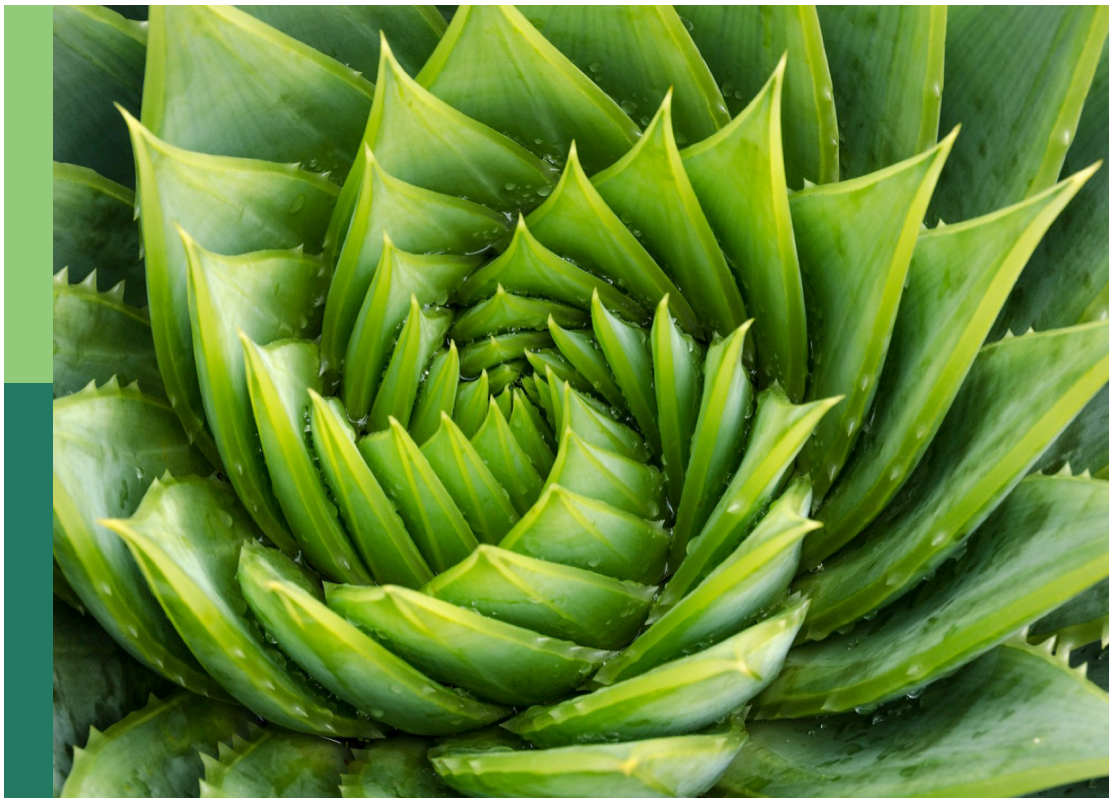
frontiers | Research Topics

Advances in environmental stress biology and important agronomic traits improvement in non-staple crops

Edited by
Meng Kou, Kaixuan Zhang, Jian Sun
and Ivan Kreft

Published in
Frontiers in Plant Science



ISSN 1664-8714
ISBN 978-2-8325-4906-3
DOI 10.3389/978-2-8325-4906-3

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

Advances in environmental stress biology and important agronomic traits improvement in non-staple crops

Topic editors

Meng Kou — Xuzhou Institute of Agricultural Sciences in Jiangsu Xuhuai District, China

Kaixuan Zhang — Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, China

Jian Sun — Jiangsu Normal University, China

Ivan Kreft — University of Ljubljana, Slovenia

Citation

Kou, M., Zhang, K., Sun, J., Kreft, I., eds. (2024). *Advances in environmental stress biology and important agronomic traits improvement in non-staple crops*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-4906-3

Table of contents

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 10 January 2023
DOI 10.3389/fpls.2022.1119049

