. Finally, we thank Grasslands Innovation Ltd, for providing access to the germplasm used in this study.

Conflict of interest

Author JF is employed by PGG Wrightson Seeds Ltd.

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1095359/full#supplementary-material

References

Abernethy, G. A., and McManus, M. T. (1999). Tissue-specific changes in the pattern of ubiquitin conjugation of leaf proteins in *Festuca novae-zelandiae* in response to a water deficit. *J. Plant Physiol.* 154, 404–407. doi: 10.1016/S0176-1617(99)80188-6

Anderson, C. B., Franzmayr, B. K., Hong, S. W., Larking, A. C., Van Stijn, T. C., Tan, R., et al. (2018). Protocol: a versatile, inexpensive, high-throughput plant genomic DNA extraction method suitable for genotyping-by-sequencing. *Plant Methods* 14, 75. doi: 10.1186/s13007-018-0336-1

Annicchiarico, P., and Piano, E. (1995). Variation within and among ladino white clover ecotypes for agronomic traits. *Euphytica* 86, 135–142. doi: 10.1007/BF00022019

Antao, T., Lopes, A., Lopes, R. J., Beja-Pereira, A., and Luikart, G. (2008). LOSITAN: A workbench to detect molecular adaptation based on a Fst-outlier method. *BMC Bioinf.* 9, 323. doi: 10.1186/1471-2105-9-323

Arojju, S. K., Barth, S., Milbourne, D., Conaghan, P., Velmurugan, J., Hodkinson, T. R., et al. (2016). Markers associated with heading and aftermath heading in perennial ryegrass full-sib families. *BMC Plant Biol.* 16, 160. doi: 10.1186/s12870-016-0844-y

Ballicora, M. A., Iglesias, A. A., and Preiss, J. (2004). ADP-glucose pyrophosphorylase: A regulatory enzyme for plant starch synthesis. *Photosynthesis Res.* 79, 1–24. doi: 10.1023/B:PRES.0000011916.67519.58

Balloux, F., and Lugon-Moulin, N. (2002). The estimation of population differentiation with microsatellite markers. *Mol. Ecol.* 11, 155–165. doi: 10.1046/j.0962-1083.2001.01436.x

Biazzi, E., Nazzicari, N., Pecetti, L., Brummer, E. C., Palmonari, A., Tava, A., et al. (2017). Genome-wide association mapping and genomic selection for alfalfa (*Medicago sativa*) forage quality traits. *PLoS One* 12, e0169234. doi: 10.1371/journal.pone.0169234

Bonferroni, C. E. (1936). Teoria statistica delle classi e calcolo delle probabilità. *Pubblicazioni del R Istituto Superiore di Sci. Economiche e Commerciali di Firenze* 8, 3–62. doi: 10.1007/978-1-4419-9863-7_1213

Burgess, D., Penton, A., Dunsmuir, P., and Dooner, H. (1997). Molecular cloning and characterization of ADP-glucose pyrophosphorylase cDNA clones isolated from pea cotyledons. *Plant Mol. Biol.* 33, 431–444. doi: 10.1023/A:1005752311130

Büttner, M. (2007). The monosaccharide transporter(-like) gene family in *Arabidopsis*. *FEBS Lett.* 581, 2318–2324. doi: 10.1016/j.febslet.2007.03.016

Caradus, J. R., Mcnabb, W., Woodfield, D. R., Waghorn, G. C., and Keogh, R. (1995). "Improving quality characteristics of white clover," in *Agronomy society of New Zealand - proceedings, twenty-fifth annual conference 1995/96*. Eds. J. G. HAMPTON and K. M. POLLOCK (Lincoln Canterbury: Agronomy Soc New Zealand Inc).

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behav. Res.* 1, 245–276. doi: 10.1207/s15327906mbr0102_10

Chardon, F., Bedu, M., Calenge, F., Klemens, P. A.W., Spinner, L., Clement, G., et al. (2013). Leaf fructose content is controlled by the vacuolar transporter SWEET17 in *Arabidopsis*. *Curr. Biol.* 23, 697–702. doi: 10.1016/j.cub.2013.03.021

Charlton, J. F. L., and Stewart, A. V. (1999). Pasture species and cultivars used in New Zealand - a list. *Proc. New Z. Grassland Assoc.* 61, 147–166. doi: 10.33584/jnzg.1999.61.2328

Chen, H., and Boutros, P. C. (2011). VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in r. *BMC Bioinf.* 12, 35. doi: 10.1186/1471-2105-12-35

Chen, L.-Q., Hou, B.-H., Lalonde, S., Takanaga, H., Hartung, M. L., Qu, X.-Q., et al. (2010). Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature* 468, 527–532. doi: 10.1038/nature09606

Chen, L.-Q., Qu, X.-Q., Hou, B.-H., Sosso, D., Osorio, S., Fernie, A. R., et al. (2012). Sucrose efflux mediated by SWEET proteins as a key step for phloem transport. *Science* 335, 207. doi: 10.1126/science.1213351

Chessel, D., Dufour, A. B., and Thioulouse, J. (2004). The *ade4* package - I: One-table methods. *R News* 4, 5–10.

Chiou, T. J., and Bush, D. R. (1996). Molecular cloning, immunochemical localization to the vacuole, and expression in transgenic yeast and tobacco of a putative sugar transporter from sugar beet. *Plant Physiol.* 110, 511–520. doi: 10.1104/pp.110.2.511

Collins, R. P., Helgadóttir, Á., Frankow-Lindberg, B. E., Skøt, L., Jones, C., and Skøt, K. P. (2012). Temporal changes in population genetic diversity and structure in red and white clover grown in three contrasting environments in northern Europe. *Ann. Bot.* 110, 1341–1350. doi: 10.1093/aob/mcs058

Corson, D. C., Waghorn, G. C., Ulyatt, M. J., and Lee, J. (1999). "NIRS: Forage analysis and livestock feeding," in *Proceedings of the New Zealand Grassland Association*, (Hawkes Bay, New Zealand: New Zealand Grassland Association) 61, 127–132. doi: 10.33584/jnzg.1999.61.2340

Cosgrove, G. P., Koolaard, J., Luo, D., Burke, J. L., and Pacheco, D. (2009). "The composition of high sugar ryegrasses," in *Proceedings of the New Zealand Grassland Association*, (Waitangi, New Zealand: New Zealand Grassland Association) 71, 187–193. doi: 10.33584/jnzg.2009.71.2747

Dalmannsdóttir, S., Helgadóttir, Á., and Gudleifsson, B. E. (2001). Fatty acid and sugar content in white clover in relation to frost tolerance and ice-encasement tolerance. *Ann. Bot.* 88, 753–759. doi: 10.1006/anbo.2001.1465

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

Dodds, K. G., Mcewan, J. C., Brauning, R., Anderson, R. M., Van Stijn, T. C., Kristjánsson, T., et al. (2015). Construction of relatedness matrices using genotyping-by-sequencing data. *BMC Genomics* 16, 1047. doi: 10.1186/s12864-015-2252-3

Easton, H. S., Stewart, A. V., Lyons, T. B., Parris, M., and Charrier, S. (2009). "Soluble carbohydrate content of ryegrass cultivars," in *Proceedings of the New Zealand Grassland Association*, (Waitangi, New Zealand: New Zealand Grassland Association) 71, 161–166. doi: 10.33584/jnzg.2009.71.2745

Eckert, A. J., Bower, A. D., González-Martínez, S. C., Wegrzyn, J. L., Coop, G., and Neale, D. B. (2010a). Back to nature: ecological genomics of loblolly pine (*Pinus taeda*, pinaceae). *Mol. Ecol.* 19, 3789–3805. doi: 10.1111/j.1365-294X.2010.04698.x

Eckert, A. J., Van Heerwaarden, J., Wegrzyn, J. L., Nelson, C. D., Ross-Ibarra, J., González-Martínez, S. C., et al. (2010b). Patterns of population structure and environmental associations to aridity across the range of loblolly pine (*Pinus taeda* l., pinaceae). *Genetics* 185, 969. doi: 10.1534/genetics.110.115543

Ellison, N. W., Liston, A., Steiner, J. J., Williams, W. M., and Taylor, N. L. (2006). Molecular phylogenetics of the clover genus (*Trifolium* - leguminosae). *Mol. Phylogenet. Evol.* 39, 688–705. doi: 10.1016/j.ympev.2006.01.004

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., et al. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PloS One* 6, e19379. doi: 10.1371/journal.pone.0019379

Endelman, J. B. (2011). Ridge regression and other kernels for genomic selection with r package rrBLUP. *Plant Genome* 4, 250–255. doi: 10.3835/plantgenome2011.08.0024

Endler, A., Meyer, S., Schelbert, S., Schneider, T., Weschke, W., Peters, S. W., et al. (2006). Identification of a vacuolar sucrose transporter in barley and *Arabidopsis* mesophyll cells by a tonoplast proteomic approach. *Plant Physiol.* 141, 196–207. doi: 10.1104/pp.106.079533

Excoffier, L., Hofer, T., and Foll, M. (2009). Detecting loci under selection in a hierarchically structured population. *Heredity* 103, 285. doi: 10.1038/hdy.2009.74

Excoffier, L., Smouse, P. E., and Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131, 479–491. doi: 10.1093/genetics/131.2.479

Faville, M. J., Ganesh, S., Cao, M., Jahufer, M. Z. Z., Bilton, T. P., Easton, H. S., et al. (2018). Predictive ability of genomic selection models in a multi-population perennial ryegrass training set using genotyping-by-sequencing. *Theor. Appl. Genet.* 131, 703–720. doi: 10.1007/s00122-017-3030-1

Foll, M., and Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics* 180, 977–993. doi: 10.1534/genetics.108.092221

George, J., Dobrowolski, M. P., Jong, E. V. Z. D., Cogan, N. O. I., Smith, K. F., and Forster, J. W. (2006). Assessment of genetic diversity in cultivars of white clover (*Trifolium repens* l.) detected by SSR polymorphisms. *Genome* 49, 919–930. doi: 10.1139/g06-079

Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., et al. (2014). TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. *PLoS One* 9, e90346. doi: 10.1371/journal.pone.0090346

Goudet, J., and Jombart, T. (2015). *Hierfstat: Estimation and tests of hierarchical f-statistics* (R package version 0), 04–22. Available at: https://CRAN.R-project.org/package=hierfstat.

Griffiths, A. G., Moraga, R., Tausen, M., Gupta, V., Bilton, T. P., Campbell, M. A., et al. (2019). Breaking free: The genomics of allopolyploidy-facilitated niche expansion in white clover. *Plant Cell* 31, 1466–1487. doi: 10.1105/tpc.18.00606

Guo, X., Cericola, F., Fè, D., Pedersen, M. G., Lenk, I., Jensen, C. S., et al. (2018). Genomic prediction in tetraploid ryegrass using allele frequencies based on genotyping by sequencing. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.01165

Gustine, D. L., and Huff, D. R. (1999). Genetic variation within among white clover populations from managed permanent patures of the northeastern USA. *Crop Sci.* 39, 524–530. doi: 10.2135/cropsci1999.0011183X003900020037x

Hamrick, J. L., and Godt, M. J. W. (1996). Effects of life history traits on genetic diversity in plant species. philosophical transactions of the royal society of London. *Ser. B: Biol. Sci.* 351, 1291–1298. doi: 10.1098/rstb.1996.0112

Hartl, D. L., and Clark, A. G. (2007). *Principles of population genetics* (Sinauer: Sutherland).

Hermisson, J. (2009). Who believes in whole-genome scans for selection? *Heredity* 103, 283–284. doi: 10.1038/hdy.2009.101

Hirschhorn, J. N., and Daly, M. J. (2005). Genome-wide association studies for common diseases and complex traits. *Nat. Rev. Genet.* 6, 95–108. doi: 10.1038/nrg1521

Inostroza, L., Bhakta, M., Acuña, H., Vásquez, C., Ibáñez, J., Tapia, G., et al. (2018). Understanding the complexity of cold tolerance in white clover using temperature gradient locations and a GWAS approach. *Plant Genome* 11, 170096. doi: 10.3835/plantgenome2017.11.0096

Jahufer, M. Z. Z., Cooper, M., Bray, R. A., and Ayres, J. F. (1999). Evaluation of white clover (*Trifolium repens* l.) populations for summer moisture stress adaptation in Australia. *Aust. J. Agric. Res.* 50, 561–574. doi: 10.1071/A98141

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., Mcginnis, S., and Madden, T. L. (2008). NCBI BLAST: a better web interface. *Nucleic Acids Res.* 36, W5–W9. doi: 10.1093/nar/gkn201

Jombart, T. (2008). *adegenet*: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403–1405. doi: 10.1093/bioinformatics/btn129

Jombart, T., Devillard, S., and Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* 11, 94. doi: 10.1186/1471-2156-11-94

Kagan, I. A., Anderson, M. L., Kramer, K. J., Seman, D. H., Lawrence, L. M., and Smith, S. R. (2020). Seasonal and diurnal variation in water-soluble carbohydrate concentrations of repeatedly defoliated red and white clovers in central Kentucky. *J. Equine Veterinary Sci.* 84, 102858. doi: 10.1016/j.jevs.2019.102858

Kamvar, Z. N., Tabima, J. F., and Grünwald, N. J. (2014). *Poppr*: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2, e281. doi: 10.7717/peerj.281

Kang, Y., Sakiroglu, M., Krom, N., Stanton-Geddes, J., Wang, M., Lee, Y.-C., et al. (2015). Genome-wide association of drought-related and biomass traits with HapMap SNPs in medicago truncatula. *Plant Cell Environ.* 38, 1997–2011. doi: 10.1111/pce.12520

Kerepesi, I., and Galiba, G. (2000). Osmotic and salt stress-induced alteration in soluble carbohydrate content in wheat seedlings. *Crop Sci.* 40, 482–487. doi: 10.2135/cropsci2000.402482x

Khanlou, K. M., Vandepitte, K., Asl, L. K., and Bockstaele, E. V. (2011). Towards an optimal sampling strategy for assessing genetic variation within and among white clover (*Trifolium repens* l.) cultivars using AFLP. *Genet. Mol. Biol.* 34, 252–258. doi: 10.1590/s1415-47572011000200015

Kim, S. J., and Kim, W. T. (2013). Suppression of *Arabidopsis* RING E3 ubiquitin ligase *AtATL78* increases tolerance to cold stress and decreases tolerance to drought stress. *FEBS Lett.* 587, 2584–2590. doi: 10.1016/j.febslet.2013.06.038

Kiyosue, T., Abe, H., Yamaguchi-Shinozaki, K., and Shinozaki, K. (1998). ERD6, a cDNA clone for an early dehydration-induced gene of *Arabidopsis*, encodes a putative sugar transporter. *Biochim. Biophys. Acta (BBA) - Biomembranes* 1370, 187–191. doi: 10.1016/S0005-2736(98)00007-8

Kiyosue, T., Yoshiba, Y., Yamaguchi-Shinozaki, K., and Shinozaki, K. (1996). A nuclear gene encoding mitochondrial proline dehydrogenase, an enzyme involved in proline metabolism, is upregulated by proline but downregulated by dehydration in *Arabidopsis*. *Plant Cell* 8, 1323–1335. doi: 10.1105/tpc.8.8.1323

Knaus, B. J., and Grünwald, N. J. (2017). Vcfr: a package to manipulate and visualize variant call format data in r. *Mol. Ecol. Resour.* 17, 44–53. doi: 10.1111/1755-0998.12549

Kooyers, N. J., and Olsen, K. M. (2012). Rapid evolution of an adaptive cyanogenesis cline in introduced north American white clover (*Trifolium repens* l.). *Mol. Ecol.* 21, 2455–2468. doi: 10.1111/j.1365-294X.2012.05486.x

Kooyers, N. J., and Olsen, K. M. (2013). Searching for the bull's eye: agents and targets of selection vary among geographically disparate cyanogenesis clines in white clover (*Trifolium repens* l.). *Heredity* 111, 495–504. doi: 10.1038/hdy.2013.71

Lenth, R. V., Buerkner, P., Giné-Vázquez, I., Herve, M., Jung, M., Love, J., et al. (2020). *Emmeans: Estimated marginal means, aka least-squares means* (R package version 1), 4.2. Available at: https://CRAN.R-project.org/package=emmeans.

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

Livingston, D. P., Hincha, D. K., and Heyer, A. G. (2009). Fructan and its relationship to abiotic stress tolerance in plants. *Cell. Mol. Life Sci. CMLS* 66, 2007–2023. doi: 10.1007/s00018-009-0002-x

Lotterhos, K. E., and Whitlock, M. C. (2014). Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Mol. Ecol.* 23, 2178–2192. doi: 10.1111/mec.12725

Lotterhos, K. E., and Whitlock, M. C. (2015). The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Mol. Ecol.* 24, 1031–1046. doi: 10.1111/mec.13100

Luo, J., Sun, X. Z., Pacheco, D., Ledgard, S. F., Lindsey, S. B., Hoogendoorn, C. J., et al. (2015). Nitrous oxide emission factors for urine and dung from sheep fed either fresh forage rape (*Brassica napus* l.) or fresh perennial ryegrass (*Lolium perenne* l.). *Animal* 9, 534–543. doi: 10.1017/S1751731114002742

Luu, K., Bazin, E., and Blum, M. G. B. (2017). *pcadapt*: an R package to perform genome scans for selection based on principal component analysis. *Mol. Ecol. Resour.* 17, 67–77. doi: 10.1111/1755-0998.12592

Malinowski, D. P., Belesky, D. P., and Fedders, J. (1998). Photosynthesis of white clover (*Trifolium repens* l.) germplasms with contrasting leaf size. *Photosynthetica* 35, 419–427. doi: 10.1023/A:1006920520169

Michell, P. J. (1973). Relations between fibre and water soluble carbohydrate contents of pasture species and their digestibility and voluntary intake by sheep. *Aust. J. Exp. Agric. Anim. Husbandry* 13, 165–170. doi: 10.1071/EA9730165

Narum, S. R., and Hess, J. E. (2011). Comparison of FST outlier tests for SNP loci under selection. *Mol. Ecol. Resour.* 11, 184–194. doi: 10.1111/j.1755-0998.2011.02987.x

Nybom, H. (2004). Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Mol. Ecol.* 13, 1143–1155. doi: 10.1111/j.1365-294x.2004.02141.x

Olsen, K. M., Sutherland, B. L., and Small, L. L. (2007). Molecular evolution of the li/li chemical defence polymorphism in white clover (*Trifolium repens* l.). *Mol. Ecol.* 16, 4180–4193. doi: 10.1111/j.1365-294x.2007.03506.x

Patel, M., Milla-Lewis, S., Zhang, W., Templeton, K., Reynolds, W. C., Richardson, K., et al. (2015). Overexpression of ubiquitin-like LpHUB1 gene confers drought tolerance in perennial ryegrass. *Plant Biotechnol. J.* 13, 689–699. doi: 10.1111/pbi.12291

Poland, J. A., Brown, P. J., Sorrells, M. E., and Jannink, J.-L. (2012b). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-Sequencing approach. *PloS One* 7, e32253. doi: 10.1371/journal.pone.0032253

Poland, J., Endelman, J., Dawson, J., Rutkoski, J., Wu, S., Manes, Y., et al. (2012a). Genomic selection in wheat breeding using genotyping-by-Sequencing. *Plant Genome* 5, 103–113. doi: 10.3835/plantgenome2012.06.0006

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795

R Core Team (2019). *R: A language and environment for statistical computing. 3.6.1* (Vienna, Austria: R Foundation for Statistical Computing). Available at: https://www.R-project.org/.

Reinert, S., Osthoff, A., Léon, J., and Naz, A. A. (2019). Population genetics revealed a new locus that underwent positive selection in barley. *Int. J. Mol. Sci.* 20, 202. doi: 10.3390/ijms20010202

Royston, P. (1995). Remark AS R94: A remark on algorithm AS 181: The W-test for normality. *J. R. Stat. Society Ser. C (Applied Statistics)* 44, 547–551. doi: 10.2307/2986146

Ruckle, M. E., Bernasconi, L., Kölliker, R., Zeeman, S. C., and Studer, B. (2018). Genetic diversity of diurnal carbohydrate accumulation in white clover (*Trifolium repens* l.). *Agronomy* 8, 47. doi: 10.3390/agronomy8040047

Ruckle, M. E., Meier, M. A., Frey, L., Eicke, S., Kölliker, R., Zeeman, S. C., et al. (2017). Diurnal leaf starch content: An orphan trait in forage legumes. *Agronomy* 7, 16. doi: 10.3390/agronomy7010016

Sakiroglu, M., and Brummer, E. C. (2017). Identification of loci controlling forage yield and nutritive value in diploid alfalfa using GBS-GWAS. *Theor. Appl. Genet.* 130, 261–268. doi: 10.1007/s00122-016-2782-3

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682. doi: 10.1038/nmeth.2019

Schulze, W. X., Reinders, A., Ward, J., Lalonde, S., and Frommer, W. B. (2003). Interactions between co-expressed *Arabidopsis* sucrose transporters in the split-ubiquitin system. *BMC Biochem.* 4, 3–3. doi: 10.1186/1471-2091-4-3

Selbie, D. R., Buckthought, L. E., and Shepherd, M. A. (2015). "Chapter four - the challenge of the urine patch for managing nitrogen in grazed pasture systems," in *Advances in agronomy*. Ed. D. L. SPARKS (Academic Press) 129, 229–292. doi: 10.1016/bs.agron.2014.09.004

Sul, J. H., Martin, L. S., and Eskin, E. (2018). Population structure in genetic studies: Confounding factors and mixed models. *PloS Genet.* 14, e1007309–e1007309. doi: 10.1371/journal.pgen.1007309

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 47, D607–d613. doi: 10.1093/nar/gky1131

Tahir, F., Hassani, A., Kouadria, M., and Rezzoug, W. (2019). Study of morpho-physiological and biochemical behavior of cultivated legume (*Lens culinaris medik* ssp *culinaris*) in dry area of Algeria. *Ukrainian J. Ecol.* 9, 535–541. doi: 10.15421/2019_786

Taiz, L., and Zeiger, E. (2010). "Chapter 6: Solute transport," in *Plant physiology. 5th ed* (Sunderland, MA, USA: Sinauer Associates).

The Uniprot Consortium (2018). UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* 47, D506–D515. doi: 10.1093/nar/gky1049

Turner, S. D. (2018). *qqman*: An R package for visualizing GWAS results using Q-Q and manhattan plots. *J. Open Source Software* 3 (25), 731. doi: 10.21105/joss.00731

Ulyatt, M. J. (1997). "Can protein utilisation from pasture be improved?," in *Proceedings of the New Zealand Society of Animal Production*, (Palmerston North, New Zealand: New Zealand Society of Animal Production) 57, 4–8.