OPEN ACCESS

EDITED BY
Meng Kou,
Sweet Potato Research Institute (CAAS), China

REVIEWED BY
Lixiang Wang,
Shanxi Agricultural University, China
Zaijun Yang,
China West Normal University, China

*CORRESPONDENCE
Qi Wu
jerviswuqi@126.com

SPECIALTY SECTION
This article was submitted to Crop and Product Physiology, a section of the journal Frontiers in Plant Science

RECEIVED 08 December 2022
ACCEPTED 19 December 2022
PUBLISHED 10 January 2023

CITATION
Wu Q, Bai X, Nie M, Li L, Luo Y, Fan Y, Liu C, Ye X and Zou L (2023) Genome-wide identification and expression analysis disclose the pivotal *PHOSPHATIDYLETHANOLAMINE BINDING PROTEIN* members that may be utilized for yield improvement of *Chenopodium quinoa*. *Front. Plant Sci.* 13:1119049. doi: 10.3389/fpls.2022.1119049

Genome-wide identification and expression analysis disclose the pivotal *PHOSPHATIDYLETHANOLAMINE BINDING PROTEIN* members that may be utilized for yield improvement of *Chenopodium quinoa*

Qi Wu[1,2,3]*, Xue Bai[1,2,3], Mengping Nie[1,2,3], Li Li[1,2,3], Yiming Luo[1,2,3], Yu Fan[1,2,3], Changying Liu[1,2,3], Xueling Ye[1,2,3] and Liang Zou[1,2,3]

[1]Key Laboratory of Coarse Cereal Processing, Ministry of Agriculture and Rural Affairs, Chengdu University, Chengdu, Sichuan, China, [2]Sichuan Engineering & Technology Research Center of Coarse Cereal Industralization, Chengdu University, Chengdu, Sichuan, China, [3]School of Food and Biological Engineering, Chengdu University, Chengdu, Sichuan, China

Quinoa (*Chenopodium quinoa*) is a prospective orphan crop that needs yield improvement. Previous studies indicate *PHOSPHATIDYLETHANOLAMINE BINDING PROTEIN* (*PEBP*) family genes are highly associated with the key agronomic traits of crops. Characterizing the pivotal *PEBP* genes will speed up the domestication and yield improvement of quinoa. Previous investigations on *PEBP* genes of *Chenopodium* species indicated that, the *PEBP* genes, despite in the same subclade, may have experienced functional diversification. Especially, the allotetraploidy (AABB) and numerous segmental duplications and chromosomal rearrangements in quinoa make it more difficult to understand the functions of *PEBP* genes. More recently, 6 quinoa *FT* subfamily genes were predicted to be related to flowering of quinoa. However, investigation on the whole *PEBP* family members is still lacking. In this study, we obtained 23 *PEBP* genes, including 5 *MFT*, 11 *FTL* and 7 *TFL* genes. We found 7 orthologous gene pairs, from sub-genome A and sub-genome B, respectively, showing collinearities with sugar beet. Evolution analysis on *PEBP* genes of two quinoa sub-genomes, sugar beet and relatives of diploid ancestors indicated that, the reasons for gene duplication events varied and 4 tandem duplications are the major reason for *PEBP* family expansion. Tissue-specific expression analysis suggested that expression patterns are mostly differing between orthologous gene pairs. Analysis on gene expressions at 6 stages suggested the possible positive roles of *CqFTL1*/*CqFTL2*, *CqFTL5*, *CqFTL8*, *CqFTL6*/*CqFTL9* and *CqTFL6*/*CqTFL7*, and negative roles of *CqTFL1*/*CqTFL2*/*CqTFL3*,

CqTFL4/*CqTFL5* in inflorescence branching. Expression analysis in ABA-treated seed, in combination with the *cis*-acting element distribution analysis, indicated that *CqMFT2*, *CqMFT3* and *CqMFT4* may regulate seed germination *via* ABA signaling. Observations on responses to night break and photoperiod changes highlighted the roles of *CqFTL5* and *CqFTL8* under short day, and *CqFTL6* under long day for quinoa flowering. Further, co-expression network analysis indicated that 64 transcription factors may act upstream of *CqFTL5* and *CqFTL8* to regulate flowering. Together, this study will help us identify the pivotal *PEBP* genes that may be utilized for quinoa breeding in future.