Venables, W. N., and Ripley, B. D. (2002). *Modern applied statistics with s* (New York: Springer). doi: 10.1007/978-0-387-21706-2

Wang, W. Y. S., Barratt, B. J., Clayton, D. G., and Todd, J. A. (2005). Genome-wide association studies: theoretical and practical concerns. *Nat. Rev. Genet.* 6, 109–118. doi: 10.1038/nrg1522

Wei, X., Liu, F., Chen, C., Ma, F., and Li, M. (2014). The malus domestica sugar transporter gene family: identifications based on genome and expression profiling related to the accumulation of fruit sugars. *Front. Plant Sci.* 5. doi: 10.3389/fpls.2014.00569

Weir, B. S., and Cockerham, C. C. (1984). Estimating f-statistics for the analysis of population structure. *Evolution* 38, 1358–1370. doi: 10.1111/j.1558-5646.1984.tb05657.x

Whitlock, M. C., and Lotterhos, K. E. (2015). Reliable detection of loci responsible for local adaptation: Inference of a null model through trimming the distribution of F_{ST}. *Am. Nat.* 186, S24–S36. doi: 10.1086/682949

Widdup, K. H., Ford, J. L., Barrett, B. A., and Woodfield, D. R. (2010). "Development of white clover populations with higher concentrations of water soluble carbohydrate," in *Proceedings of the New Zealand Grassland Association*, (Lincoln, New Zealand: New Zealand Grassland Association) 72, 277–282. doi: 10.33584/jnzg.2010.72.2795

Widdup, K. H., Ford, J. L., Cousins, G., Woodfield, D., Caradus, J. R., and Barrett, B. (2015). A comparison of New Zealand and overseas white clover

cultivars under grazing in New Zealand. *J. New Z. Grasslands* 77, 51–56. doi: 10.33584/jnzg.2015.77.483

Williams, W. M. (1983). "Chapter 25 white clover," in *Plant breeding in New Zealand*. Eds. G. S. WRATT and H. C. SMITH (Wellington, NZ: Butterworths/DSIR).

Woodfield, D. R., and Clark, D. A. (2009). Do forage legumes have a role in modern dairy farming systems? *Irish J. Agric. Food Res.* 48, 137–147.

Woodfield, D. R., Clifford, P. T. P., Cousins, G. R., Ford, J. L., Baird, I. J., Miller, J. E., et al. (2001). "Grasslands kopu II and crusader: new generation white clovers," in *Proceedings of the New Zealand Grassland Association*, (Waikato, New Zealand: New Zealand Grassland Association) 63, 103–108. doi: 10.33584/jnzg.2001.63.2446

Wright, S. (1978). Evolution and the genetics of populations.*Variability within and among natural populations* (Chicago: University of Chicago Press) Vol 4.

Wright, S. J., Zhou, C. D., Kuhle, A., and Olsen, K. M. (2017). Continent-wide climatic variation drives local adaptation in north American white clover. *J. Heredity* 109, 78–89. doi: 10.1093/jhered/esx060

Yamazaki, M., Shimada, T., Takahashi, H., Tamura, K., Kondo, M., Nishimura, M., et al. (2008). *Arabidopsis* VPS35, a retromer component, is required for vacuolar protein sorting and involved in plant growth and leaf senescence. *Plant Cell Physiol.* 49, 142–156. doi: 10.1093/pcp/pcn006

Yang, J., Benyamin, B., Mcevoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., et al. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* 42, 565–569. doi: 10.1038/ng.608

Zeven, A. C. (1991). Four hundred years of cultivation of Dutch white clover landraces. *Euphytica* 54, 93–99. doi: 10.1007/BF00145635

Zhang, T., Yu, L.-X., Zheng, P., Li, Y., Rivera, M., Main, D., et al. (2015). Identification of loci associated with drought resistance traits in heterozygous autotetraploid alfalfa (*Medicago sativa* l.) using genome-wide association studies with genotyping by sequencing. *PLoS One* 10, e0138931. doi: 10.1371/journal.pone.0138931

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 10 January 2023
DOI 10.3389/fpls.2022.1104755

OPEN ACCESS

EDITED BY
Linkai Huang,
Sichuan Agricultural University, China

REVIEWED BY
Dan Luo,
Sichuan Agricultural University, China
Shiyong Chen,
Southwest Minzu University, China
Tao Hu,
Lanzhou University, China

*CORRESPONDENCE
Yan Fan
34629646@qq.com
Bing Zeng
zbin78@163.com

†These authors have contributed equally to this work

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

RECEIVED 22 November 2022
ACCEPTED 20 December 2022
PUBLISHED 10 January 2023

CITATION
Qu M, Zheng Y, Bi L, Yang X, Shang P, Zhou X, Zeng B, Shen B, Li W, Fan Y and Zeng B (2023) Comparative transcriptomic analysis of the gene expression and underlying molecular mechanism of submergence stress response in orchardgrass roots. *Front. Plant Sci.* 13:1104755. doi: 10.3389/fpls.2022.1104755

Comparative transcriptomic analysis of the gene expression and underlying molecular mechanism of submergence stress response in orchardgrass roots

Minghao Qu[1†], Yuqian Zheng[1], Lei Bi[1], Xingyun Yang[1], Panpan Shang[1], Xiaoli Zhou[1], Bing Zeng[1], Bingna Shen[1], Wenwen Li[1], Yan Fan[2*] and Bing Zeng[1,3*]

[1]College of Animal Science and Technology, Southwest University, Chongqing, China, [2]Institute of Prataculture, Chongqing Academy of Animal Science, Chongqing, China, [3]Chongqing University Herbivore Engineering Research Center, Chongqing, China

Introduction: Submergence stress creates a hypoxic environment. Roots are the first plant organ to face these low-oxygen conditions, which causes damage and affects the plant growth and yield. Orchardgrass (*Dactylis glomerata* L.) is one of the most important cold-season forage grasses globally. However, their submergence stress-induced gene expression and the underlying molecular mechanisms of orchardgrass roots are still unknown.

Methods: Using the submergence-tolerant 'Dianbei' and submergence-sensitive 'Anba', the transcriptomic analysis of orchardgrass roots at different time points of submergence stress (0 h, 8 h, and 24 h) was performed.

Results: We obtained 118.82Gb clean data by RNA-Seq. As compared with the control, a total of 6663 and 9857 differentially expressed genes (DEGs) were detected in Dianbei, while 7894 and 11215 DEGs were detected in Anba at 8 h and 24 h post-submergence-stress, respectively. Gene Ontology (GO) enrichment analysis obtained 986 terms, while Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis obtained 123 pathways. Among them, the DEGs in plant hormones, mitogen-activated protein kinase (MAPK) and Ca^{2+} signal transduction were significantly differentially expressed in Dianbei, but not in Anba.

Discussion: This study was the first to molecularly elucidate the submergence stress tolerance in the roots of two orchardgrass cultivars. These findings not only enhanced our understanding of the orchardgrass submergence tolerance, but also provided a theoretical basis 36 for the cultivation of submergence-tolerant forage varieties.

KEYWORDS

***Dactylis glomerata* L., transcriptome, submergence stress, differentially expressed genes, molecular mechanisms**

Introduction

Water is essential for plants, but excessive water, waterlogging or flooding can cause stress by preventing the exchange of gases between the soil and the atmosphere, which negatively impacts plant growth and development (Striker and Colmer, 2016; Wang et al., 2017). Although oxygen accounts for ~20% in the air, it has extremely low solubility in water. The diffusion rate of oxygen in flooded soil is 10^4 times slower than in air. Furthermore, respiration of soil microorganisms and plant roots rapidly depletes the oxygen level (Kaur et al., 2021). Root is the first organ facing the decline of oxygen tension. Since they can't produce oxygen through photosynthesis, it will soon face hypoxia under submergence stress, ultimately causing more damage to the roots (Panozzo et al., 2019). Plants with submergence-tolerant depend heavily on multiple morphological and physiological changes that are regulated by different genes. RNA-seq has been widely used in plant field in recent years, especially in the research of plant resistance (Li et al., 2018). Transcriptome analysis revealed that ethylene and ABA synthesis in *Phalaris arundinacea* were inhibited by waterlogging stress (Wang et al., 2021). In deep-water rice (*Oryza sativa*) cultivar, the elevated ethylene levels promoted the expression of *SK1/SK2*, and reduced abscisic acid (ABA) biosynthesis, which further increased gibberellin (GA) content and induced cell elongation (Zhou et al., 2020). In maize (*Zea mays*), invertase and hexokinase expression were up-regulated under waterlogged stress in roots, providing more energy for self-stabilization (Arora et al., 2017). The insufficient energy resulting from the stress-induced oxygen shortage triggers reactive oxygen species (ROS) formation (Yeung et al., 2018; Anee et al., 2019). The hypoxic environment inhibits ATP production, restricts mitochondrial respiration, and alters the ratios of ATP to ADP and ATP to AMP (Sairam et al., 2008). Submergence causes plants to grow under hypoxic conditions, which limits the gas exchange capability of plants and impairs the metabolic balance and nutrient transport of plant roots, resulting in an 'energy crisis' (Herzog et al., 2016; Zhou et al., 2020). In a word, submergence can cause plant morphological damage, increasing the susceptibility of plant to diseases and insect pests, and even cause plant death (Phukan et al., 2016). RNA-seq can be used to study the molecular mechanism of plant response to submergence stress.

Grasslands are distributed worldwide except for Antarctica. The grassland covers an area of 52,544,000 km^2, which is 40.5% of the world's land area. It is the second largest land type for human habitation after agricultural land (Karunarathna et al., 2021). Orchardgrass is a cold-season, and perennial forage grass native to Eurasia and North Africa. As one of the four major forages in the world, it can be used as pasture or hay grass and has gained good ecological and economic benefits in China (Hirata et al., 2011). As an important forage grass for cultivating livestock worldwide, orchardgrass has the advantages of fast growth, high yield, good palatability, rich nutrition and strong stress resistance (Feng et al., 2018). Orchardgrass is mainly distributed in Sichuan, Chongqing, Yunnan, Guizhou, and other regions in China. Because of rich germplasm resources, orchardgrass has been widely planted and applied. It is an important species for land management and restoration in Southwest China (Peng et al., 2008). Therefore, in the world of increasing forage resources demand, submergence-tolerant is a key for orchardgrass breeding programs.

In recent years, with the rapid development of the next-generation sequencing (NGS) technology and the assembly of orchardgrass genome, a better platform for evaluating molecular and genetic information of orchardgrass is available (McCombie et al., 2019; Huang et al., 2020). *WRKY*, *GRAs*, and C2H2-type zinc finger gene family have previously been studied in orchardgrass (Xu et al., 2020; Ren et al., 2021; Shuai et al., 2022). Studies on biotic and abiotic stress reports of orchardgrass and other cold-season forage grasses had mainly focused on drought, high temperature, and rust (Huang et al., 2015; Ji et al., 2018; Yan et al., 2013, 2016; Hu et al., 2014; Sun et al., 2020). Although there are some studies on submergence stress in orchardgrass (Klaas et al., 2019; Qiao et al., 2020; Zeng et al., 2020), but it's underlying submergence stress mechanisms in roots have not been reported. In the current study, the aim of this study was to understand the gene expression and underlying molecular mechanism of different submergence stress times response in two orchardgrass cultivars roots by RNA-seq. These findings can provide a basis for further study of the submergence tolerance mechanism of orchardgrass.

Materials and methods

Plant materials and growth conditions

In this experiment, the submergence-tolerant 'Dianbei' (DB) and submergence-sensitive 'Anba' (AB) were selected as test materials. The seeds of orchardgrass were obtained from the College of Animal Science and Technology of Southwest University, China. The seeds of orchardgrass were germinated in culture dish. Orchardgrass with the same growth trend were transplanted to the flowerpot (diameter 15.0 cm, height 13.5 cm) containing vermiculite, vegetative soil, and perlite (1:3:1, v/v/v). The orchardgrass cultivars were grown in an incubator with humidity of 70-85%, temperature of 22/15°C (day/night), photoperiod 14 h/10 h (day/night), and 5000 lux light intensity (Xu et al., 2017). Subsequently, the plants were irrigated with 1/2 Hoagland nutrient solution twice a week.

At the 3-4 leaf stage, the submergence treatment was started. Well-grown orchardgrass were selected for testing. Orchardgrass were placed in a water tank (length 80 cm, width 57 cm, height 50

cm), and water was added until the orchardgrass were completely submerged. Root phenotype and root length were used as growth characteristics data. The root length was measured and photographed at 0 h, 8 h, 24 h, 48 h and 72 h after submergence stress, and each treatment was repeated three times. Samples were taken at 0 h, 8 h, and 24 h under submergence stress, frozen in liquid nitrogen immediately, and stored at −80°C for further analysis. Each treatment was repeated for three times, with a total of 18 samples.

RNA extraction and sequencing

Total RNA was extracted from roots at each time point (three biological replicates per treatment) using Trizol reagent (Invitrogen) according to the manufacture's protocol. Total amounts and integrity of RNA were assessed using the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, CA, USA). Samples with a RIN > 6.5 were used in RNA-seq library preparation. Total RNA was used as the input material for the mRNA sample preparations. The mRNA was purified from the total RNA by using poly-T oligo-attached magnetic beads. Fragmentation was carried out using divalent cations under elevated temperature in First Strand Synthesis Reaction Buffer(5X). First strand cDNA was synthesized using random hexamer primer and M-MuLV Reverse Transcriptase, then use RNaseH to degrade the RNA.Second strand cDNA synthesis was subsequently performed using DNA Polymerase I and dNTP. Remaining overhangs were converted into blunt ends *via* exonuclease/polymerase activities. After adenylation of 3' ends of DNA fragments, Adaptor with hairpin loop structure were ligated to prepare for hybridization. In order to select cDNA fragments of preferentially 370~420 bp in length, the library fragments were purified with AMPure XP system (Beckman Coulter, Beverly, USA). Then PCR amplification, the PCR product was purified by AMPure XP beads, and the library was finally obtained. After the construction of the library, the library was initially quantified by Qubit2.0. After the insert size met the expectation, qRT-PCR was used to accurately quantify the effective concentration of the library (the effective concentration of the library was higher than that of 2nM) to ensure the quality of the library.

After the library was qualified, the different libraries were pooling according to the effective concentration and the target amount of data of the machine, then being sequenced by the Illumina NovaSeq 6000 (Illumina, San Diego, CA, USA). The end reading of 150 bp pairing was generated. Library construction and transcriptome sequencing were conducted by the Novogene Bioinformatics Institute (Beijing, China). All reads have been deposited in the sequence read archive (SRA) under the accession numbers PRJNA897027 and PRJNA896863.

RNA-Seq data analysis

The raw data size was at least 6 Gb for each sample. The end reading of 150 bp pairing was generated. Clean reads were obtained by removing reads containing adapter, reads containing N base, and low quality reads from raw reads. Reference genome and gene model annotation files were downloaded from the genome website (https://orchardgrassgenome.sicau.edu.cn/) (Huang et al., 2020). Building of the reference genome index and alignment of the paired-end clean reads were done using Hisat2 (v2.0.5) (Kim et al., 2019). The mapped reads of each sample were assembled by StringTie (v1.3.3b) (Pertea et al., 2015). The featureCounts v1.5.0-p3 (Liao et al., 2014) was used to count the reads numbers mapped to each gene. Then the fragments per kilobase of exon per million mapped fragments (FPKM) of each gene was calculated based on the gene length and the reads count mapped to it. Differential expression analysis of two groups was performed using the DESeq2 R package (1.20.0) (Love et al., 2014). Padj < 0.05 and $|\log_2(\text{fold-change})| \geq 1$ were set as the threshold for significantly differential expression. ClusterProfiler (3.8.1) (Yu et al., 2012) was used for GO and KEGG enrichment analysis.The overall workflow for transcriptomic analysis is illustrated in Supplementary Figure S1.

Validation of RNA-Seq data by qRT-PCR

Eight DEGs were randomly selected for qRT-PCR. The qRT-PCR reaction volume was 10 µl, containing 1 µl cDNA, 5 µl TB Green Premix Ex Taq II (Tli RNaseH Plus) (2×), 0.4 µl ROX Reference Dye II (50×), 0.8 µl of the forward and reverse primers, and 7 µl ddH_2O. *Actin* was used as the endogenous reference gene. PCR reaction system: 95°C 30s; 95°C 5s, 60°C 34s, 40 Cycles; 95°C 15s, 60°C 1min, 95°C 15s. Primers were designed using Primer 5.0 (Supplementary Table S1). The application of the $2^{-\Delta\Delta C(t)}$ method converts the instrument-generated threshold cycle value output into the relative gene expression level (Livak and Schmittgen, 2001). Three biological replicates were generated and three measurements were performed for each replicate.

Results

Different responses to submergence stress between Dianbei and Anba

The morphologies of Dianbei and Anba under submergence stress were compared at 0 h, 8 h, 24 h, 48 h, and 72 h (Figures 1A–E).

10.3389/fpls.2022.1104755

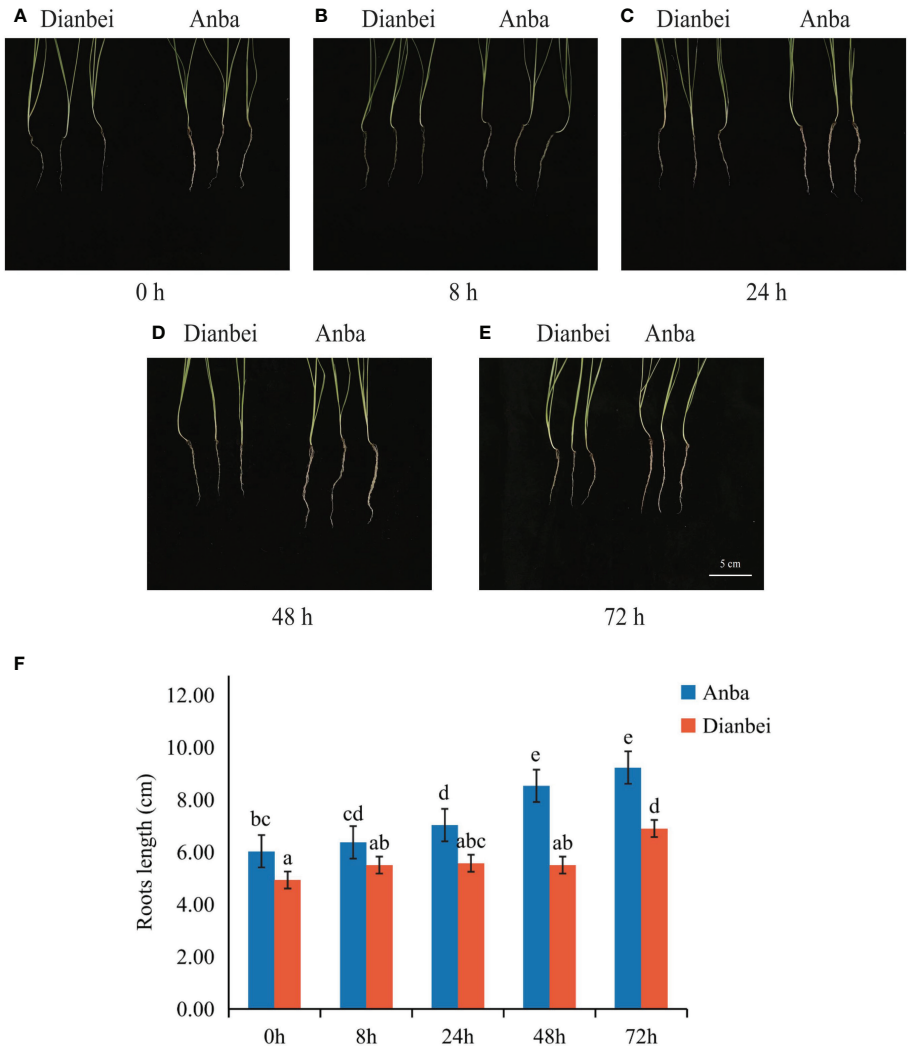


FIGURE 1
Root phenotype and length in Dianbei and Anba. **(A-E)** The phenotype of Dianbei and Aaba at 0 h **(A)**, 8 h **(B)**, 24 h **(C)**, 48 h **(D)**, and 72 h **(E)**. **(F)** Effects of submergence on the root lengths of Dianbei and Anba. Error bars indicate the standard error, n = 3. Different letters indicate the significant differences. Statistical analysis using one-way analysis of variance (ANOVA) using Duncan's multiple range test ($P < 0.05$).

At 0 h, the root of Anba was 1.10 cm longer than Dianbei. After 72 h submergence stress, it was 2.33 cm longer than Dianbei. The roots of Dianbei and Anba gradually elongated with the increasing of submergence time. Anba showed a significant difference at 24 h, while Dianbei showed a significant difference at 72 h. As compared with 0 h, the roots of Dianbei and Anba increased post-submergence-stress by 39.96% and 53.07%, respectively (Figure 1F). Anba responded more positively to submergence stress. Thus, Dianbei is more resistant to submergence than Anba.

Data analysis of RNA-Seq

To explore the gene expression and underlying molecular response mechanism of orchardgrass under short-term submergence stress. Illumina Novaseq 6000 was used to sequence the transcriptome of two orchardgrass cultivars roots at three different time points under submergence stress (0 h, 8 h, and 24 h), and 18 qualified RNA libraries were separately constructed (three library repeats for each time-point). After filtering out low-quality reads and reads containing N base,

792,201,680 clean reads and 118.82 Gb clean bases were obtained. For each sample, the Q30 values exceeded 93%, GC content was between 52.40 to 56.10%, and the error rate was only 0.03 (Supplementary Table S2). The distribution of gene expression levels for each sample was similar, and the overall gene expression level was high (Figure 2A, Supplementary Figure S2). The pearson correlation between samples was more than 0.8 (Figure 2B). Therefore, RNA-seq data was confirmed to be reliable and could be used in subsequent analyses.

HISAT2 software was used to map the clean reads to orchardgrass reference genome (Huang et al., 2020). Moreover, the clean reads average mapping rate was 68.18%, the unique mapping rate was between 60.92% to 70.88% (Figure 2C), and more than 78.8% of the clean reads were mapped to the exon region (Figure 2D). By comparing the FPKM of each sample, it was found that at each time point ~39.67% of the genes (DB_0h 39.05%, DB_8h 39.91%, DB_24h 40.96%, AB_0h 38.09%, AB_8h 39.21%, and AB_24h 40.8%) were expressed (FPKM ≥ 1), with over 3.10% of the genes were highly expressed (FPKM > 60) (Figure 2E).

Differential expression analysis of Dianbei and Anba under submergence stress

To explore the DEGs of Dianbei and Anba in response to submergence stress, the DEGs of two orchardgrass cultivars were screened with the thresholds of $|\log_2$ (fold-change)$|$ > 1 and padj

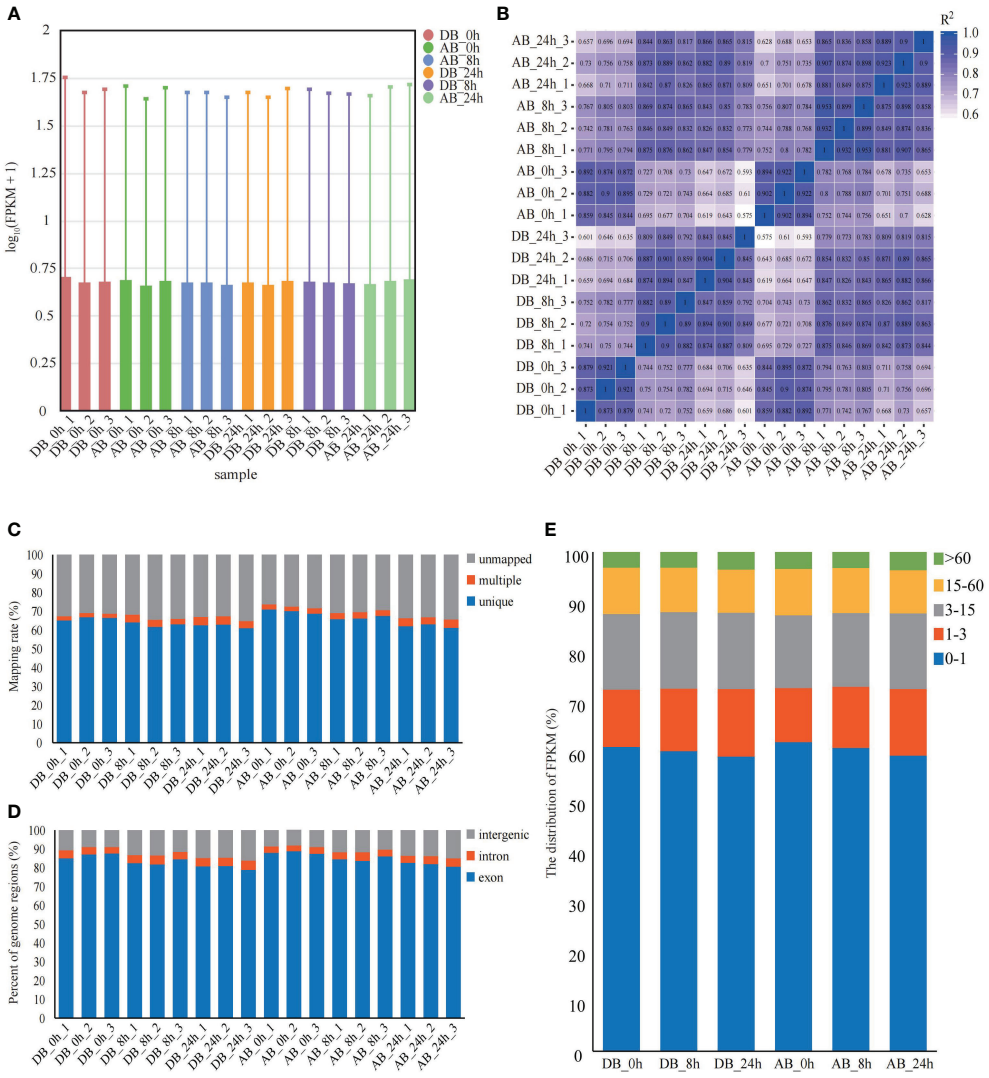


FIGURE 2
Mapping results of Dianbei and Anba. **(A)** The FPKM of Dianbei and Anba. The abscissa in the graph indicates different samples; the ordinate indicates the logarithmic values of the sample expression FPKM. **(B)** Pearson correlation between Dianbei and Anba. **(C-E)** Mapping ratio **(C)**, percent of genome regions **(D)**, and distribution of FPKM **(E)** in the roots of Dianbei and Anba under submergence stress.

< 0.05. As compared with the control, a total of 6663 DEGs (3096 up-regulated and 3567 down-regulated) and 9857 DEGs (4779 up-regulated and 5078 down-regulated) were detected in Dianbei, while 7894 DEGs (3801 up-regulated and 4093 down-regulated) and 11215 DEGs (5752 up-regulated and 5463 down-regulated) were detected in Anba at 8 h and 24 h post-submergence-stress (Figure 3A, Supplementary Figure S3, Supplementary Table S3), respectively. Venn analysis was performed to examine the differential expression of roots tissue at different time points under submergence stress (Figures 3B, C). With 0 h as a control, 5455 and 6202 DEGs were found to be expressed in Dianbei and Anba at both 8 h and

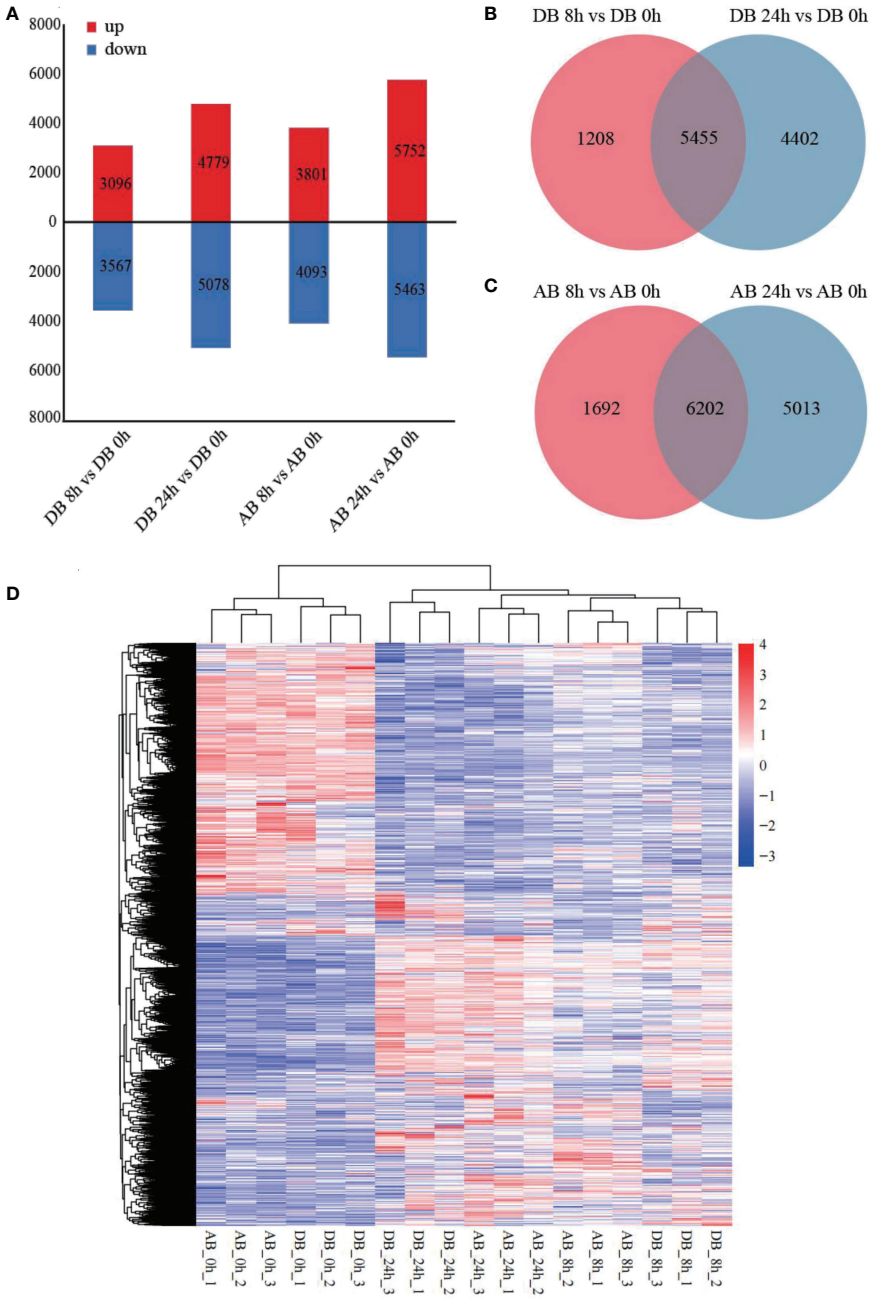


FIGURE 3
DEGs under submergence stress. **(A)** the number of up- and down-regulated genes in four pair-wise sampling stages, including DB 8 h vs DB 0 h, DB 24 h vs DB 0 h, AB 8 h vs AB 0 h, and AB 24 h vs AB 0 h **(B, C)** Venn analysis of the DEGs at 8 h and 24 h of post-submergence-stress induction as compared to control in Dianbei **(B)**, and Anba **(C)**. **(D)** the heatmap shows the respective expression levels of DEGs in each sample group, based on the average FPKM of biological replicates. The y-axis shows the cluster dendrogram of DEGs, and the x-axis shows the sample groups.

24 h, respectively. With the increasing submergence stress time, the number of DEGs in the two orchardgrass cultivars gradually increased.

Gene cluster analysis of the two orchardgrass cultivars under submergence stress showed that Dianbei and Anba had similar expression patterns at the same stress time points, thus indicating that there were noticeable differences in the expression patterns of DEGs at different stress time points (Figure 3D). The differential expression of DEGs in Dianbei was higher than that in Anba at 8 h and 24 h post-submergence-stress, thus indicating that Dianbei, as a submergence-tolerant cultivar, has more DEGs for responding to submergence stress.

GO and KEGG analysis of two orchardgrass cultivars DEGs under submergence stress

To explore the functional significance of DEGs at different time points in Dianbei and Anba under submergence stress, we subjected GO enrichment analysis and obtained 986 terms. The DEGs were classified into three categories: biological processes (BP), molecular functions (MF), and cellular components (CC). The top 10 enriched terms of each category were selected for plotting. DEGs were mainly enriched in terms of biosynthesis and metabolism process, transcription factor activity, enzyme activity and coenzyme binding, oxidative response, microtubule, Golgi apparatus and vesicle (Figure 4, Supplementary Table S4). It is speculated that submergence stress may stimulate the expression of related genes in orchardgrass.

In addition, KEGG enrichment analysis was performed and DEGs were mapped to 123 pathways (Figure 5, Supplementary Table S5). The results showed that pathways like phenylpropanoid biosynthesis, glycolysis/gluconeogenesis, and nitrogen metabolism were significantly enriched in both Dianbei and Anba. Enrichment of these pathways might indicate that submergence stress promoted their activation. It is worth mentioning that plant hormone signal transduction, MAPK signaling pathway, and plant-pathogen interaction were significantly enriched in Dianbei, but not in Anba. Therefore, it could be speculated that these pathways and genes might be related to the submergence tolerance in Dianbei and Anba.

Plant hormone signal transduction in orchardgrass roots under submergence stress

Plant hormones regulate plant growth and other biological processes as along with stress adaptation (Verma et al., 2016). In this study, the DEGs under submergence stress condition were analyzed for their KEGG pathway enrichment (Figure 5). The results showed that submergence stress activated various plant hormones (auxin, cytokinin, ethylene, abscisic acid (ABA), jasmonic acid (JA), salicylic acid (SA), brassinosteroids (BR), gibberellin (GA), etc.) signal pathways in orchardgrass roots (Figure 6, Supplementary Table S6). In Dianbei, multiple genes related to ARF and GH3 in auxin, CRE1 in cytokinin, GID1 in GA, PP2C, SnRK2 and ABF in ABA, and NPR1 in SA, were down-regulated at 8 h and 24 h post-submergence-stress, triggering speculation that these genes were inhibited under submergence stress. Multiple genes related to PYR/PYL in ABA, ETR and EBF1/2 in ethylene, and BRI1 in BR, were up-regulated at 8 h and 24 h post-submergence-stress, indicating that these genes were activated under submergence stress. Additionally, multiple genes related to AUX1 in auxin, DELLA in GA, SIMKK in ethylene, and TCH4 in BR, were up-regulated at 24 h, and PR-1 in SA, JAR1 and MYC2 in JA, were down-regulated at 24 h, but had no significant change at 8 h post-submergence-stress. These results indicate that these genes may need a certain time to respond to submergence stress. Most DEGs were down-regulated post submergence stress. Thus, we speculated that submergence stress may inhibit plant hormones signal transduction, the inhibitory effect increased with the increasing submergence time. In addition, among the DEGs involved in plant hormone signal transduction, only 10 DEGs (DG2C00767, DG2C00772, DG2C01118, DG2C02515, DG2C06118, DG4C03948, DG5C01306, DG5C01122, DG6C01437, and DG7C03758) had higher differential expression in Anba than in Dianbei, while the other DEGs showed the opposite, thus indicating that the differential expression of these genes may be one of the reasons for the different submergence tolerance capabilities between Dianbei and Anba.

MAPK signal transduction in orchardgrass roots under submergence stress

MAPK signal transduction is one of the most well-studied plant signaling mechanisms, and plant MAPK cascades play pivotal roles in signaling plant defense against biotic and abiotic stresses. (Zhang and Klessig, 2001). The MAPK cascade is minimally composed of different combinations of at least three protein kinases: MAPKKK (MAP3K/MEKK/MKKK), MAPKK (MKK/MEK), and MAPK (MPK), which activate each other sequentially *via* phosphorylation (Danquah et al., 2014) (Figures 7A, B). In this study, DEGs related to MAPK signal transduction, FLS2 was up-regulated and activated the transmembrane transport of flg22, while most DEGs related to MAPK cascade signal were down-regulated, thus indicating that submergence stress may inhibit the amplification of the MAPK cascade signal. Two DEGs, DG1C06357 and DG1C04245, showed significant differential expression at 8 h, but had no significant change at 24 h of post-submergence-stress, thus indicating that these two DEGs may play an important role in

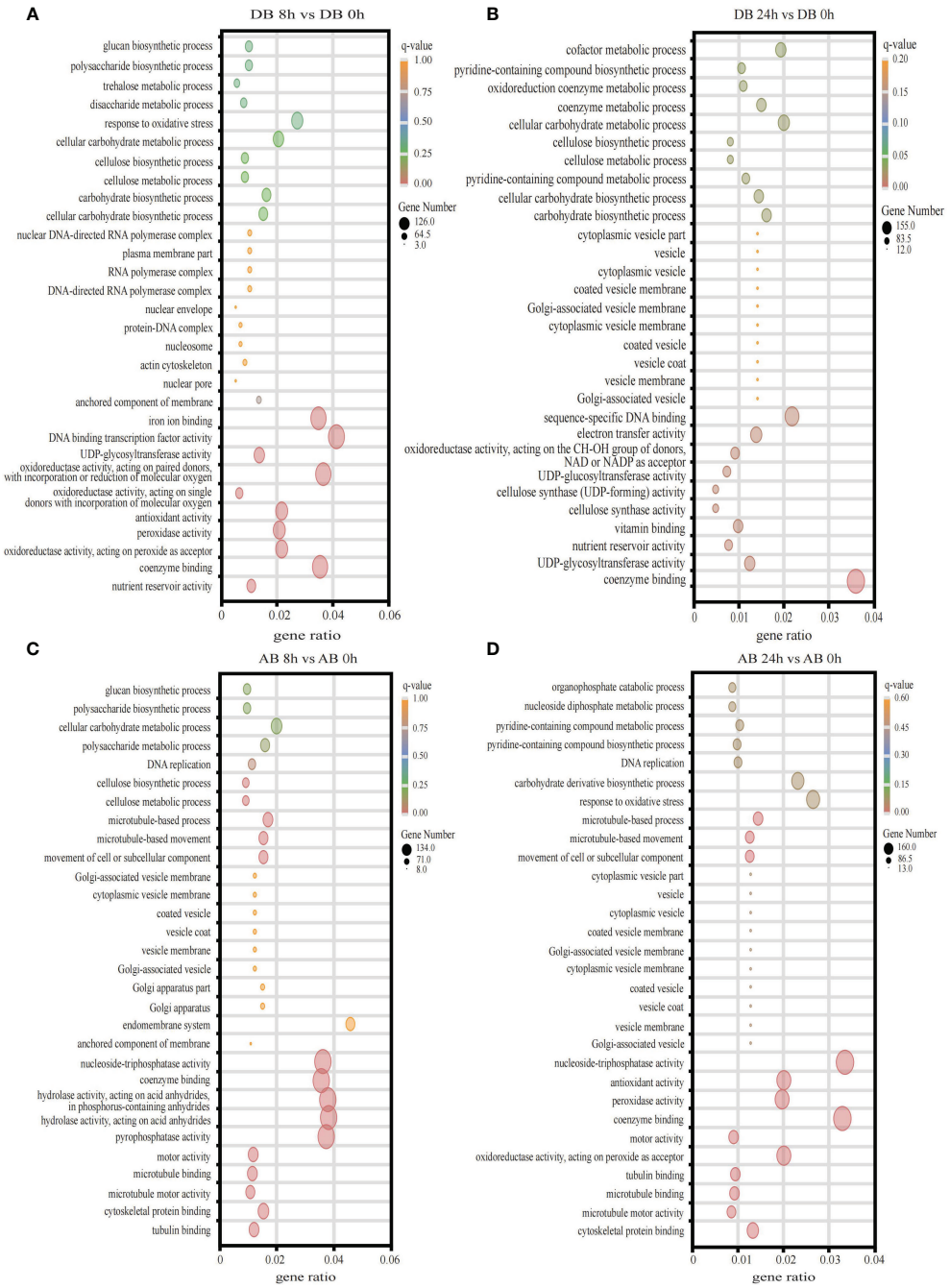


FIGURE 4
The GO enrichment analysis. **(A)** DB 8 h vs DB 0 h **(B)** DB 24 h vs DB 0 h **(C)** AB 8 h vs AB 0 h **(D)** AB 24 h vs AB 0 h The ordinate indicates the GO id, and the abscissa indicates the gene ratio. The size of the dot indicates the number of DEGs in the pathway, and the color indicates the different q-value.

the early submergence stress response. The differential expression of other DEGs showed an increasing trend with the increasing submergence time. Out of these DEGs, seven DEGs (DG6C02930, DG2C04602, DG3C07121, DG6C02696, DG4C06428, DG5C04434, and DG4C03116) had no significant change at 8 h, but showed significant differential expression at 24 h post-submergence-stress, thus indicating that these genes may need some time to respond to submergence stress. In addition, eight DEGs (DG0C00183, DG0C00192, DG0C00190, DG5C03912, DG2C05975, DG6C00846, DG6C00463, and DG1C05223) were differentially expressed at 8 h and 24 h post-submergence-stress in Dianbei, while showing no significant expression in Anba, thus indicating that these DEGs may be one of the reasons for the different submergence

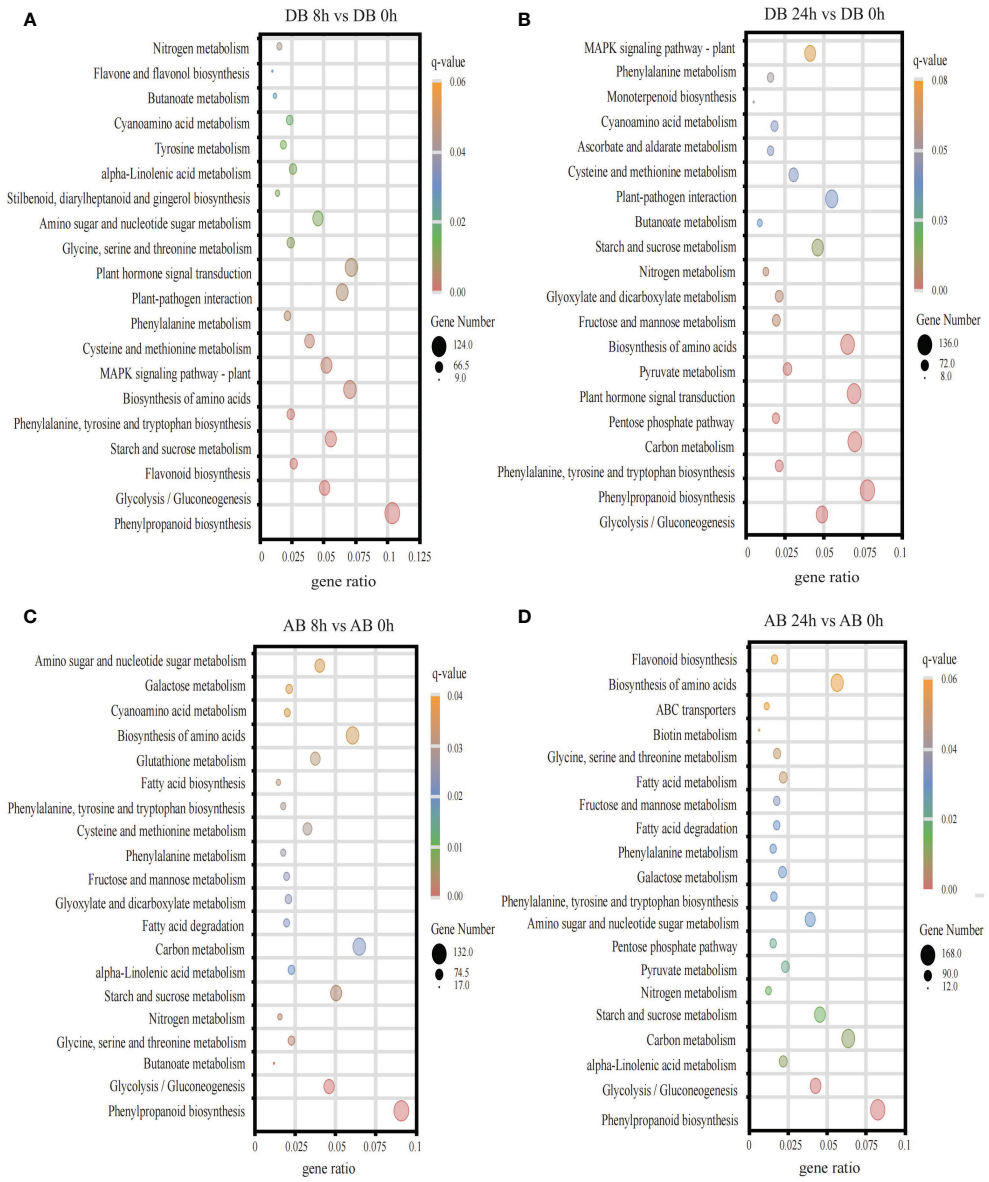


FIGURE 5
The KEGG pathway enrichment analysis. **(A)** DB 8 h vs DB 0 h **(B)** DB 24 h vs DB 0 h **(C)** AB 8 h vs AB 0 h **(D)** AB 24 h vs AB 0 h The ordinate indicates metabolic pathway, while the abscissa indicates the gene ratio. The size of the dot indicates the number of DEGs in the pathway, while the color indicates the different q-value.

tolerance abilities between Dianbei and Anba (Figures 7B, C, Supplementary Table S6).