KEYWORDS

FT, TFL, MFT, flowering, yield, Chenopodium quinoa

Background

Quinoa (*Chenopodium quinoa*) is a prospective orphan crop due to the nutritional components in its seed and high tolerance to various abiotic stresses that could ensure its growth in marginal lands (Lopez-Marques et al., 2020). Nowadays, because of the increasing global demands for quinoa, yield improvement per unit area and expansion of cultivation area should be achieved. As demonstrated in many studies, flowering time regulation not only is tightly associated with inflorescence morphology and yield, but also is related to the adaption in higher latitudes (Song et al., 2015). Thus, identifying the genes highly associated with flowering and inflorescence morphology is essential for fast domestication and yield improvement of quinoa.

Transition from vegetive to reproductive stage is influenced by both internal and environmental cues (Song et al., 2015). Photoperiod is the major pathway regulating floral transition. In favorable season, day length signal is transmitted to the major floral integrator *FLOWERING LOCUS T* (*FT*) (Taoka et al., 2013; Tsuji et al., 2013; Holt et al., 2014; Kyozuka et al., 2014; Song et al., 2015). *FT* encodes for florigen protein that moves through the phloem from leaves to the shoot apical meristem (SAM), where partners with bZIP transcription factor FD and 14-3-3 proteins to form floral activating complex (FAC), which in turn promotes expressions of the floral identity genes *APETALA1* (*AP1*) and *LEAFY* (*LFY*) (Taoka et al., 2013; Tsuji et al., 2013; Holt et al., 2014; Kyozuka et al., 2014; Song et al., 2015). *FT* belongs to *PHOSPHATIDYLETHANOLAMINE BINDING PROTEIN* (*PEBP*) family. *PEBP* family contains three subfamilies, namely *FT*, *TERMINAL FLOWER 1* (*TFL1*) and *MOTHER OF FT AND TFL1* (*MFT*), all of which exert important roles in plant growth and development (Karlgren et al., 2011; Wickland and Hanzawa, 2015). In contrast to the floral-inducing role of *FT*, *TFL1* functions as a floral repressor and maintains vegetative growth by repressing *AP1* and *LFY* (Hanano and Goto, 2011; Kaneko-Suzuki et al., 2018; Goretti et al., 2020). The spatio-temporal expression of *FT*/*TFL1* affects flowering time, inflorescence architecture and final yield.

Up to date, a good many evidences have demonstrated the important roles of *FT*-, *TFL1*- and *MFT-like* genes in agronomic traits regulation. Overexpression of the rice (*Oryza saltiva*) *TFL1* homologs, *RICE CENTRORADIALIS-like 1/2* (*RCN1/2*), led to delayed heading and generation of higher-order panicle branches (Nakagawa et al., 2002), whereas RNA interfering (RNAi) of rice *TFL1* resulted in advanced heading and reduced branches (Liu et al., 2013). The maize (*Zea mays*) plants ectopic expressing *ZEA CENTRORADIALIS 2/4/5* (*ZCN2/4/5*) produced increased lateral branches (Danilevskaya et al., 2010). The wheat (*Triticum aestivum*) plants overexpressing *TaTFL1-2D* generated increased spikelets (Wang et al., 2017). In the background of mutant *self-pruning* (*sp*, homolog of *TFL1*), the tomato (*Solanum lycopersicum*) plants with heterozygous *single flower truss* (*sft*/+) (*sft*, homolog of *FT*) produced heterosis and dramatically increased number of fruits (Krieger et al., 2010; Jiang et al., 2013). A natural variant of the *TFL1* homolog *CENTRORADIALIS* (*HvCEN*) contributed to the spring growth habit of cultivated barley (*Hordeum vulgare*) (Comadran et al., 2012). *HvCEN* interacts with *HvFT3* to control spikelet initiation and grain number of barley (Mulki et al., 2018; Bi et al., 2019). In addition to the classical florigen function, the rice *FT* homolog *HEADING DATE 3A* (*HD3A*) also regulates shoot branching (Tamaki et al., 2007; Tsuji et al., 2015). In potato (*Solanum tuberosum*), the *FT* homolog *StSP6A* is required for tuberization transition (Navarro et al., 2011). *MFT*, ancestral gene of *FT* and *TFL1*, is a key regulator in seed germination. *OsMFT2* negatively regulates seed germination *via* ABA pathway and the knock-out mutant exhibited pre-harvest sprouting phenotype (Song et al., 2020). The wheat *MFT* was revealed to be tightly linked with the seed dormancy QTL *QPhs.ocs-3A.1* (Nakamura et al., 2011). The conserved functions of *PEBP* homologs were found among