Additionally, studies have been shown that MAP kinases are involved in ABA-mediated stomatal closure. The active SnRK2 kinase phosphorylates the NADPH oxidase RbohF, leading to ROS accumulation. The ROS-mediated activation of MAPKs positively regulates the ABA-mediated stomatal closure (Danquah et al., 2014). In this study, three types of DEGs were found to participate in this pathway (Figures 6H, 7C). Among them, the genes related to PYR/PYL receptor proteins were up-regulated, while the genes related to PP2C and SnRK2 kinase biosynthesis were down-regulated. Dianbei showed a greater variation than Anba, and the differential expression of DEGs increased with the increasing submergence time. It indicated that submergence stress activated the PYR/PYL expression in the upstream, and inhibited the expression of PP2C and SnRK2 in the downstream of ABA pathway, thus inhibiting ABA-mediated stomatal closure and increasing the respiration of orchardgrass under submergence stress, and promoting more oxygen uptake. Therefore, Dianbei can respond more quickly and positively to submergence stress, which may be one of the reasons for the different submergence tolerance between Dianbei and Anba.

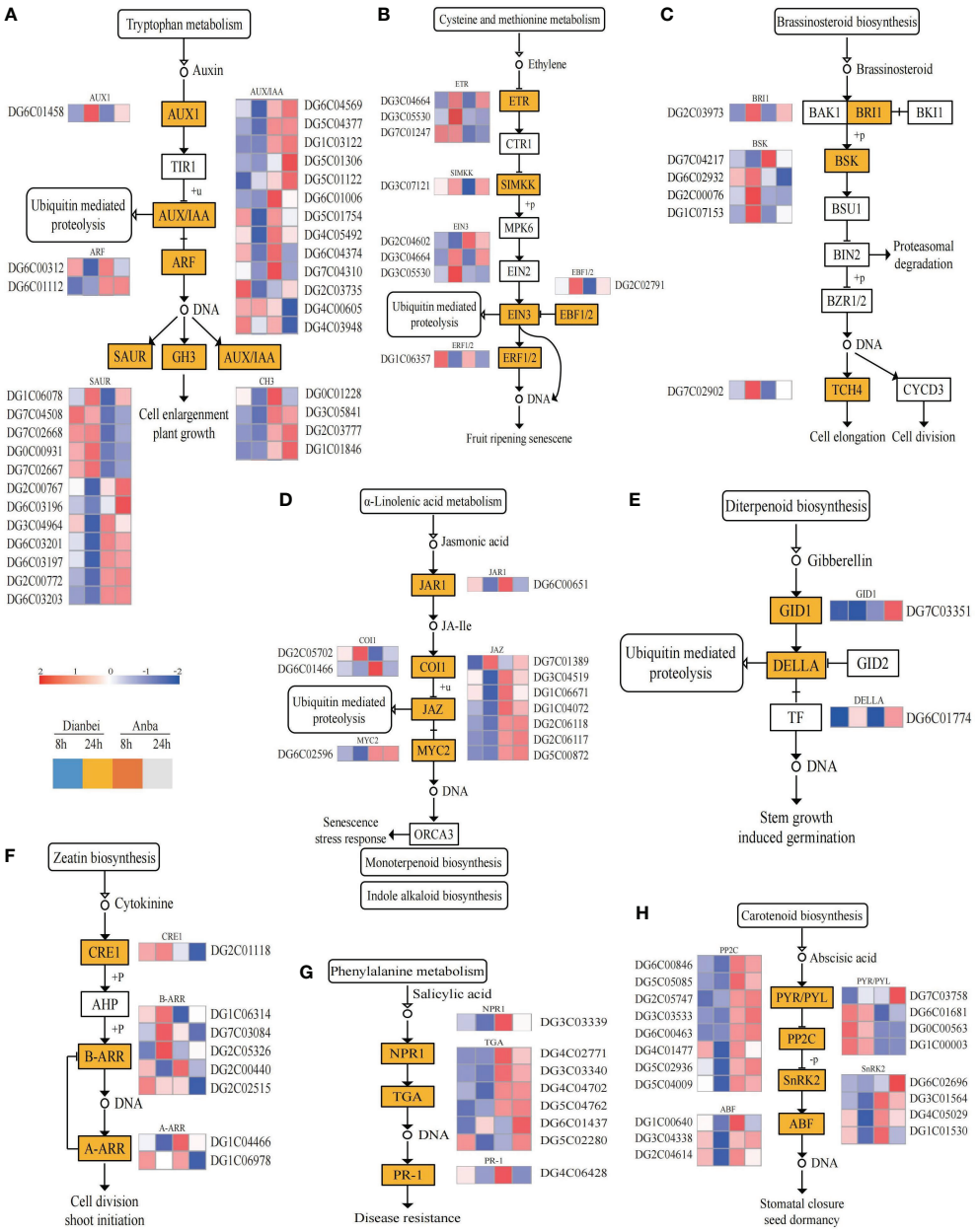


FIGURE 6
Plant hormone signal transduction pathway. **(A–H)** Auxin **(A)**, Ethylene **(B)**, Brassinosteroid **(C)**, Jasmonic acid **(D)**, Gibberellin **(E)**, Cytokinine **(F)**, Salicylic acid **(G)**, and Abscisic acid **(H)** biosynthesis pathway. The y-axis shows the cluster dendrogram of DEGs, and the x-axis shows the sample groups.

Ca^{2+} signal transduction in orchardgrass roots under submergence stress

During the KEGG pathway enrichment analysis under submergence stress in orchardgrass, it was found that in plant-pathogen interaction signal transduction, 39 DEGs also participated in the Ca^{2+}-mediated hypersensitive response (HR) besides MAPK signal transduction pathway (Figure 8, Supplementary Table S6). Similar to the plant hormone and MAPK signal transduction pathways, the differential expression of DEGs increased with the increase of submergence time except for DG4C04884 and DG2C02769. The amplitude of variation in Dianbei was greater than that in Anba, which again indicated that Dianbei could respond more rapidly and positively than Anba under submergence stress. Interestingly, the genes related to Ca^{2+} transmembrane CNGCs receptor proteins and NOS (NO synthesis

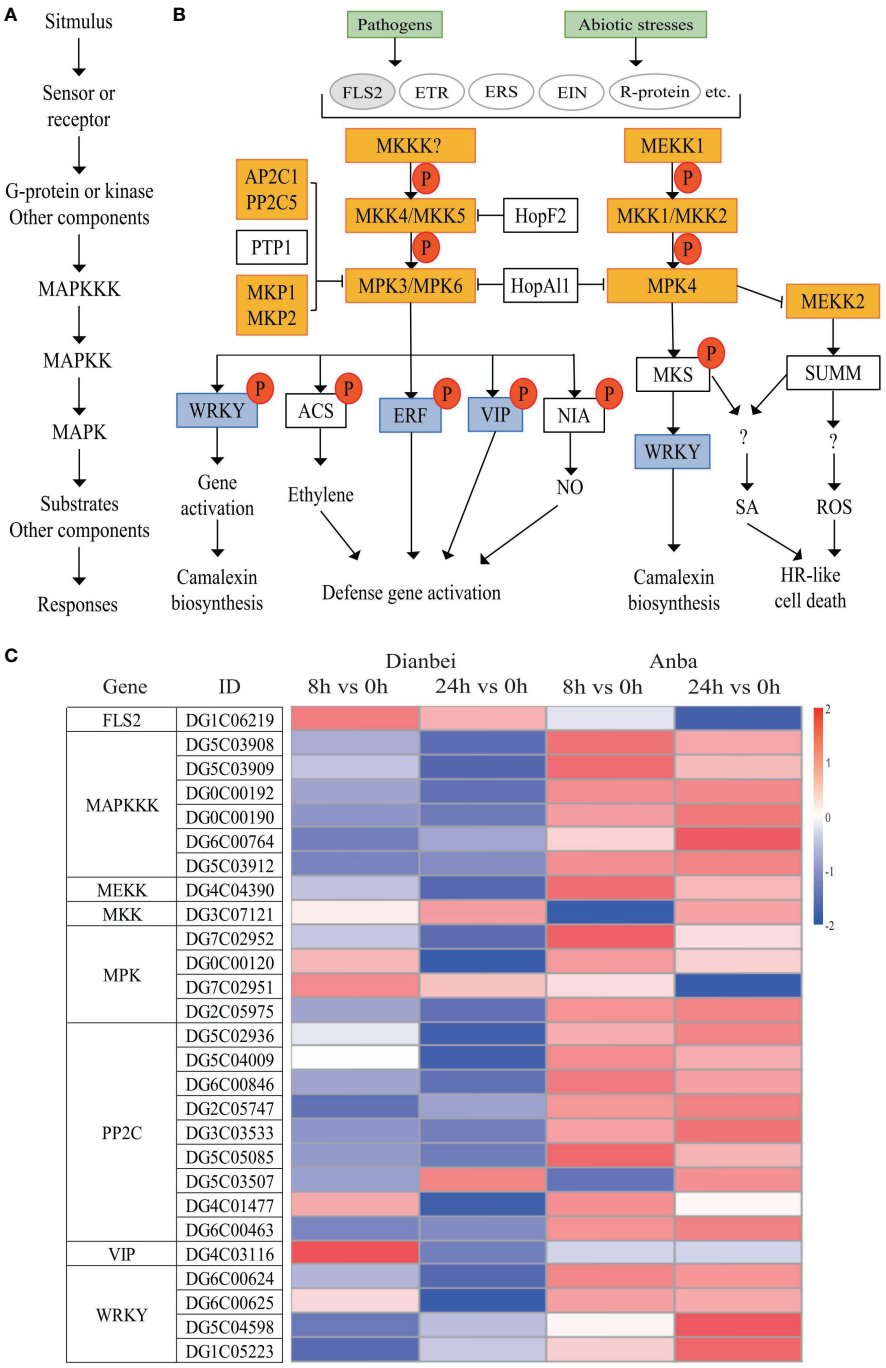


FIGURE 7
MAPK signal transduction pathway. **(A)** Plant MAPK cascades pattern. **(B)** Possible functions of various MAPK cascades in plants. **(C)** Heatmap of DEGs involved in the MAPK signaling pathway. The y-axis shows the cluster dendrogram of DEGs, and the x-axis shows the sample groups.

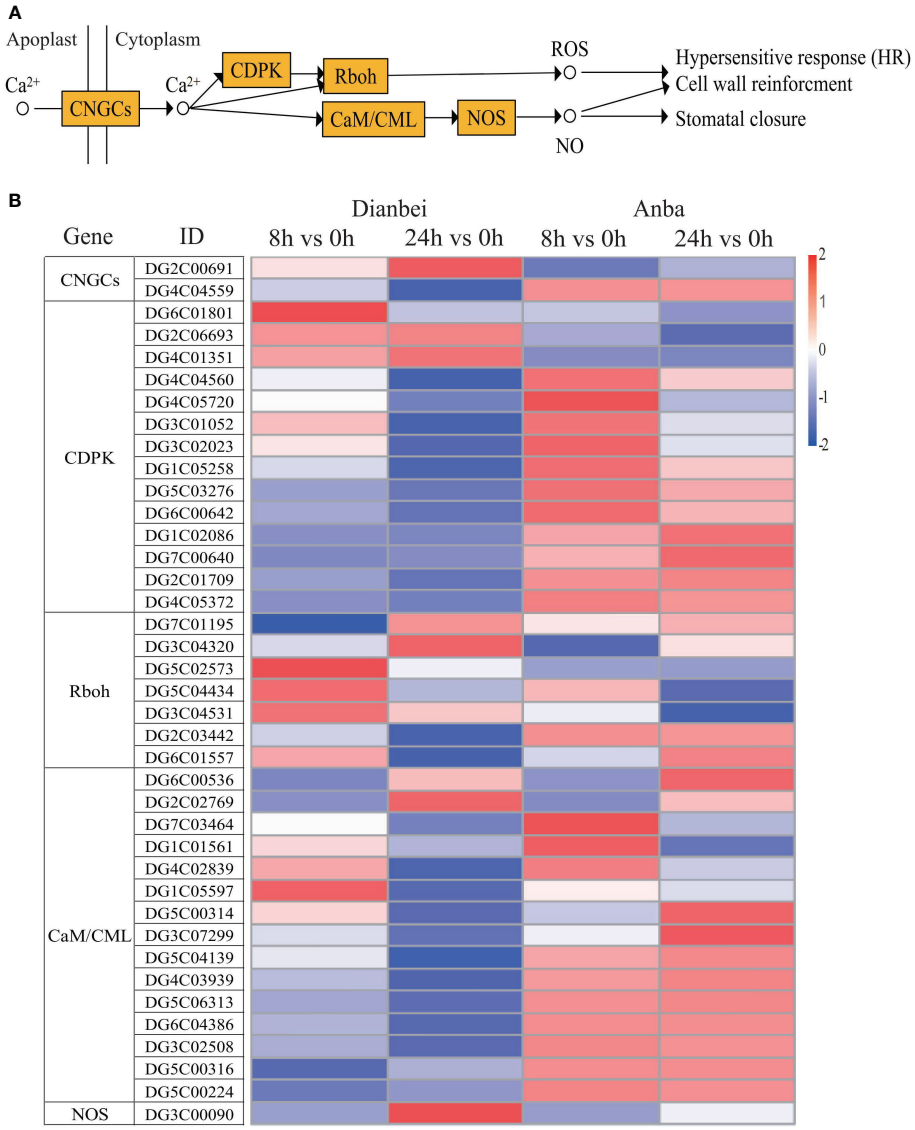


FIGURE 8
Ca^{2+} signal transduction pathway. **(A)** Schematic diagram of Ca^{2+} signaling pathway. **(B)** Heatmap of DEGs involved in Ca^{2+} signaling pathway. The y-axis shows the cluster dendrogram of DEGs, while the x-axis shows the sample groups.

rate limiting enzymes) were up-regulated, which activated Ca^{2+} transmembrane transport, NO synthesis, and induced HR production in orchardgrass under submergence stress.

Validation of transcriptome sequencing data by qRT-PCR

To verify the reliability of transcriptome results, we selected eight genes for qRT-PCR (Figure 9). The expression profiles of these genes were consistent with the sequencing results, indicating that our analysis based on transcriptome data is reliable.

Discussion

Water is essential for all plants, but excessive water or submergence results in stress and prevents gaseous exchange between the soil and atmosphere. Excessive water can inhibit the growth and development and even lead to death. In this study, a model was constructed by transcriptome analysis to understand the response of orchardgrass under submergence stress (Figure 10).

In one day of waterlogging, the partial pressure of O_2 fell from 20.9 to 1 kPa, leading to anoxia (Phukan et al., 2016). To self-repair and maintain stability, tissues or organs specific functions cells require the mitochondrial oxidative phosphorylation of ATP (Phukan et al., 2016). Therefore, in

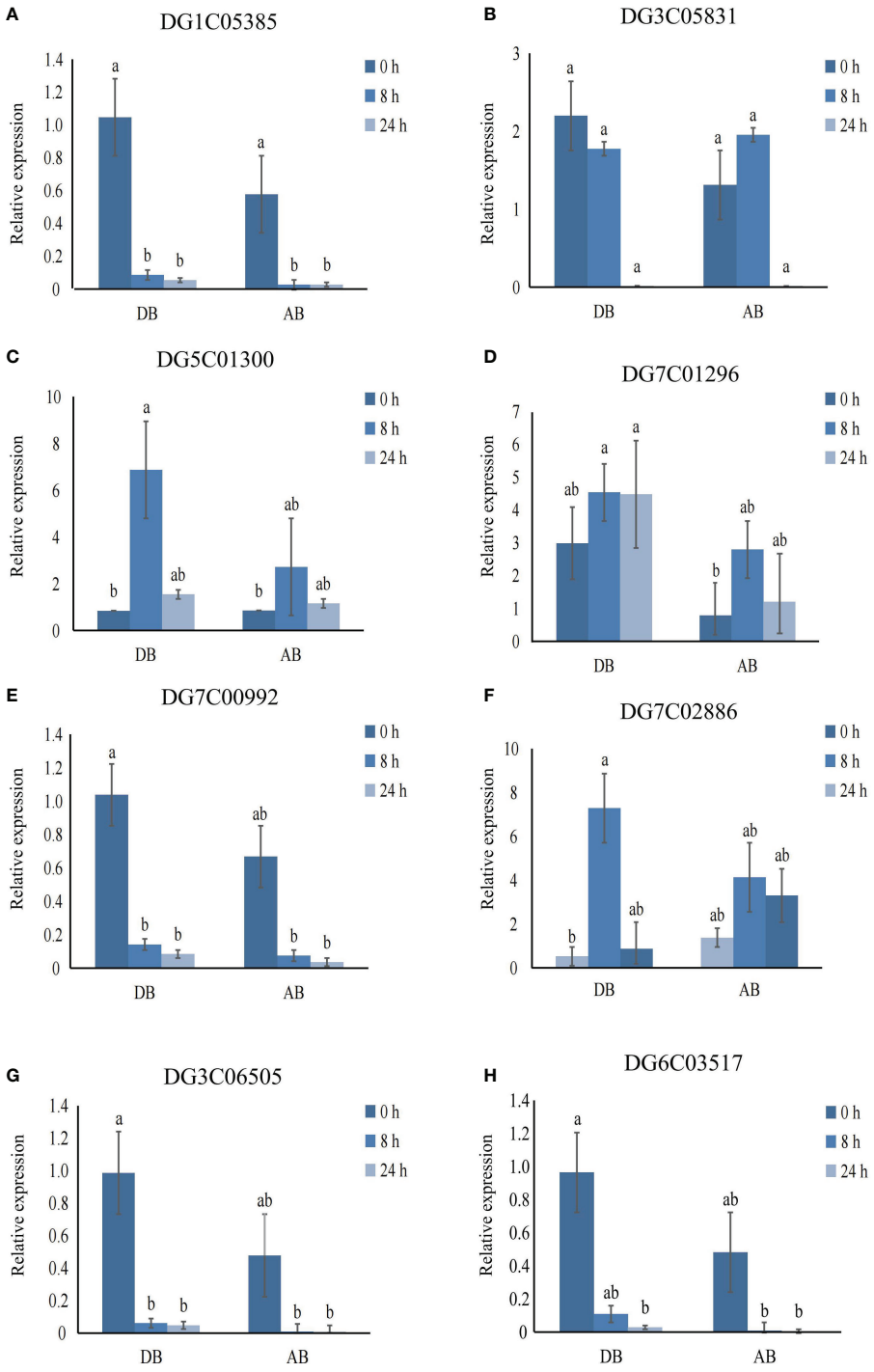


FIGURE 9
Validation of genes expression using qRT-PCR. **(A)** DG1C05385. **(B)** DG3C05831. **(C)** DG5C01300. **(D)** DG7C01296. **(E)** DG7C00992. **(F)** DG7C02886. **(G)** DG3C06505. **(H)** DG6C03517. Error bars indicate the standard error, n = 3. Different letters indicate the significant differences. Statistical analysis using one-way analysis of variance (ANOVA) using Duncan's multiple range test ($P < 0.05$).

order to survive the energy crisis in waterlogged or submerged conditions, ATP related processes, like DNA synthesis and cell division, are inhibited (Gibbs and Greenway, 2003). In this study, we found that DEGs of two orchardgrass cultivars were significantly expressed in biosynthesis and metabolism process, transcription factor activity, enzyme activity and coenzyme binding, oxidative response and cellular components (microtubules, Golgi, and vesicles) terms under submergence

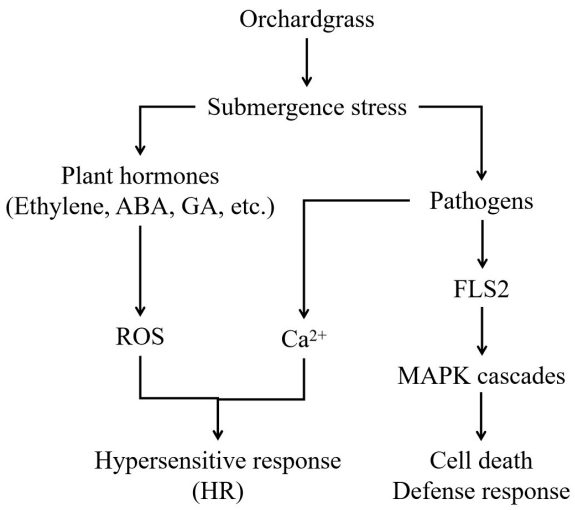


FIGURE 10
A proposed model for orchardgrass under submergence stress. The figure depicts the interplay and crosstalk of plant hormones, pathogen interaction, and MAPK cascades in orchardgrass during submergence stress.

stress (Figure 4). Therefore, the results indicated that the synthesis and metabolism of intracellular molecules are important in the submergence stress response of orchardgrass.

Waterlogging reduces the soil permeability, inhibits the plant TCA cycle, promotes the fermentation of the glycolysis-generated pyruvate acid to produce ethanol and NAD^+, thus altering the energy metabolism and causing the accumulation of toxic substances in roots (Judge and Dodd, 2020). Glycolysis and galactose metabolism are important glucose metabolism processes in plants, which provide energy for the plant growth and development. Strengthening the glycolysis pathway can alleviate the hypoxia-induced energy shortage under waterlogging stress (Kolahi et al., 2021). In quinoa (*Chenopodium quinoa* Willd.), the expression of saccharides and alcohol-related genes in Dianli-1299 (waterlogging-resistant cultivar) increased (Guo et al., 2022). Nitrogen application can effectively improve the development of rapeseed roots and reduce the waterlogging stress-induced damage (Men et al., 2020). In this study, it was found that, phenylpropanoid biosynthesis, glycolysis/gluconeogenesis, and nitrogen metabolism were significantly enriched in two orchardgrass cultivars under submergence stress (Figure 5), indicating that phenylpropanoid biosynthesis, glycolysis/gluconeogenesis and nitrogen metabolism participate in maintaining the balance of energy metabolism in orchardgrass roots under submergence stress. In addition, plant hormone signal transduction, MAPK signaling pathway, and plant-pathogen interaction were significantly enriched in Dianbei, but not in Anba (Figure 5). It is speculated that these pathways might be related to the submergence tolerance in Dianbei and Anba.

Plant hormones mediate plants abiotic stress response (Shu et al., 2018). The interaction of plant hormones is crucial in waterlogging and submergence tolerance. Some plant endogenous hormones, like abscisic acid, ethylene, auxin and cytokinin, are very sensitive to waterlogging stress and can effectively alleviate its adverse effects (Bashar, 2018; Qi et al., 2019; Hu et al., 2020). The level of ethylene increases in waterlogged plants, which has been identified as a signal that regulates the early response to submergence stress (Phukan et al., 2016). ERF transcription factors are regulated by ethylene. Exogenous ethylene significantly promotes ERF transcription activity in *Arabidopsis* and soybean (*Glycine max* L.) (Hess et al., 2011; Tamang et al., 2014). ETR1 activates EIN2/EIN3, which induces the expression of ERF transcription factors, and promotes ethylene production under waterlogging stress (Pierik et al., 2006). Plants, such as rice, which slowly produces ethylene in its roots, positively respond to ethylene. Ethylene reduces the level of ABA level and enhances the sensitivity of GA in tissue (van der Knaap and Kende, 1995). In rice, *Sub1A* inhibited the accumulation of DELLA, a GA-responsive protein, and prolonged the submergence tolerance time by preserving ATP and CHO content (Peña-Castro et al., 2011). The endogenous ABA concentration did not decrease under flooding in soybean, but it was decreased in untreated plants. The transcriptional expression system in soybean which provided exogenous ABA was better, thus indicating that exogenous ABA enhanced the waterlogging tolerance in soybean (Komatsu et al., 2013). In this study, in Dianbei, the genes related to ETR, SIMKK, EIN3, and EBF1/2 in Ethylene were up-regulated, GID1 and DELLA in GA were down-regulated, and PYR/PYL were up-regulated, which inhibited the downstream expression of PP2C, SnRK2, and ABF in ABA under submergence stress (Figure 6). This indicates that submergence stress stimulates ethylene biosynthesis and metabolism, and reduces ABA and GA in Dianbei. These results were similar to the results in rice and soybean under waterlogging stress (Peña-Castro et al., 2011; Komatsu et al., 2013). Most DEGs related to plant hormone signal transduction were differentially expressed in Dianbei, and the differential expression was down-regulated with the increasing submergence treatment time (Figure 6). Therefore, we speculate that the differential expression of these genes might be one of the reasons for the different submergence tolerance in Dianbei and Anba. In addition, the submergence stress may inhibit plant hormone signal transduction, which increases with the increasing submergence time.

Under waterlogging stress, an efficient carbohydrate mobilization mechanism in plant roots enables cells to survive prolonged hypoxia (Phukan et al., 2016). Cytosolic Ca^{2+} increased rapidly in maize and *Arabidopsis*, which ultimately altered the expression of hypoxia-responsive genes in hypoxia conditions (Subbaiah and Sachs, 2003; Bailey-Serres and Chang, 2005). Hypoxia leads to ROS production, in turn, ROS signal

transduction promotes various defense responses (Steffens et al., 2012). In *Arabidopsis*, ROS activates MAPK6 to improve the survival under hypoxic conditions. (Chang et al., 2012). However, excessive ROS causes oxidative damage. Therefore, the ROS defense and signaling cascade are strictly regulated in plants to maintain an appropriate balance between survival and stress tolerance (Sachdev et al., 2021). Waterlogging could induce MAPK cascades, promote the regulation of plant hormones, and forme aerenchyma tissues, which modulates the morphological adaptations in maize roots (Kaur et al., 2021). In *Arabidopsis*, MAPK led to phosphorylation of ACC synthase and participated in aerenchyma formation (Liu and Zhang, 2004). In this study, CNGCs, Ca^{2+} transmembrane transduction-mediating receptor proteins, were up-regulated and activated Ca^{2+} transmembrane transport in orchardgrass under submergence stress. Most genes in MAPK signal transduction were down-regulated, which inhibited the amplification of MAPK cascade signaling (Figures 7, 8). Moreover, with the increasing submergence time, the amplitude of differential expression of genes gradually increased, and Dianbei was greater than Anba. Therefore, we speculate that the differential expression of these genes might be one of the reasons for the different submergence tolerance abilities in Dianbei and Anba.

Conclusion

In this study, we tested the morphology indexes of roots of two orchardgrass cultivars at 0 h, 8 h, 24 h, 48 h and 72 h. Simultaneously, we obtained 118.82 Gb clean data from RNA-seq analysis of orchardgrass roots. At 8 h and 24 h post-submergence-stress, Dianbei identified 6663 and 9857 DEGs, while Anba identified 7894 and 11215 DEGs, respectively. With 0 h as a control, 5455 and 6202 DEGs were differentially expressed at 8 h and 24 h in Dianbei and Anba, respectively. These results indicate that Anba has more genes to respond to submergence stress. Moreover, genes related to biosynthesis and metabolism, cellular component, transcription factor activity, enzyme activity and coenzyme binding, phenylpropanoid biosynthesis, glycolysis/gluconeogenesis, and nitrogen metabolism were involved in the submergence stress response in orchardgrass. The expression of genes involved in plant hormone, MAPK, and Ca^{2+} signal transduction were significantly in Dianbei, but not in Anba. The differential expression of these genes and pathways may be the main reasons behind the different submergence tolerance abilities in Dianbei and Anba. Since the submergence tolerance of orchardgrass is controlled by multiple genes, RNA-seq can be used for comprehensively exploring submergence tolerance-related genes and pathways. Therefore, the study can facilitate further understanding of the molecular regulatory mechanism in orchardgrass roots under submergence stress conditions.

Data availability statement

The datasets presented in this study can be found in online repositories. All datasets can be found in the article/Supplementary Material.

Author contributions

BZ (teacher, corresponding author), YF, MQ, and YZ conceived and designed the project and the strategy. MQ, LB, and BZ (student, co-author) contributed to plant samples collection. MQ, XY, and PS performed experiments. MQ, BS, and WL analyzed data. BZ (teacher, corresponding author), YF, MQ, and XZ wrote the manuscript.

Funding

This research work was funded by the Science and Technology Project of Guizhou Province (Qian Kehe Support [2021] General 226), the Chongqing Modern Mountain Characteristic Efficient Agricultural Industry Technology System (herbivorous livestock) (2023), and the Student Innovation and Entrepreneurship Training Program of Southwest University (X202210635542 and X202210635544).

Acknowledgments

We thank the College of Animal Science and Technology of Southwest University for providing the experimental conditions. We also thank the Chongqing Academy of Animal Science for its advice on the experiment.

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.2022.1104755/full#supplementary-material

SUPPLEMENTARY TABLE 1
The Primer design of genes.

SUPPLEMENTARY TABLE 2
Transcriptome analysis of 18 orchardgrass roots.

SUPPLEMENTARY TABLE 3
The list of differentially expressed genes.

SUPPLEMENTARY TABLE 4
The list of GO enrichment.

SUPPLEMENTARY TABLE 5
The list of KEGG pathway enrichment.

SUPPLEMENTARY TABLE 6
The log_2(fold-change) in plant hormone, MAPK, and Ca^{2+} signal transduction pathway.

References

Anee, T. I., Nahar, K., Rahman, A., Mahmud, J. A., Bhuiyan, T. F., Alam, M. U., et al. (2019). Oxidative damage and antioxidant defense in *Sesamum indicum* after different waterlogging durations. *Plants* 8, 196. doi: 10.3390/plants8070196

Arora, K., Panda, K. K., Mittal, S., Mallikarjuna, M. G., Rao, A. R., Dash, P. K., et al. (2017). RNAseq revealed the important gene pathways controlling adaptive mechanisms under waterlogged stress in maize. *Sci. Rep.* 7, 10950. doi: 10.1038/s41598-017-10561-1

Bailey-Serres, J., and Chang, R. (2005). Sensing and signalling in response to oxygen deprivation in plants and other organisms. *Ann. Bot.* 96, 507–518. doi: 10.1093/aob/mci206

Bashar, K. K. (2018). Hormone dependent survival mechanisms of plants during post-waterlogging stress. *Plant Signal Behav.* 13, e1529522. doi: 10.1080/15592324.2018.1529522

Chang, R., Jang, C. J. H., Branco-Price, C., Nghiem, P., and Bailey-Serres, J. (2012). Transient MPK6 activation in response to oxygen deprivation and reoxygenation is mediated by mitochondria and aids seedling survival in *Arabidopsis*. *Plant Mol. Biol.* 78, 109–122. doi: 10.1007/s11103-011-9850-5

Danquah, A., de Zelicourt, A., Colcombet, J., and Hirt, H. (2014). The role of ABA and MAPK signaling pathways in plant abiotic stress responses. *Biotechnol. Adv.* 32, 40–52. doi: 10.1016/j.biotechadv.2013.09.006

Feng, G., Xu, L., Wang, J., Nie, G., Bushman, B. S., Xie, W., et al. (2018). Integration of small RNAs and transcriptome sequencing uncovers a complex regulatory network during vernalization and heading stages of orchardgrass (*Dactylis glomerata* l.). *BMC Genomics* 19, 727. doi: 10.1186/s12864-018-5104-0

Gibbs, J., and Greenway, H. (2003). Review: Mechanisms of anoxia tolerance in plants. i. growth, survival and anaerobic catabolism. *Funct. Plant Biol.* 30, 353. doi: 10.1071/PP98095

Guo, Y., Wang, Q., Zhang, H., Huang, T., Zhang, X., Xie, H., et al. (2022). Responses to waterlogging stress in quinoa seedlings based on metabolomic and transcriptomic analysis. *Environ. Exp. Bot.* 203, 105044. doi: 10.1016/j.envexpbot.2022.105044

Herzog, M., Striker, G. G., Colmer, T. D., and Pedersen, O. (2016). Mechanisms of waterlogging tolerance in wheat - a review of root and shoot physiology. *Plant Cell Environ.* 39, 1068–1086. doi: 10.1111/pce.12676

Hess, N., Klode, M., Anders, M., and Sauter, M. (2011). The hypoxia responsive transcription factor genes ERF71/HRE2 and ERF73/HRE1 of *Arabidopsis* are differentially regulated by ethylene. *Physiol. Plant* 143, 41–49. doi: 10.1111/j.1399-3054.2011.01486.x

Hirata, M., Yuyama, N., and Cai, H. (2011). Isolation and characterization of simple sequence repeat markers for the tetraploid forage grass *Dactylis glomerata*. *Plant Breed.* 130, 503–506. doi: 10.1111/j.1439-0523.2010.01831.x

Huang, L., Feng, G., Yan, H., Zhang, Z., Bushman, B. S., Wang, J., et al. (2020). Genome assembly provides insights into the genome evolution and flowering regulation of orchardgrass. *Plant Biotechnol. J.* 18, 373–388. doi: 10.1111/pbi.13205

Huang, L. K., Yan, H. D., Zhao, X. X., Zhang, X. Q., Wang, J., Frazier, T., et al. (2015). Identifying differentially expressed genes under heat stress and developing molecular markers in orchardgrass (*Dactylis glomerata* l.) through transcriptome analysis. *Mol. Ecol. Resour.* 15, 1497–1509. doi: 10.1111/1755-0998.12418

Hu, J., Ren, B., Dong, S., Liu, P., Zhao, B., and Zhang, J. (2020). Comparative proteomic analysis reveals that exogenous 6-benzyladenine (6-BA) improves the defense system activity of waterlogged summer maize. *BMC Plant Biol.* 20, 44. doi: 10.1186/s12870-020-2261-5

Hu, T., Sun, X., Zhang, X., Nevo, E., and Fu, J. (2014). An RNA sequencing transcriptome analysis of the high-temperature stressed tall fescue reveals novel insights into plant thermotolerance. *BMC Genomics* 15, 1147. doi: 10.1186/1471-2164-15-1147

Ji, Y., Chen, P., Chen, J., Pennerman, K. K., Liang, X., Yan, H., et al. (2018). Combinations of small RNA, RNA, and degradome sequencing uncovers the expression pattern of microRNA–mRNA pairs adapting to drought stress in leaf and root of *Dactylis glomerata* l. *Int. Mol. Sci.* 19, 3114. doi: 10.3390/ijms19103114

Judge, A., and Dodd, M. S. (2020). Metabolism. *Essays Biochem.* 64, 607–647. doi: 10.1042/EBC20190041

Karunarathna, A., Tibpromma, S., Jayawardena, R. S., Nanayakkara, C., Asad, S., Xu, J., et al. (2021). Fungal pathogens in grasslands. *Front. Cell Infect. Microbiol.* 11. doi: 10.3389/fcimb.2021.695087

Kaur, G., Vikal, Y., Kaur, L., Kalia, A., Mittal, A., Kaur, D., et al. (2021). Elucidating the morpho-physiological adaptations and molecular responses under long-term waterlogging stress in maize through gene expression analysis. *Plant Sci.* 304, 110823. doi: 10.1016/j.plantsci.2021.110823

Kim, D., Paggi, J. M., Park, C., Bennett, C., and Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* 37, 907–915. doi: 10.1038/s41587-019-0201-4

Klaas, M., Haiminen, N., Grant, J., Cormican, P., Finnan, J., Arojju, S. K., et al. (2019). Transcriptome characterization and differentially expressed genes under flooding and drought stress in the biomass grasses *Phalaris arundinacea* and *Dactylis glomerata*. *Ann. Bot.* 124, 717–730. doi: 10.1093/aob/mcz074

Kolahi, M., Faghani, E., Kazemian, M., Goldson-Barnaby, A., and Dodangi, S. (2021). Changes in secondary metabolites and fiber quality of cotton (*Gossypium hirsutum*) seed under consecutive water stress and in silico analysis of cellulose synthase and xyloglucan endotransglucosylase. *Physiol. Mol. Biol. Plant.* 27, 1837–1857. doi: 10.1007/s12298-021-01033-y

Komatsu, S., Han, C., Nanjo, Y., Altaf-Un-Nahar, M., Wang, K., He, D., et al. (2013). Label-free quantitative proteomic analysis of abscisic acid effect in early-stage soybean under flooding. *J. Proteome Res.* 12, 4769–4784. doi: 10.1021/pr4001898

Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. doi: 10.1093/bioinformatics/btt656

Li, J.-R., Liu, C.-C., Sun, C.-H., and Chen, Y.-T. (2018). Plant stress RNA-seq nexus: a stress-specific transcriptome database in plant cells. *BMC Genomics* 19, 966. doi: 10.1186/s12864-018-5367-5

Liu, Y., and Zhang, S. (2004). Phosphorylation of 1-Aminocyclopropane-1-Carboxylic acid synthase by MPK6, a stress-responsive mitogen-activated protein kinase, induces ethylene biosynthesis in *Arabidopsis*. *Plant Cell.* 16, 3386–3399. doi: 10.1105/tpc.104.026609

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

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

McCombie, W. R., McPherson, J. D., and Mardis, E. R. (2019). Next-generation sequencing technologies. *Cold Spring Harb. Perspect. Med.* 9, a036798. doi: 10.1101/cshperspect.a036798

Men, S., Chen, H., Chen, S., Zheng, S., Shen, X., Wang, C., et al. (2020). Effects of supplemental nitrogen application on physiological characteristics, dry matter and nitrogen accumulation of winter rapeseed (*Brassica napus* l.) under waterlogging stress. *Sci. Rep.* 10, 10201. doi: 10.1038/s41598-020-67260-7

Panozzo, A., Dal Cortivo, C., Ferrari, M., Vicelli, B., Varotto, S., and Vamerali, T. (2019). Morphological changes and expressions of *AOX1A*, *CYP81D8*, and putative *PFP* genes in a Large set of commercial maize hybrids under extreme waterlogging. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.00062

Peña-Castro, J. M., van Zanten, M., Lee, S. C., Patel, M. R., Voesenek, L. A. J. C., Fukao, T., et al. (2011). Expression of rice SUB1A and SUB1C transcription factors in *Arabidopsis* uncovers flowering inhibition as a submergence tolerance mechanism. *Plant J.* 67, 434–446. doi: 10.1111/j.1365-313X.2011.04605.x

Peng, Y., Zhang, X., Deng, Y., and Ma, X. (2008). Evaluation of genetic diversity in wild orchardgrass (*Dactylis glomerata* l.) based on AFLP markers. *Hereditas* 145, 174–181. doi: 10.1111/j.0018-0661.2008.02038.x

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

Phukan, U. J., Mishra, S., and Shukla, R. K. (2016). Waterlogging and submergence stress: affects and acclimation. *Crit. Rev. Biotechnol.* 36, 956–966. doi: 10.3109/07388551.2015.1064856

Pierik, R., Tholen, D., Poorter, H., Visser, E. J. W., and Voesenek, L. A. C. J. (2006). The janus face of ethylene: growth inhibition and stimulation. *Trends Plant Sci.* 11, 176–183. doi: 10.1016/j.tplants.2006.02.006

Qiao, D., Zhang, Y., Xiong, X., Li, M., Cai, K., Luo, H., et al. (2020). Transcriptome analysis on responses of orchardgrass (*Dactylis glomerata* l.) leaves to a short term flooding. *Hereditas* 157, 20. doi: 10.1186/s41065-020-00134-0

Qi, X., Li, Q., Ma, X., Qian, C., Wang, H., Ren, N., et al. (2019). Waterlogging-induced adventitious root formation in cucumber is regulated by ethylene and auxin through reactive oxygen species signalling. *Plant Cell Environ.* 42, 1458–1470. doi: 10.1111/pce.13504

Ren, J., Hu, J., Zhang, A., Ren, S., Jing, T., Wang, X., et al. (2021). The whole-genome and expression profile analysis of *WRKY* and *RGAs* in *Dactylis glomerata* showed that *DG6C02319.1* and dg *WRKYs* may cooperate in the immunity against rust. *PeerJ* 9, e11919. doi: 10.7717/peerj.11919

Sachdev, S., Ansari, S. A., Ansari, M. I., Fujita, M., and Hasanuzzaman, M. (2021). Abiotic stress and reactive oxygen species: Generation, signaling, and defense mechanisms. *Antioxidants* 10, 277. doi: 10.3390/antiox10020277

Sairam, R. K., Kumutha, D., Ezhilmathi, K., Deshmukh, P. S., and Srivastava, G. C. (2008). Physiology and biochemistry of waterlogging tolerance in plants. *Biol. Plant* 52, 401–412. doi: 10.1007/s10535-008-0084-6

Shuai, Y., Feng, G., Yang, Z., Liu, Q., Han, J., Xu, X., et al. (2022). Genome-wide identification of C2H2-type zinc finger gene family members and their expression during abiotic stress responses in orchardgrass (*Dactylis glomerata*). *Genome* 65, 189–203. doi: 10.1139/gen-2020-0201

Shu, K., Zhou, W., Chen, F., Luo, X., and Yang, W. (2018). Abscisic acid and gibberellins antagonistically mediate plant development and abiotic stress responses. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.00416

Steffens, B., Kovalev, A., Gorb, S. N., and Sauter, M. (2012). Emerging roots alter epidermal cell fate through mechanical and reactive oxygen species signaling. *Plant Cell.* 24, 3296–3306. doi: 10.1105/tpc.112.101790

Striker, G. G., and Colmer, T. D. (2016). Flooding tolerance of forage legumes. *J. Exp. Bot.* 68, 1851–1872. doi: 10.1093/jxb/erw239

Subbaiah, C. C., and Sachs, M. M. (2003). Molecular and cellular adaptations of maize to flooding stress. *Ann. Bot.* 91, 119–127. doi: 10.1093/aob/mcf210

Sun, M., Huang, D., Zhang, A., Khan, I., Yan, H., Wang, X., et al. (2020). Transcriptome analysis of heat stress and drought stress in pearl millet based on Pacbio full-length transcriptome sequencing. *BMC Plant Biol.* 20, 323. doi: 10.1186/s12870-020-02530-0

Tamang, B. G., Magliozzi, J. O., Maroof, M. A. S., and Fukao, T. (2014). Physiological and transcriptomic characterization of submergence and reoxygenation responses in soybean seedlings: Soybean responses to submergence and reoxygenation. *Plant Cell Environ.* 37, 2350–2365. doi: 10.1111/pce.12277

van der Knaap, E., and Kende, H. (1995). Identification of a gibberellin-induced gene in deepwater rice using differential display of mRNA. *Plant Mol. Biol.* 28, 589–592. doi: 10.1007/BF00020405

Verma, V., Ravindran, P., and Kumar, P. P. (2016). Plant hormone-mediated regulation of stress responses. *BMC Plant Biol.* 16, 86. doi: 10.1186/s12870-016-0771-y

Wang, X., Deng, Z., Zhang, W., Meng, Z., Chang, X., and Lv, M. (2017). Effect of waterlogging duration at different growth stages on the growth, yield and quality of cotton. *PloS One* 12, e0169029. doi: 10.1371/journal.pone.0169029

Wang, X., He, Y., Zhang, C., Tian, Y., Lei, X., Li, D., et al. (2021). Physiological and transcriptional responses of *Phalaris arundinacea* under waterlogging conditions. *J. Plant Physiol.* 261, 153428. doi: 10.1016/j.jplph.2021.153428

Xu, X., Chen, M., Ji, J., Xu, Q., Qi, X., and Chen, X. (2017). Comparative RNA-seq based transcriptome profiling of waterlogging response in cucumber hypocotyls reveals novel insights into the *de novo* adventitious root primordia initiation. *BMC Plant Biol.* 17, 129. doi: 10.1186/s12870-017-1081-8

Xu, X., Feng, G., Huang, L., Yang, Z., Liu, Q., Shuai, Y., et al. (2020). Genome-wide identification, structural analysis and expression profiles of GRAS gene family in orchardgrass. *Mol. Biol. Rep.* 47, 1845–1857. doi: 10.1007/s11033-020-05279-9

Yan, H., Zeng, B., Zhang, X., Cheng, L., Miller, S., and Huang, L. (2013). Screening orchardgrass (*Dactylis glomerata* l.) germplasm for high rust resistance and high summer survival rate in a subtropical climate. *Grassl. Sci.* 59, 205–210. doi: 10.1111/grs.12030

Yan, H., Zhang, Y., Zeng, B., Yin, G., Zhang, X., Ji, Y., et al. (2016). Genetic diversity and association of EST-SSR and SCoT markers with rust traits in orchardgrass (*Dactylis glomerata* l.). *Molecules* 21, 66. doi: 10.3390/molecules21010066

Yeung, E., van Veen, H., Vashisht, D., Sobral Paiva, A. L., Hummel, M., Rankenberg, T., et al. (2018). A stress recovery signaling network for enhanced flooding tolerance in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U. S. A.* 115, E6085–E6094. doi: 10.1073/pnas.1803841115

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

Zeng, B., Zhang, Y., Zhang, A., Qiao, D., Ren, J., Li, M., et al. (2020). Transcriptome profiling of two *Dactylis glomerata* l. cultivars with different tolerance in response to submergence stress. *Phytochemistry* 175, 112378. doi: 10.1016/j.phytochem.2020.112378

Zhang, S., and Klessig, D. F. (2001). MAPK cascades in plant defense signaling. *Trends Plant Sci.* 6, 520–527. doi: 10.1016/S1360-1385(01)02103-3

Zhou, W., Chen, F., Meng, Y., Chandrasekaran, U., Luo, X., Yang, W., et al. (2020). Plant waterlogging/flooding stress responses: From seed germination to maturation. *Plant Physiol. Biochem.* 148, 228–236. doi: 10.1016/j.plaphy.2020.01.020

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 31 January 2023
DOI 10.3389/fpls.2023.1100691

OPEN ACCESS

EDITED BY
Linkai Huang,
Sichuan Agricultural University, China

REVIEWED BY
Umesh K. Reddy,
West Virginia State University, United States
Hanna Bolibok-Bragoszewska,
Warsaw University of Life Sciences, Poland

*CORRESPONDENCE
Bingjiang Liu
ycbjliu@163.com
Xiong Wu
wutta2014@163.com

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

RECEIVED 17 November 2022
ACCEPTED 16 January 2023
PUBLISHED 31 January 2023

CITATION
Li Y, Huo Y, Yang Y, Wang Z, Sun Y, Liu B and Wu X (2023) Construction of a high-resolution genetic map and identification of single nucleotide polymorphism markers relevant to flower stalk height in onion. *Front. Plant Sci.* 14:1100691. doi: 10.3389/fpls.2023.1100691

Construction of a high-resolution genetic map and identification of single nucleotide polymorphism markers relevant to flower stalk height in onion

Yanwei Li, Yumeng Huo, Yanyan Yang, Zhenbao Wang, Yaling Sun, Bingjiang Liu* and Xiong Wu*

Key Laboratory for Biology of Greenhouse Vegetables of Shandong Province/National Center for Vegetable Improvement (Shandong Branch), Vegetable Research Institute, Shandong Academy of Agricultural Sciences, Jinan, China

Introduction: Onion (Allium cepa L., 2n=16) is an economically and nutritionally important vegetable crop worldwide. Construction of a high-resolution genetic map and map-based gene mining in onion have lagged behind other vegetable crops such as tomato and pepper.

Methods: In this study, we constructed a high-resolution genetic map of onion using 321 F2 individuals from a cross between two double haploid lines DH-1×DH-17 and employing specific length amplified fragment (SLAF)-seq technology. The genetic map containing 10,584 polymorphic SLAFs with 21,250 single nucleotide polymorphism (SNP) markers and 8 linkage groups was developed for onion, which spanned 928.32 cM, with an average distance of 0.09 cM between adjacent markers.

Results: Using this map, we carried out QTL mapping of Ms locus related to the male-fertile trait and reproduced previous mapping results, which proved that this map was of good quality. Then, four QTLs (located on LG2, LG5, and LG8) were detected for flower stalk height, explaining 26.60% of the phenotypic variance. Among them, we proposed that 20 SLAF markers (in three QTLs) of flower stalk height trait were effective favorable allelic variant markers associated with heterosis.

Discussion: Overall, the genetic map was structured using SLAF-seq based on DH lines, and it is the highest-quality and highest-resolution linkage map of onion to date. It lays a foundation for the fine mapping and candidate gene identification of flower stalk height, and provides new insights into the developmental genetic mechanisms in onion breeding.

KEYWORDS

double haploid (DH), SLAF sequencing, genetic map, QTL analysis, onion

Introduction

Onion (*Allium cepa* L.) is a diploid species (2n=16) that belongs to the genus *Allium*, family Amaryllidaceae, and it is an economically and nutritionally important vegetable crop worldwide (Khosa et al., 2016). Onion is widely distributed, grown throughout the North and South, and is one of the main domestic vegetables, but also one of the main export vegetable crops in China. According to the statistics of the Food and Agriculture Organization (FAO, 2020), China has the world's largest onion production area and yield, accounting for approximately 30% of the world's total area and yield. Onion is a biennial plant, and decreasing the long breeding cycle is desirable for efficient onion breeding. Therefore, many studies focus on the molecular markers of fertility identification related to actual breeding (King et al., 1998; Park et al., 2013).

Our team previously developed molecular markers that are co-segregating with the male sterility gene and restorer gene, such as cDNA marker WHR240 (Huo et al., 2012), SCAR markers DNF-566 and DNS-357 (Yang et al., 2013), and multiple PCR marker AcSKP1 (Huo et al., 2015). These markers have been verified in open-pollinated (OP) populations with different genetic backgrounds, and can be directly used to identify the genotype of nuclear male sterility locus in onion. However, research on genetic linkage mapping and markers in the onion requires a great deal of time, due to its huge genome size (approximately 15 Gb), large number of repetitive sequences, and high heterozygous genetic background (Labani and Elkington, 1987; King et al., 1998; Jakse et al., 2005; Finkers et al., 2021).

To date, progress on onion genomic and genetic maps has been reported as follows (King et al., 1998; Martin et al., 2005; Duangjit et al., 2013; Jo et al., 2017; Choi et al., 2020; Cho et al., 2021; Finkers et al., 2021). The first onion genetic map was constructed by King et al. (1998) using dominant markers [randomly amplified polymorphic DNA (RAPD) and amplified fragment length polymorphism (AFLP) markers], which consisted of 116 markers, distributed in 12 linkage groups with an average genetic distance of 9 cM. Subsequently, Martin et al. (2005) developed 100 new molecular markers based on the rice expressed sequence tags library with high similarity and integrated them into the onion genetic map reported above using the chromatid replacement line of scallion, constructing a genetic map of 14 linkage groups with a total length of 1,907 cM. Duangjit et al. (2013) constructed a genetic map covering 936 single nucleotide polymorphism (SNP) markers and 10 linkage groups by library analysis on the Roche 454 platform of the leaves, bulbs, roots, and flower buds from two inbred lines. Jo et al. (2017) developed 202 SNP markers using an F_2 population by genotyping by sequencing (GBS) to construct a genetic map with a genetic distance of 1,383 cM and an average marker distance of 8.08 cM. Choi et al. (2020) performed GBS and high-resolution melting (HRM) analyses on F_2 onion plants and constructed an onion genetic linkage map with 319 SNPs and 34 HRM markers, consisting of 8 linkage groups and covering 881.4 cM with an average marker interval of 2.5 cM. Through quantitative trait loci (QTL) analysis, a major QTL, *qAS7.1*, for onion anthocyanin synthesis and two significant QTLs, *qAC4.1* and *qAC4.2*, for anthocyanin content were identified. Subsequently, Cho et al. (2021) developed 652 molecular markers using the above same F_2 population for constructing genetic maps with increased resolution. Fujito et al. (2021) constructed a high-density linkage map in *Allium cepa* using 1,435 unigene markers obtained from the transcriptome information of a 96 F_2 plants (from a cross between the A. *cepa* DH lines). The first *de novo* genome sequence (14.9 Gb with an N50 of 464 Kb) of onion was assembled based on Illumina short read sequencing by Finkers et al. (2021), and covered approx. 91% of the expected genome size with 72.4% of the genome identified as repetitive sequences and 20% of the putative (retro) transposons. Of this, 2.4 Gb was ordered into eight linkage groups, which was a valuable resource for research and breeding. However, fine mapping of onion traits is difficult due to the low densities of the genetic markers.

Advances in genotyping-by-sequencing have provided an excellent opportunity to increase the resolution of QTL mapping by increasing the numbers of markers in recent years (Varshney et al., 2009). Specific length amplified fragment (SLAF) sequencing (SLAF-seq), a high-resolution strategy of large-scale *de novo* SNP discovery and genotyping, was first described by Sun et al. (2013). SLAF-seq has been proved to be an effective method for genetic map construction and has wide applications in cucumber (Wei et al., 2014; Zhu et al., 2016), cotton (Zhang et al., 2016; Ali et al., 2018), peanut (Hu et al., 2018), pea (Zheng et al., 2018), soybean (Zhang et al., 2020; Mendoza et al., 2021), and eggplant (Wei et al., 2020). Compared with GBS, SLAF-seq has the following advantages: I. SLAF library was prepared by the double enzyme protocol and double barcode system, while GBS by the single enzyme. The DNA fragments sheared by double enzyme digestion were more uniformly distributed in the genome and the flexible enzymes combinations made it easier to control the number of markers; II. the marker density of SLAF was of 50-80/Mb, while of GBS was 5-40/Mb (Elshire et al., 2011; Sun et al., 2013). Therefore, SLAF-seq technology could provide a high-resolution strategy for SNP discovery, which is more suitable for widespread use in large-scale genotyping study. Considering the number of markers, the quality of markers, number of samples (321 samples) and research funding (the huge and complex genome of onion), SLAF-seq is the optimal choice for large-scale molecular marker development and high-resolution linkage map construction, especially in onion for which no reference genome information is available.

In this study, we constructed an F_2 population from a cross between two double haploid lines, DH-1 and DH-17. Using the F_2 population containing 321 individuals and SLAF-seq technology, we constructed a high-resolution genetic map in onion with 21,250 SNP markers. To verify the quality of the map, we performed QTL mapping of a male fertile trait. Then, we identified the QTLs related to flower stalk height (FSH) trait and explored their QTL effects to reveal the genetic mechanism of the FSH of onion, initially explaining its heterosis mechanism.

Materials and methods

Plant materials and phenotypic evaluation

Two onion double haploid (DH) lines were used in this study as parents. DH-1, whose male-sterility locus was S (*msms*) and whose phenotype was sterile, was induced from the Japanese hybrid ATON (Takii Seed, Figure 1A) *via in vitro* gynogenesis (Liu et al., 2012).

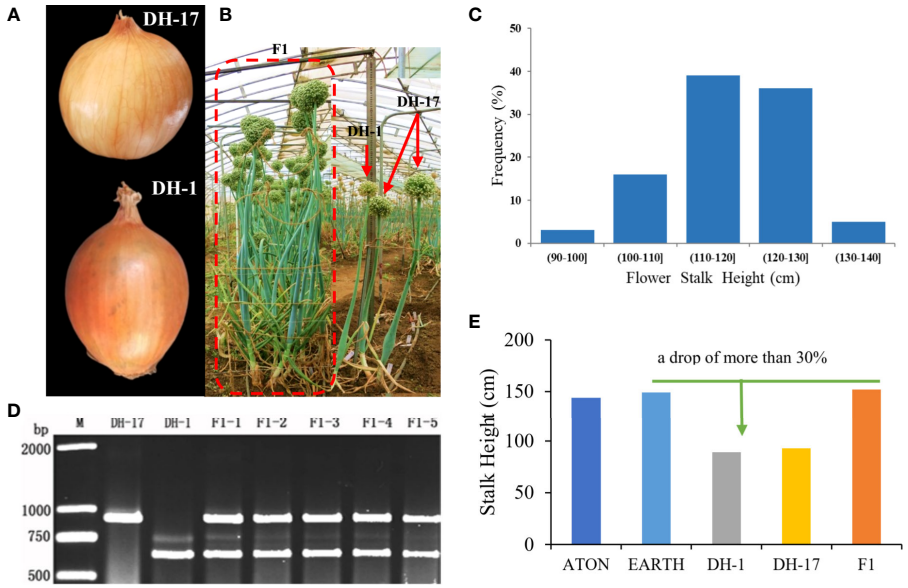


FIGURE 1
Phenotypes of the parental inbred lines and the F_1, F_2 individuals. **(A)** Paternal line DH-17 and maternal line DH-1; **(B)** Plant phenotypes of both parents and F_1 individuals; **(C)** Phenotypes evaluation of flower stalk height for the F_2 individuals; **(D)** Genotypes of both parents and partial F_1 individuals by the PCR markers DNF-566 and DNS-357; **(E)** Height of flower stalk for the original hybrids, DH lines and F_1.

Through flow cytometry analysis of leaf tissue, we preliminarily identified DH lines DH-1 and DH-17. To ensure that the DH lines were naturally doubled and no other sources such as somatic cells, we also performed PCR detection with *Ms*/*ms* co-segregated markers DNF-566 (F: 5′-TACAGATTTGTTTATCTTCTTCTTCT-3′; R: 5′-TTCATTTGTTAGGATGTACTCTTACC-3′) and DNS-357 (F: 5′-TCAGTATCAATAGAAGGAATCAC-3′; R: 5′-GTATACCATT GGTACTTGATGCA-3′, Yang et al., 2013). DH-17, which was for S (*MsMs*) with a fertile phenotype, was derived from the Japanese hybrid EARTH (Takii Seed, Figure 1A). F_1 hybrid came from a crossed between DH-1 (female parent) and DH-17 (male parent) were grown in greenhouses in early September of 2015; then a total of 321 F_2 individuals was derived from a self-crossing of F_1 were randomly selected to construct the genetic map. The parents and 321 F_2 individuals were grown in greenhouses in early September of 2017 at the Shandong Academy of Agricultural Sciences, Jinan, China. The genotype was identified for each individual by DNF-566 and DNS-357 markers (Yang et al., 2013), and individuals were identified as male sterile or fertile through inspection of pollen at mature anthers stage in June of 2019. The phenotypic data for FSH was collected in June of 2019. The FSH is defined as the height from the ground surface to the top of the stem, excluding the inflorescence.

DNA extraction

Young leaves from the 321 F_2 individuals and two parents were collected, and then stored in a -80°C freezer. The total genomic DNA was extracted from each sample according to the cetyltrimethyl ammonium bromide (CTAB) method (Murray and Thompson, 1980). The concentration and quality of total DNA were evaluated by electrophoresis on 1% agarose gels and an ND-1000 spectrophotometer (NanoDrop, Wilmington, DE, USA).

SLAF library construction and sequencing

SLAF-seq with some improvements was applied to genotype onion 321 F_2 individuals and two parents in 2018 (Sun et al., 2013). In the absence of reference genome, the dendrobium (https://www.ncbi.nlm.nih.gov/assembly/GCA_001605985.2), a closely related species, was finally selected as the reference genome for electronic enzymatic digestion prediction according to the onion genome size and GC content information. *Hae*III and *Eco*RV-HF (New England Biolabs, NEB, USA) were selected as endonucleases to digest the genomic DNA. A single nucleotide (A) overhang was added to the digested fragments using Klenow Fragment (3′→5′, exo-) and dATP at 37°C. Then, duplex tag-labeled sequencing adapters were ligated to the A-tailed fragments using T4 DNA ligase. PCR reactions were run using diluted restriction-ligation DNA, dNTPs, Q5® high-fidelity DNA polymerase, and universal PCR primers (F: 5′-AATGATACGGCGACCACCGA-3′, R: 5′-CAAGCAGAAGACGGCATACG-3′; PAGE-purified, Life Technologies). PCR products were purified using Agencourt AMPure XP beads (Beckman Coulter, High Wycombe, UK), and pooled samples were separated by 2% agarose gel electrophoresis. The digested fragments ranging from 414-444 base pairs (with indexes and adaptors) in size were excised and purified using a QIAquick gel extraction kit (Qiagen), and then, pair-end sequencing (100 bp) of the gel-purified products was performed on an Illumina HiSeq 2500 system (Illumina, Inc., San Diego, CA, USA) according to the manufacturer's recommendations.

SLAF-seq data analysis and genotyping

SLAF-seq data were assessed by the software developed by Sun et al. (2013), and SLAF marker identification and genotyping were performed according to procedures described by Sun et al. (2013).

Firstly, after filtering out low-quality reads (quality score < 20e), raw reads were sorted to each progeny according to duplex barcode sequences; then trimming the barcodes and the terminal 5-bp positions from each high-quality reads, clean reads were clustered by similarity above 90%. Sequences clustered together were defined as one SLAF locus (Zhang et al., 2015). Secondly, single nucleotide polymorphism (SNP) loci of each SLAF locus were detected between parents. SLAFs with more than 5 SNPs were filtered out, and SNP loci of each SLAF locus were detected between parents by using the GATK software kit (McKenna et al., 2010). Alleles of each SLAF locus were then defined according to parental reads with sequence depth >319.07-fold, while for each progeny the reads with sequence depth >29.12-fold were used to define alleles. For diploid species, one SLAF locus can contain at most 4 genotypes, so SLAF loci with more than 4 alleles were defined as repetitive SLAFs and discarded. Only SLAFs with 2 to 4 alleles were identified as polymorphic and considered potential markers. All polymorphism SLAFs loci were genotyped with consistency in the parental and progeny SNP loci. Because the mapping populations were derived from two onion DH parents with a genotype of aa or bb, only the polymorphic SLAFs were analyzed on the basis of the F_2 population type (aa × bb).

Genotype scoring was performed by a Bayesian approach to ensure the genotyping quality (Sun et al., 2013). *A posteriori* conditional probability was calculated using the coverage of each allele and the number of SNP, and genotyping quality score translated from the probability was used to select qualified markers. Then, low-quality markers were counted and the worse marker or individual were deleted. High-quality SLAF markers for the genetic mapping were filtered by the following criteria: (1) contained more than 5 SNPs; (2) markers with more than 70% missing data, (i.e., complete degree of markers in each progeny below 70% discarded); (3) markers with significant separation distortion ($p > 0.05$) by using chi-square test; (4) parents' sequence depth of less than 10×. The remaining high-quality SLAFs were subsequently used for constructing the linkage map.

Linkage map construction

To ensure efficient construction of the high-resolution and high-quality map, the SLAF markers were correctly ordered by the HighMap strategy, and genotyping errors were corrected by the SMOOTH algorithm (van Os et al., 2005; Liu et al., 2014). First, the recombination rate and maximized logarithm of odds (MLOD) value between two points were calculated. Then, the interval of the MLOD value was determined by arranging the MLOD values of those markers from the smallest to largest, and the highest MLOD values among the markers were in the same linkage group (LG). Finally, all SLAF markers (MLOD > 5) were divided into 8 onion linkage groups (LGs; each linkage group was considered as a chromosome) to construct the genetic map (Zhang et al., 2015). Taking LG as a unit, the regression algorithm/minimum spanning tree (MST) algorithm was selected to estimate the genetic distance between adjacent markers, and the optimal linear order of markers in LGs was obtained by regression sequencing analysis (Liu et al., 2014). Then, the k-nearest neighbor algorithm was applied to impute missing genotypes (Huang et al., 2011), and the Kosambi mapping function was used for estimating map distances (cM) (Kosambi, 1943).

QTL mapping

The composite internal mapping (CIM) analysis within the R/QTL software was applied for QTL mapping of plant male fertility (Broman and Sen, 2009). The logarithm of odds (LOD) threshold was determined using the 1000 permutations test (PT, $P<0.05$) to evaluate the statistical significance of each QTL. To ensure that major and minor effect QTLs could be identified, different LOD scores were adopted. First, an LOD threshold corresponding to 0.99 confidence was considered, followed by 0.95 and 0.90. Then, if no QTL interval was detected, the PT result was not considered, and the threshold was manually lowered to 3.0, 2.5, and 2.0 successively. QTLs were normally named according to their linkage group locations and trait names.

Results

SLAF-seq data and SLAF marker analysis

SLAF sequencing of the 321 F_2 population, including the two parents, generated a total of 857.58 GB clean data comprising 4.29 G reads of approximately 100-bp paired-end sequencing (Table 1). Among these reads, 95.70% achieved or exceeded a quality score of 30 (Q30, indicating a 0.1% chance of an error and 99.9% confidence), and the guanine-cytosine (GC) content was 40.54%. There were 131,741,300 reads from male parent DH-17 and 119,231,892 reads from female parent DH-1, and the average number of reads from the F_2 population was 11,567,116. Specifically, the reads of rice (*Oryza sativa* L. var. Japonica, genome size ≈ 382 Mb), which were used as a control to estimate the validity of library construction, were 6,852,482 with 95.46% Q30 bases and 45.60% GC (Table 1).

The number of SLAFs (based on the SLAF development process, with the reads at the same location being defined as a SLAF tag) in the male, female parent, and F_2 population were 533,049, 458,885, and 323,656, respectively. The average sequencing depths of the male, the

TABLE 1 Summary of the SLAF-sequencing data.

Sample ID	Total Reads (Mb)	Total Bases (Gb)	Q30 (%)	GC (%)
DH-17	131.74	26.35	95.56	41.41
DH-1	119.23	23.85	95.86	39.88
F_2-progeny	11.57	2.31	95.70	40.53
Total	4,287.90	857.58	95.70	40.54

10.3389/fpls.2023.1100691

female parent, and F_2 were 90.84×, 100.19×, and 13.68×, respectively. Among the 692,676 SLAFs developed, 155,067 were polymorphic SLAFs with a rate of 22.39% (Supplementary Table S1). Of those 155,067 polymorphic SLAF tags, 93,539 were successfully encoded and grouped into eight segregation patterns (ab×cd, ef×eg, ab×cc, cc×ab, hk×hk, lm×ll, nn×np and aa×bb; Figure 2; Supplementary Table S2). Due to the F_2 population obtained by selfing the F_1 of a cross between two parents with the homozygous genotype of aa or bb, only the F_2 plants with the aa×bb segregation pattern were valid polymorphic SLAFs. There were 77,766 SLAFs accounting for 83.14% of the total encoded markers that were categorized with a segregation pattern of aa×bb in this study (Figure 2).

A high-resolution genetic linkage map

To produce a high-quality map, low-quality SLAFs with a parental sequencing depth < 10×, SNP number > 5, segregation distortion $p < 0.05$, and MOLD between neighboring markers < 5 were discarded. To construct a genetic map, 10,584 high-quality SLAFs including 21,250 SNPs were retained (Table 2). For the map, the total distance of the genetic map was 928.32 cM, and the average interval between markers was 0.09 cM (Table 2). The average depths of the SNPs for female, male parent, and F_2 population were 313.74×, 324.40×, and 29.12×, respectively.

Based on those analyses, we constructed a new high-resolution onion genetic map containing 8 linkage groups (LGs, Figure 3A). The linear arrangements and the genetic distances of markers in each LG were analyzed using the HighMap software. The average integrity of the mapped markers was 96.01%, which adequately ensured the accuracy of the genotyping. The genetic distances in the map spanned 96.24 cM (LG7) to 140.43 cM (LG3), with the number of SLAF markers in each LG ranging from 198 (LG6) to 1,704 (LG2) and SNP marker numbers ranging from 352 (LG6) to 3,448 (LG8, Table 2). The average number of SLAF and SNP markers in all LGs were 1,323 and 2,656, respectively, and markers spanned an average length of 116.04 cM. Satisfactory uniformity in the distribution was obtained, with gaps < 5 cM constituting 99.94% of the total LGs, and the average max gap was 2.91 cM (Table 2).

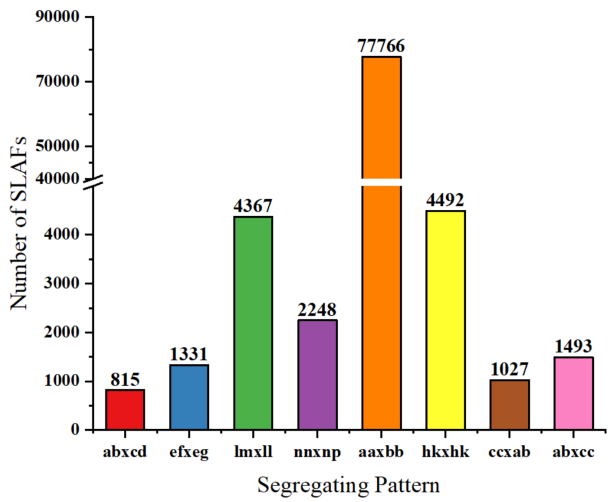


FIGURE 2
Genotype distribution of SLAF markers in onion. The Y-axis indicates the number of SLAF markers, x-axis indicates the segregation patterns of markers.

Male-fertile phenotypic characterization of the genetic population

Performing PCR to detect the DH lines genotype with *Ms*/*ms* co-segregated markers DNF-566 and DNS-357 (Yang et al., 2013), the results showed that the *Ms* locus of DH-1 was S *(msms)*, which was a sterile phenotype; the genotype of DH-17 was S (*MsMs*) with a fertile phenotype (Figure 1D). According to plant genotype, the F_2 population was divided into the following three groups (Supplementary Excel S1): *MsMs* of 75 individuals, *Msms* of 166 individuals, and *msms* of 80 individuals, with a 1:2:1 segregation ratio. By chi-square test, the separation ratio of fertile (241 individuals) and sterile plants (81 individuals) was 3:1. These results suggest that the onion *Ms*/*ms* trait is controlled by one dominant loci, which was consistent with previous conclusions (Jones and Clarke, 1943).

QTL mapping reproducibility in the F_2 population

Before conducting QTL mapping for onion agronomic traits on the basis of the high-resolution genetic map, we performed QTL

TABLE 2 Basic information for high-resolution genetic map of onion.

LG ID	SLAF Num.	SNP Num.	Total Dis. (cM)	Average Dis. (cM)	Gaps<=5	Max Gap (cM)
LG1	1,185	2,336	115.29	0.10	100%	3.60
LG2	1,704	3,396	131.33	0.08	99.94%	6.27
LG3	1,689	3,369	140.43	0.08	100%	2.18
LG4	1,425	2,922	102.40	0.07	100%	2.92
LG5	1,209	2,368	100.99	0.08	100%	1.35
LG6	198	352	103.59	0.53	100%	2.59
LG7	1,500	3,059	96.24	0.06	100%	2.12
LG8	1,674	3,448	138.05	0.08	100%	2.27

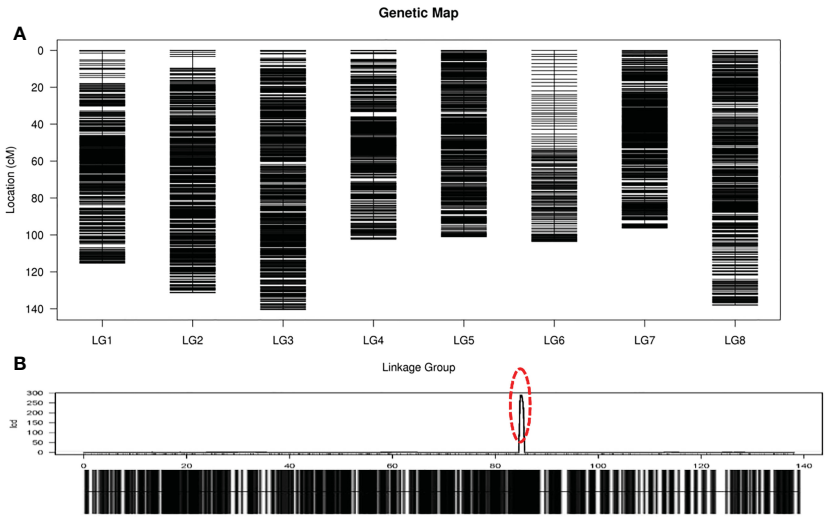


FIGURE 3
High-resolution genetic map of onion. **(A)** Genetic linkage map of onion. The black bars on each linkage group represent SLAF markers on the map; **(B)** Locations of male-fertile trait on genetic map. The red circle indicates the site of male-fertile trait.

mapping of a male fertile trait to verify the quality of the map. Only one QTL locus for male-fertile trait, designated *MS8.1*, was mapped on LG8 (Figure 3B), which could explain 110.31% of the observed phenotypic variation (PVE). The genetic distance interval of *qtl-MS8.1* ranged from 84.83 to 85.42 cM, and 11 SLAF markers were uncovered within this locus. The additive and dominant effects for *qtl-MS8.1* were -0.49 and 0.49, respectively, and the LOD threshold value of the QTL was 241.64. Based on the study of the DNF-566 and RNS-357 markers by Yang et al. (2013), (Kim et al., 2014; Kim et al., 2015) developed one co-dominant marker jnurf13 [an insertion deletion (indel) marker] in 2014 and proposed in 2015 that the bulb onion DNA mismatch repair (*AcPMS1*) gene was the most probable candidate for regulating fertility restoration. Huo et al. (2015) identified the putative S-phase kinase-associated protein 1 (SKP1) protein in linkage disequilibrium with *Ms* locus. In conclusion, it is likely that the male-fertility QTL locus located in this study is the same as that mapped by King et al. (1998). The SLAF marker distribution results in this study are presented in Figure 4. Detailed sequence information and the positions of SNPs uncovered within SLAFs are presented in Supplementary Table S3.

QTLs mapping of the FSH trait

The FSH phenotypic data for all onion materials (including two original induced hybrids, two parents, and F_1 and F_2 populations) are presented in Figures 1B, C. The FSH of double haploids DH-1 and DH-17 was lower than FSH of the original inducing hybrids from which the parental DH lines were developed by more than 30%. The FSH of F_1 plants from cross with DH-1×DH-17 was restored to FSH of the original inducing hybrids, and the height of the F_1 plants were highly uniform and consistent (Figure 1E). The height of the flower stalks in this study were measured using a ruler. FSH ranged in a spectrum from short stalk to tall stalk in the segregating populations and was thus assessed on a 1-5 height scale. There were a tiny percentage of F_2 individuals within the range of 90.01-100 cm (3%, 11 individuals) and 130.01-140 cm (5%, 17 individuals), then 16% in the range of 100.01-110 cm (51 individuals), and the majority percentage within the range of 110.01-120 cm (39%, 125 individuals) and 120.01-130 cm (36%, 117 individuals). Analyses of the FSH data showed a normal distribution in the F_2 population (Figure 1C; Supplementary Excel S1). According to the genotype of the onion F_2 population from the SLAF-seq data, 49 SLAF markers (located in 4 QTLs) were detected and considered as favorable allelic variant (FAV) markers for flower stalk height. The LGs, starting positions, ending positions, SLAF markers, LODs, additive values, dominance values, and PVEs are shown in Figures 4, 5, Table 3 and Supplementary Table S4.

To avoid false-positive sites caused by repetitive sequences or splicing errors, and further narrow down the range of FAV markers, in our opinion, those SLAF markers that were beneficial to the flower stalk height of heterozygous progeny could be considered as effective FAV markers and should be preferentially selected for analysis. Since there was no significant difference in FSH between the parents used to construct the population, we only calculated the proportion of heterozygote progeny carrying FAV markers for the perspective of heterosis effect. The results showed that the proportion of heterozygous individuals carrying 20 markers were positively correlated with the height of flower stalk in F_2 population were selected from the all 49 SLAF markers (Supplementary Table S5). After evaluating the proportion of F_2 heterozygous individuals that carried the FAV marker by two measures, average proportion and total proportion in F_2 population, we proposed that there were 20 SLAF markers (in 3 QTLs) that could be considered as effective FAV markers of FSH (Figure 6A; Table 4). Those 3 QTLs related to the FSH trait were distributed on linkage groups LG2 and LG8 (Figures 4, 5; Table 3). The genetic distance interval of the QTLs ranged from 0 to 0.59 cM. The PVE by the 3 QTLs ranged from 5.17-7.19%, and the accumulated PVE was 19.90%. The LOD value of the 20 SLAF markers ranged from 5.17-6.84 (Table 5). The number of SNP loci

10.3389/fpls.2023.1100691

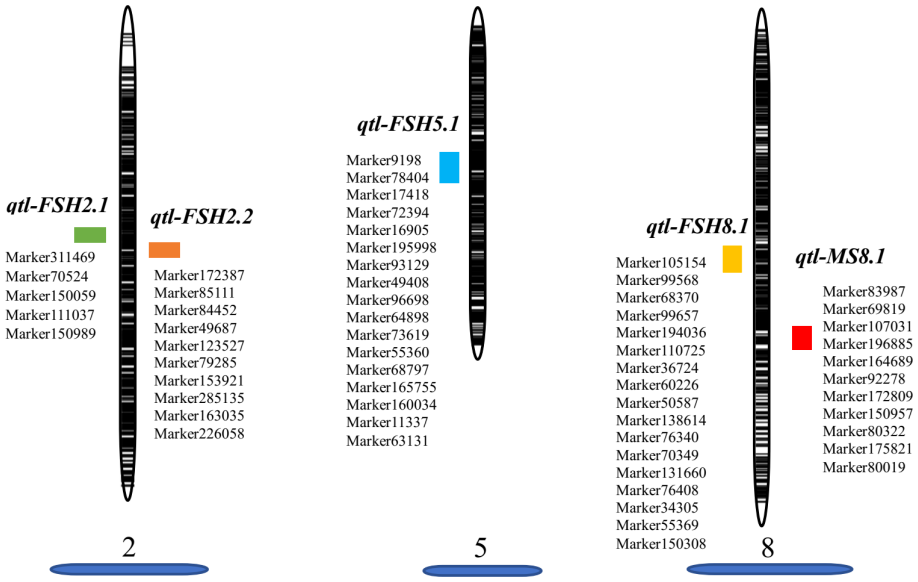


FIGURE 4
Graphic view of the distribution of markers associated with related traits on onion LGs 2, 5, 8.

within each SLAF varied from 1 to 5. Detailed sequence information and alignment positions for all markers are presented in Table 3 and Figure 4. Furthermore, we analyzed the number of effective FAVs carried by 17 F_2 individuals ranging in height from 130 to 140 cm. As shown in Figure 6B, 11 out of 17 individual plants contained more than 50% of effective FAV markers, and one of the finest F_2-98 individuals carried 20 effective FAV markers of FSH.

Discussion

Construction of a high-resolution genetic map based on onion double haploid lines

To analyze an onion that is a cross-pollinated plant, with a large, repeated genome (Khosa et al., 2016; Finkers et al., 2021), DH lines

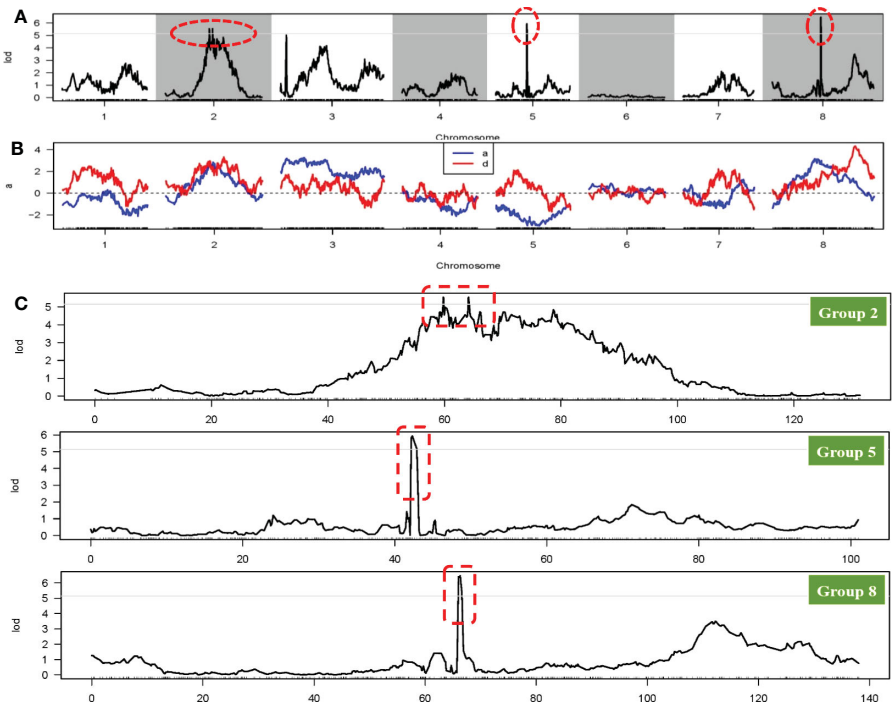


FIGURE 5
QTL locations of flower stalk height trait using the genetic map. **(A)** QTL analysis of the flower stalk height trait on eight linkage groups; **(B)** Phenotypic contribution rate for variation of the flower stalk height trait; **(C)** LOD scores for the QTLs of the flower stalk height trait on the 2th, 5th and 8th LG. The red circle or box all indicate the QTL regions related to flower stalk height trait.

TABLE 3 Analysis of quantitative trait loci (QTL) for flower stalk height in onion.

QTL	LG ID	Linkage Map Pos. (cM)	Interval Size (cM)	LOD	Max LOD	SLAF Num.	ADD	DOM	PVE%
qtl-FSH2.1	LG2	59.82-59.83	0	5.15	5.55	5	2.80	2.98	7.19
qtl-FSH2.2	LG2	64.11-64.27	0.15	5.15	5.55	10	2.86	2.48	6.65
qtl-FSH5.1	LG5	42.16-42.90	0.74	5.15	5.94	17	-2.99	0.83	6.45
qtl-FSH8.1	LG8	66.01-66.60	0.59	5.15	6.48	17	3.00	1.60	6.28
Total	–	–	–	–		49		7.89	26.57

LOD, logarithm of odds; ADD, additive; DOM, dominance; PVE, phenotypic variance explained.

are the most ideal first choice for mapping segregated populations. A sterile yellow doubled haploid line (DH-1) and a fertile yellow doubled haploid line (DH-17) were used to produce an F_2 mapping population (containing 321 individuals) of onions in this study. The SLAF-seq strategy we applied was based on high-throughput sequencing technologies and developed as a simplified genome sequencing method (Sun et al., 2013). The advantages of using SLAF-seq technology rather than other available methods (such as RAD, GBS) is that SLAF adopted of a flexible double enzyme protocol and duplex barcode system, which could ensure the filtered molecular markers more uniformly distributed on the genome (Baird et al., 2008; Elshire et al., 2011; Sun et al., 2013), and this method uses high depth sequencing to define *de novo* genotypes without reference genome sequence and reference SNP, which is suitable for constructing onion genetic map.

In recent years, SLAF-seq has been widely applied in constructing high-resolution genetic maps not only for crops with reference genomes (Dai et al., 2018; Wang et al., 2018; Wei et al., 2020; Tao et al., 2022), but also for plants without a reference genome, such as orchardgrass (Zhao et al., 2016) and pea (Zheng et al., 2018). Detailed information on *Allium cepa* DHCU066619 genome has not yet been published, and its huge genome (14.9 Gb) consists of complex highly repetitive regions with (retro) transposons (Finkers et al., 2021). Therefore, we used the SLAF-seq technique to build a high-quality genetic map of onion without reference genome in 2018. To date, although several onion genetic maps have been constructed (Duangjit et al., 2013; Jo et al., 2017; Choi et al., 2020; Ryu et al., 2021), we used 10,584 polymorphic markers with the complete degree of 96.01% to construct onion genetic linkage map of 8 linkage groups (LGs) with a total length of 928.32 cM. The average marker distributed in the 8 LGs was 1,323, but only 198 markers were developed from LG6. Referred to the relevant literature on the onion genetic map and found that markers developed for each LG were about 50-200 (Choi et al., 2020; Cho et al., 2021; Fujito et al., 2021), and took into account large number of repetitive sequences existing in the onion genome (Finkers et al., 2021), we believed that the quality of this genetic map was not affected, although the number of markers located in the LG6 was low. As a whole, the linkage map carries the largest number of marker loci among all onion linkage maps constructed to date. Combined with the high depth of parental sequencing, low missing data rate information, average interval between markers (0.09 cM), and max gap value (Table 2; Figure 3A; Supplementary Table S6), the linkage map constructed in our study has the highest quality and resolution among onion linkage maps. Although total genetic map length was very small comparatively to the genome size of onion, our high-resolution genetic map should not only be useful for accurate and reliable mapping for QTLs, but also has important reference significance for onion DH-17 genome assembly. However, the chromosomal identity of LGs has not been known yet in this study, we would try to construct a physical chromosome assignment map based transcriptome unigene marker (Fujito et al., 2021) or carry out onion DH-17 genome assembly to resolve this issue in the future work.

Implications for onion cytoplasmic male-fertility marker

Because onion is a biennial plant with a long breeding cycle, it is crucial to develop effective molecular markers for fertility

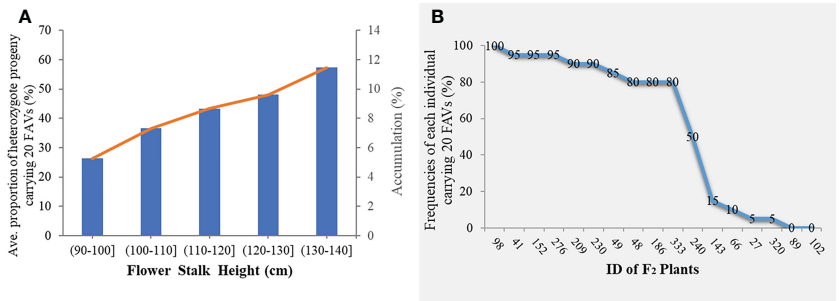


FIGURE 6
Statistical analysis of effective favorable allele varieties (FAVs) in the F_2 population. **(A)** Average (accumulation) proportion of heterozygote progeny carrying effective FAV markers in the F_2 population. The orange curve indicates the trend curve between cumulative proportion and flower stalk height; **(B)** Frequencies of the 20 FAV loci in the F_2 individuals with the height range from 130-140 cm.

TABLE 4 Twenty effective favorable allelic variant SLAFs associated with flower stalk height trait and heterosis in the F_2 population.

LG ID	SLAF ID	SNP Site	SNP in DH-17	SNP in DH-1	Proportion of heterozygote progeny in (90-100 cm]	Proportion of heterozygote progeny in (100-110 cm]	Proportion of heterozygote progeny in (110-120 cm]	Proportion of heterozygote progeny in (120-130 cm]	Proportion of heterozygote progeny in (130-140 cm]
LG2	Marker150989	103	G	A	0.27	0.33	0.42	0.47	0.47
LG2	Marker285135	157	C	T	0.27	0.33	0.38	0.44	0.35
LG2	Marker226058	73	T	A	0.27	0.25	0.39	0.53	0.53
LG8	Marker105154	31	C	A	0.36	0.37	0.43	0.43	0.65
LG8	Marker99568	114	G	A	0.27	0.39	0.46	0.47	0.53
LG8	Marker68370	68	A	G	0.27	0.39	0.42	0.47	0.65
LG8	Marker99657	175	G	T	0.27	0.35	0.46	0.46	0.59
LG8	Marker194036	54	A	G	0.18	0.29	0.43	0.41	0.47
LG8	Marker110725	107	G	A	0.27	0.33	0.41	0.44	0.65
LG8	Marker36724	137	T	C	0.27	0.39	0.47	0.56	0.59
LG8	Marker60226	101	A	G	0.27	0.39	0.45	0.44	0.47
LG8	Marker50587	44	A	G	0.27	0.37	0.42	0.52	0.65
LG8	Marker138614	81	T	C	0.18	0.37	0.43	0.49	0.65
LG8	Marker76340	119	T	C	0.18	0.35	0.46	0.43	0.53
LG8	Marker70349	90	T	C	0.27	0.39	0.43	0.54	0.65
LG8	Marker131660	47	A	G	0.27	0.39	0.42	0.48	0.59
LG8	Marker76408	112	A	T	0.27	0.41	0.44	0.51	0.65
LG8	Marker34305	27	A	G	0.27	0.41	0.47	0.52	0.65
LG8	Marker55369	114	G	A	0.36	0.41	0.49	0.54	0.65
LG8	Marker150308	172	C	A	0.18	0.35	0.38	0.48	0.53

identification that can be used during the practical breeding of onion. Cytoplasmic male-sterility (CMS), a maternally inherited trait characterized by the inability to produce viable pollen grains in higher plants, has been widely applied for the production of F_1 hybrid seeds in agronomic crops such as onion, broccoli, cabbage, and radish (Havey, 2004; Bang et al., 2011). There are two main sources of CMS systems in onion: CMS-S and CMS-T. For the CMS-S system, male fertility is controlled by a single nuclear gene and a cytoplasmic factor (Jones and Clarke, 1943), and there are three loci involved in fertility restoration in the CMS-T system (Schweisguth, 1973). Therefore, the CMS-S was more widely used for onion hybrid cultivars because of its stability in various environments (Havey, 1995; Havey, 2000). To date, there has been a great deal of research performed on molecular markers for CMS-S fertility identification that are relevant for onion breeding (Sato, 1998; Engelke et al., 2003; Kim et al., 2009; Yang et al., 2013; Huo et al., 2015).

In our study, the fertility and sterility of pollen of the F_2 population segregated in a 3:1 ratio. The segregation ratio for the *MsMs: Msms: msms* individuals that are marked by co-segregated markers DNF-566 and DNS-357 (Yang et al., 2013) was 1:2:1, which was consistent with a 3:1 Mendelian segregation ratio. Coincidentally, only one QTL locus (with higher LOD and PVE values, Figure 3B) related to male-fertile trait was detected on this genetic map, which may be the same *qtl* that was mapped by King et al. (1998). To summarize, we were able to reproduce the mapping of the onion male-fertility QTL in our F_2 population, which indicated the higher quality of our constructed genetic linkage map, and thus, it can be used to efficiently locate novel QTLs for important traits of onion.

Implications for onion QTL preliminary mapping of FSH

Suitable plant height is an important agronomic trait for elite hybrid seeds, especially for gramineal crops where the top grain yield organ (grain) is supported by a stem, such as sorghum, rice, and wheat (Zhou et al., 2012; Li et al., 2015; Jiang et al., 2017), and plays a vital role in the final yield formation. For onion, FSH directly affects lodging resistance and is closely related to seed yield and quality. In this study, double haploids DH-1 and DH-17 originated from commercial F_1 hybrids, and their FSHs were lower than those of the original inducing hybrids by more than 30% (Figure 1E). The FSH of the F_1 hybrids crossed with DH-1×DH-17 was restored to that of the original hybrids, and the heights of the hybrids were highly uniform and consistent. This phenomenon belongs in the category of heterosis. After taking into account the phenotypic analysis of FSH

TABLE 5 Detailed information of effective favorable allelic variant SLAF markers for the flower stalk height.

LG ID	SLAF Pos.	SLAF ID	SNP Num.	LOD	ADD	DOM	PVE%
LG2	59.822	Marker150989	1	5.55	2.80	2.98	7.19
LG2	64.263	Marker285135	2	5.17	2.83	2.39	6.43
LG2	64.263	Marker226058	4	5.17	2.83	2.39	6.43
LG8	66.010	Marker105154	2	6.42	2.91	1.75	6.07
LG8	66.010	Marker99568	3	6.42	2.91	1.75	6.07
LG8	66.010	Marker68370	1	6.42	2.91	1.75	6.07
LG8	66.010	Marker99657	2	6.42	2.91	1.75	6.07
LG8	66.010	Marker194036	3	6.42	2.91	1.75	6.07
LG8	66.157	Marker110725	1	6.38	2.90	1.77	6.08
LG8	66.304	Marker36724	2	6.48	3.00	1.60	6.28
LG8	66.304	Marker60226	1	6.48	3.00	1.60	6.28
LG8	66.304	Marker50587	2	6.48	3.00	1.60	6.28
LG8	66.304	Marker138614	2	6.48	3.00	1.60	6.28
LG8	66.304	Marker76340	2	6.48	3.00	1.60	6.28
LG8	66.304	Marker70349	5	6.48	3.00	1.60	6.28
LG8	66.304	Marker131660	5	6.48	3.00	1.60	6.28
LG8	66.304	Marker76408	1	6.48	3.00	1.60	6.28
LG8	66.304	Marker34305	2	6.48	3.00	1.60	6.28
LG8	66.304	Marker55369	3	6.48	3.00	1.60	6.28
LG8	66.598	Marker150308	3	5.40	2.63	1.85	5.17

and four QTLs (located on LG2, LG5, LG8), only 26.60% of the phenotypic variation of the F_2 population could be explained (Table 3). Consistent with this result, Baldwin et al. (2014) revealed genome regions on three chromosomes associated with bolting using a low-density linkage map. We inferred that the height of the onion flower stalk may be controlled by a heterosis effect between multiple dwarfing genes. Analysis of effective FAVs of onion FSH indicated that the genetic essence of heterosis may be the comprehensive effect of various genes containing a dominance effect, additive effect, epistatic effect, complementary and superdominance effect, and interaction effect that varied with various traits or genes.

Based on the high-resolution genetic map constructed by DH lines, we conducted QTL mapping for FSH. The implications are: (1) to propose the hypothesis of heterosis of FSH in onion; (2) to develop KASP markers for SNPs or genes relevant to FSH, which will lay a foundation for the construction of a molecular marker-assisted breeding system, and provide new insights into the developmental genetic mechanisms of FSH.

Conclusion

In conclusion, our study provided a high- resolution genetic map based on double haploid lines with SLAF-seq in onion. We conducted QTL mapping of onion male fertile trait to verify that the genetic map had the higher-quality and could be efficient to locate novel QTLs for important traits of onion. We also conducted QTL mapping for the flower stalk height trait and 20 out of 49 SLAF markers (in 3/4 QTLs) were considered as effective FAV markers, which were suitable to develop KASP markers for SNPs or genes relevant to flower stalk height in the future. These findings would lay a foundation for the construction of molecular marker-assisted breeding system and provide new insights into the developmental genetic mechanisms of flower stalk height in onion.

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

YL and XW conceived and designed the research. YL performed the main data analysis, manuscript drafting and revision. XW, ZW, and YH performed the experiments. BL and YH constructed the population, collected phenotypes and SLAF-seq. ZW, YY, and YS

carried out the field work. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by China Agriculture Research System (CARS-24-A-10) and Agricultural Science and Technology Innovation Project of SAAS (CXGC2022E08).

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.2023.1100691/full#supplementary-material

References

Ali, I., Teng, Z., Bai, Y., Yang, Q., Hao, Y., Hou, J., et al. (2018). A high density SLAF-SNP genetic map and QTL detection for fibre quality traits in gossypium hirsutum. *BMC Genomics* 19, 879–897. doi: 10.1186/s12864-018-5294-5

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., et al. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PloS One* 3, e3376. doi: 10.1371/journal.pone.0003376

Baldwin, S., Revanna, R., Pither-Joyce, M., Shaw, M., Wright, K., Thomson, S., et al. (2014). Genetic analyses of bolting in bulb onion (*Allium cepa* l.). *Theor. Appl. Genet.* 127, 535–547. doi: 10.1007/s00122-013-2232-4

Bang, S. W., Tsusui, K., Shim, S., and Kaneko, Y. (2011). Production and characterization of the novel CMS line of radish (*Raphanus sativus*) carrying brassica maurorum cytoplasm. *Plant Breed.* 130, 410–412. doi: 10.1111/J.1439-0523

Broman, K. W., and Sen, S. (2009). A guide to QTL mapping with r/qtl. *J. Stat. Softw.* 66, 658–659. doi: 10.1007/978-0-387-92125-9

Choi, Y., Kim, S., and Lee, J. (2020). Construction of an onion (*Allium cepa* l.) genetic linkage map using genotyping-by-sequencing analysis with a reference gene set and identification of QTLs controlling anthocyanin synthesis and content. *Plants* 9, 616–632. doi: 10.3390/plants9050616

Cho, Y., Kim, B., Lee, J., and Kim, S. (2021). Construction of a high-resolution linkage map and chromosomal localization of the loci determining major qualitative traits in onion (*Allium cepa* l.). *Euphytica* 217, 17–30. doi: 10.1007/s10681-020-02746-z

Dai, F., Wang, X. L., Zhang, X. Q., Chen, Z. H., Nevo, E., Jin, G. L., et al. (2018). Assembly and analysis of a qingke reference genome demonstrate its close genetic relation to modern cultivated barley. *Plant Biotechnol. J.* 16, 760–770. doi: 10.1111/pbi.12826

Duangjit, J., Bohanec, B., Chan, A. P., Town, C. D., and Havey, M. J. (2013). Transcriptome sequencing to produce SNP-based genetic maps of onion. *Theor. Appl. Genet.* 126, 2093–2101. doi: 10.1007/s00122-013-2121-x

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., et al. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PloS One* 6, e19379. doi: 10.1371/journal.pone.0019379

Engelke, T., Terefe, D., and Tatlioglu, T. A. (2003). PCR-based marker system monitoring CMS-(S), CMS-(T) and (N)-cytoplasm in the onion (*Allium cepa* l.). *Theor. Appl. Genet.* 107, 162–167. doi: 10.1007/s00122-003-1230-3

FAO. (2020). *Available at: https://www.fao.org/statistics/en/*.

Finkers, R., van Kaauwen, M., Ament, K., Burger-Meijer, K., Egging, R., Huits, H., et al. (2021). Insights from the first genome assembly of onion (*Allium cepa*). *G3* 11, jkab243. doi: 10.1093/g3journal/jkab243

Fujito, S., Akyol, T. Y., Mukae, T., Wako, T., Yamashita, K. I., Tsukazaki, H., et al. (2021). Construction of a high-density linkage map and graphical representation of the arrangement of transcriptome-based unigene markers on the chromosomes of onion, *Allium cepa* l. *BMC Genomics* 22, 481–489. doi: 10.1186/s12864-021-07803-y

Havey, M. J. (1995). Identification of cytoplasms using the polymerase chain reaction to aid in the extraction of maintainer lines from open-pollinated populations of onion. *Theor. Appl. Genet.* 90, 263–268. doi: 10.1007/BF00222212

Havey, M. J. (2000). Diversity among male-sterility-inducing and male-fertile cytoplasms of onion. *Theor. Appl. Genet.* 101, 778–782. doi: 10.1007/s001220051543

Havey, M. J. (2004). The use of cytoplasmic male sterility for hybrid seed production. *Mol. Biol. Biotechnol. Plant organelles.*, 234–623. doi: 10.1007/978-1-4020-3166-3-23

Huang, X., Zhao, Y., Wei, X., Li, C., Wang, A., Zhao, Q., et al. (2011). Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat. Genet.* 44, 32–39. doi: 10.1038/ng

Huo, Y. M., Liu, B. J., Yang, Y. Y., Miao, J., Gao, L. M., Kong, S. P., et al. (2015). *Acskp1*, a multiplex PCR-based co-dominant marker in complete linkage disequilibrium with the male-fertility restoration (*ms*) locus, and its application in open-pollinated populations of onion. *Euphytica* 204, 711–722. doi: 10.1007/s10681-015-1374-7

Huo, Y. M., Miao, J., Liu, B. J., Yang, Y. Y., Zhang, Y. H., and Wu, X. (2012). The expression of pectin methylesterase in onion flower buds is associated with the dominant male-fertility restoration allele. *Plant Breed.* 131, 211–216. doi: 10.1111/J.1439-0523.2011.01907.X

Hu, X. H., Zhang, S. Z., Miao, H. R., Cui, F. G., Shen, Y., Yang, W. Q., et al. (2018). High-density genetic map construction and identification of QTLs controlling oleic and linoleic acid in peanut using SLAF-seq and SSRs. *Sci. Rep.* 8, 5479–5489. doi: 10.1038/s41598-018-23873-7

Jakse, J., Martin, W., Mccallum, J., and Havey, M. (2005). Single nucleotide polymorphisms, indels, and simple sequence repeats for onion cultivar identification. *J. Amer. Soc Hortic. Sci.* 130, 912–917. doi: 10.21273/jashs.130.6.912

Jiang, Y., Schmidt, R. H., Zhao, Y., and Reif, J. C. (2017). A quantitative genetic framework highlights the role of epistatic effects for grain-yield heterosis in bread wheat. *Nat. Genet.* 49, 1741–1746. doi: 10.1038/ng.3974

Jones, H. A., and Clarke, A. E. (1943). Inheritance of male sterility in the onion and the production of hybrid seed. *Proc. Am. Soc Hortic. Sci.* 43, 189–194.

Jo, J., Purushotham, P. M., Han, K., Lee, H. R., Nah, G., and Kang, B. C. (2017). Development of a genetic map for onion (*Allium cepa* l.) using reference-free genotyping-by-sequencing and SNP assays. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.01606

Khosa, J. S., McCallum, J., Dhatt, A. S., and Macknight, R. C. (2016). Enhancing onion breeding using molecular tools. *Plant Breed.* 135, 9–20. doi: 10.1111/pbr.12330

Kim, S., Kim, M. S., Kim, Y. M., Yeom, S. I., Cheong, K., Kim, K. T., et al. (2014). Integrative structural annotation of *de novo* RNA-seq provides an accurate reference gene set of the enormous genome of the onion (*Allium cepa* l.). *DNA Res.* 22, 19–27. doi: 10.1093/dnares/dsu035

Kim, S., Kim, C. W., Park, M., and Choi, D. (2015). Identification of candidate genes associated with fertility restoration of cytoplasmic male-sterility in onion (*Allium cepa* l.) using a combination of bulked segregant analysis and RNA-seq. *Theor. Appl. Genet.* 128, 2289–2299. doi: 10.1007/s00122-015-2584-z

Kim, S., Lee, E. T., Cho, D. Y., Han, T., Bang, H., Patil, B. S., et al. (2009). Identification of a novel chimeric gene, *orf725*, and its use in development of a molecular marker for distinguishing among three cytoplasm types in onion (*Allium cepa* l.). *Theor. Appl. Genet.* 118, 433–441. doi: 10.1007/s00122-008-0909-x

King, J. J., Bradeen, J. M., Bark, O., McCallum, J. A., and Havey, M. J. (1998). A low-density genetic map of onion reveals a role for tandem duplication in the evolution of an extremely large diploid genome. *Theor. Appl. Genet.* 96, 52–56. doi: 10.1007/s001220050708

Kosambi, D. D. (1943). The estimation of map distances from recombination values. *Ann. Hum. Genet.* 12, 172–175. doi: 10.1111/j.1469-1809.1943.tb02321.x

Labani, R. M., and Elkington, T. T. (1987). Nuclear DNA variation in the genus allium l. (Liliaceae). *Heredity* 59, 119–128. doi: 10.1038/hdy.1987.103

Li, X., Li, X., Fridman, E., Tesso, T. T., and Yu, J. (2015). Dissecting repulsion linkage in the dwarfing gene *Dw3* region for sorghum plant height provides insights into heterosis. *Proc. Natl. Acad. Sci.* 112, 11823–11828. doi: 10.1073/pnas.1509229112

Liu, D., Ma, C., Hong, W., Huang, L., Liu, M., Liu, H., et al. (2014). Construction and analysis of high-density linkage map using high-throughput sequencing data. *PloS One* 9, e98855. doi: 10.1371/journal.pone.0098855

Liu, B. J., Miao, J., Huo, Y. M., Yang, Y. Y., Xu, K., and Wu, X. (2012). Haploid plant induction *via in vitro* gynogenesis and plant regeneration in onion (*allium cepa* l.). *Acta Hortic. Sin.* 39, 2265–2270. doi:

Martin, W. J., McCallum, J., Shigyo, M., Jakse, J., Kuhl, J. C., Yamane, N., et al. (2005). Genetic mapping of expressed sequences in onion and in silico comparisons with rice show scant colinearity. *Mol. Genet. Genomics* 274, 197–204. doi: 10.1007/s00438-005-0007-6

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., et al. (2010). The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303. doi: 10.1101/gr.107524.110

Mendoza, A., Moussa, A. A., Zhang, Q., Qu, J., Du, Y., Anwari, G., et al. (2021). Genome-wide association study of root and shoot related traits in spring soybean (*Glycine max* l.) at seedling stages using SLAF-seq. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.568995

Murray, M., and Thompson, W. F. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic. Acids Res.* 8, 4321–4326. doi: 10.1093/nar/8.19.4321

Park, J., Bang, H., Cho, D. Y., Yoon, M. K., Patil, B. S., and Kim, S. (2013). Construction of high-resolution linkage map of the *Ms* locus, a restorer-of-fertility gene in onion (*Allium cepa* l.). *Euphytica* 192, 267–278. doi: 10.1007/s10681-012-0851-5

Ryu, H., Kim, S., Lee, H. E., Han, J., and Lee, J. (2021). Development of single nucleotide polymorphism markers linked to quantitative trait loci controlling anthocyanin content in red bulb onion (*Allium cepa* l.). *Korean J. Breed. Sci.* 53, 116–126. doi: 10.9787/KJBS.2021.53.2.116

Sato, Y. (1998). PCR amplification of CMS-specific mitochondrial nucleotide sequences to identify cytoplasmic genotypes of onion (*Allium cepa* l.). *Theor. Appl. Genet.* 96, 367–370. doi: 10.1007/s001220050750

Schweisguth, B. (1973). Étude d'un nouveau type de stérilité male chez l'oignon, allium cepa l. *Ann. Amélior Plant* 23, 221–233.

Sun, X., Liu, D., Zhang, X., Li, W., Liu, H., Hong, W., et al. (2013). SLAF-seq: An efficient method of large-scale *de novo* SNP discovery and genotyping using high-throughput sequencing. *PloS One* 8, e58700. doi: 10.1371/journal.pone.0058700

Tao, J., Li, S., Wang, Q., Yuan, Y., Ma, J., Xu, M., et al. (2022). Construction of a high-density genetic map based on specific-locus amplified fragment sequencing and identification of loci controlling anthocyanin pigmentation in yunnan red radish. *Hortic. Res.* 9, uhab031. doi: 10.1093/hr/uhab031

van Os, H., Stam, P., Visser, R. G., and van Eck, H. J. (2005). SMOOTH: A statistical method for successful removal of genotyping errors from high-density genetic linkage data. *Theor. Appl. Genet.* 112, 187–194. doi: 10.1007/s00122-005-0124-y

Varshney, R. K., Nayak, S. N., May, G. D., and Jackson, S. A. (2009). Next-generation sequencing technologies and their implications for crop genetics and breeding. *Trends Biotechnol.* 27, 522–530. doi: 10.1016/j.tibtech.2009.05.006

Wang, X. L., Chen, Z. H., Yang, C. Y., Zhang, X. L., Jin, G. L., Chen, G., et al. (2018). Genomic adaptation to drought in wild barley is driven by edaphic natural selection at the tabigha evolution slope. *Proc. Natl. Acad. Sci.* 115, 5223–5228. doi: 10.1073/pnas.1721749115

Wei, Q., Wang, W., Hu, T., Hu, H., Wang, J., and Bao, C. (2020). Construction of a SNP-based genetic map using SLAF-seq and QTL analysis of morphological traits in eggplant. *Front. Genet.* 11. doi: 10.3389/fgene.2020.00178

Wei, Q., Wang, Y., Qin, X., Zhang, Y., Zhang, Z., Wang, J., et al. (2014). An SNP-based saturated genetic map and QTL analysis of fruit-related traits in cucumber using specific-length amplified fragment (SLAF) sequencing. *BMC Genomics* 15, 1158–1168. doi: 10.1186/1471-2164-15-1158

Yang, Y. Y., Huo, Y. M., Miao, J., Liu, B. J., Kong, S. P., Gao, L. M., et al. (2013). Identification of two SCAR markers co-segregated with the dominant *Ms* and recessive *ms* alleles in onion (*Allium cepa* l.). *Euphytica* 190, 267–277. doi: 10.1007/s10681-012-0842-6

Zhang, J., Zhang, Q., Cheng, T., Yang, W., Pan, H., Zhong, J., et al. (2015). High-density genetic map construction and identification of a locus controlling weeping trait in an ornamental woody plant (Prunus mume sieb. et zucc). *DNA. Res.* 22, 183–191. doi: 10.1093/dnares/dsv003

Zhang, S., Li, B., Chen, Y., Shaibu, A. S., Zheng, H., and Sun, J. (2020). Molecular-assisted distinctness and uniformity testing using SLAF-sequencing approach in soybean. *Genes* 11, 175–193. doi: 10.3390/genes11020175

Zhang, Z., Shang, H., Shi, Y., Huang, L., Li, J., Ge, Q., et al. (2016). Construction of a high-density genetic map by specific locus amplified fragment sequencing (SLAF-seq) and its application to quantitative trait loci (QTL) analysis for boll weight in upland cotton (*Gossypium hirsutum.*). *BMC Plant Biol.* 16, 79. doi: 10.1186/s12870-016-0741-4

Zhao, X., Huang, L., Zhang, X., Wang, J., Yan, D., Li, J., et al. (2016). Construction of high-density genetic linkage map and identification of flowering-time QTLs in orchardgrass using SSRs and SLAF-seq. *Sci. Rep.* 6, 29345–29355. doi: 10.1038/srep2934

Zheng, Y., Xu, F., Li, Q., Wang, G., Liu, N., Gong, Y., et al. (2018). QTL mapping combined with bulked segregant analysis identify SNP markers linked to leaf shape traits in *Pisum sativum* using SLAF sequencing. *Front. Genet.* 9. doi: 10.3389/fgene.2018.00615

Zhou, G., Chen, Y., Yao, W., Zhang, C., Xie, W., Hua, J., et al. (2012). Genetic composition of yield heterosis in an elite rice hybrid. *Proc. Natl. Acad. Sci.* 109, 15847–15852. doi: 10.1073/pnas.1214141109

Zhu, W. Y., Huang, L., Chen, L., Yang, J. T., Wu, J. N., Qu, M. L., et al. (2016). A high-density genetic linkage map for cucumber (*Cucumis sativus* l.): Based on specific length amplified fragment (SLAF) sequencing and QTL analysis of fruit traits in cucumber. *Front. Plant Sci.* 7. doi: 10.3389/fpls.2016.00437

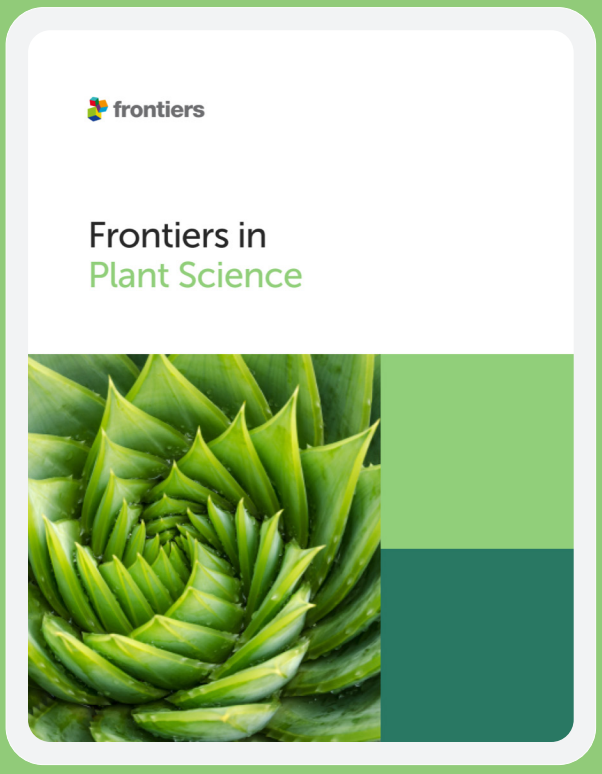
Frontiers in
Plant Science

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