different plant species. However, due to the differing expression patterns and multiple copies of *PEBP* family, the *PEBP* genes, even in the same subfamily, may have distinct functions. For example, in soybean (*Glycine max*), *FT5a* is involved in post-flowering stem growth other than the flowering-inducing role shared with *FT2a* (Takeshima et al., 2019). In rice, the *FT* homolog *HD3A* is the principal flowering regulator under short day, while the other *FT* homolog *RICE FLOWERING LOCUS T1* (*RFT1*) mainly functions under long day (Komiya et al., 2008; Komiya et al., 2009).

As the importance of *PEBP*s in yield regulation has been demonstrated in many crops, characterizing and manipulating of the significant *PEBP* genes in quinoa will help improve the yield and cultivation area of quinoa. Recently, a few studies on *FT* subclade genes were carried out in Amaranthaceae family. In sugar beet (*Beta vulgaris*), two *FT* paralogs, *BvFT1* and *BvFT2*, were identified. However, they harbor antagonistic functions in flowering (Pin et al., 2010). In *C. rubrum*, two *FT* homologs, *CrFTL1* and *CrFTL2*, were identified, in which only *CrFTL1*, rather than *CrFTL2*, was up-regulated during floral transition (Chab et al., 2008). In *Chenopodium* species, the expression of *FTL1* was correlated with floral induction in *C. suecicum* and short-day type *C. ficifolium*, whereas was not correlated with that in long-day type *C. ficifolium* (Storchova et al., 2019). These results suggested that those *PEBP*s of *Chenopodium* species, despite in the same clade, may have experienced functional diversification. Quinoa, an allotetraploid (AABB), had experienced numerous chromosomal rearrangements and chromosome fusions (Jarvis et al., 2017), which may increase the functional complexity of *PEBP* family. More recently, the quinoa *FT* subfamily genes were identified and their evolutionary relationships with other plants were assessed (Štorchová, 2020). Expressions of six *CqFT* genes were compared in early- and late-flowering quinoa accessions (Patiranage et al., 2021). Haplotypes of two *CqFT* genes were predicted to be associated with the photoperiod sensitivity of quinoa (Patiranage et al., 2021). These studies have improved our understanding of the plausible functions of quinoa *FT-like* subfamily members. Yet, functions of quinoa *TFL1-like* and *MFT-like* subfamily genes remain mysterious. Complete investigations into gene duplication, gene structure, *cis*-acting element in the promotor, and more important, the transcriptional changes of the whole *PEBP* family in various progresses are still required to further elucidate their specific roles. In this study, we analyzed the phylogenetic relationships, collinearities and duplication events between *PEBP* genes of sub-genome A and sub-genome B and relatives of diploid ancestors, and assessed their expression changes during inflorescence development and seed germination and detected their responses to night break and photoperiods, and further performed co-expression network analysis to predict the transcription factors upstream of *PEBP* genes. The results of this study will help identify the pivotal *PEBP* genes governing flowering time, inflorescence branching and seed germination, which may be utilized for quinoa breeding in future.

Materials and methods

Identification, and phylogenetic analysis of PEBP homologs from different plants

To identify the PEBP genes in various plant species, we performed BLASTP (E-value<1.0e-15) search against genomes of *Arabidopsis thaliana*, *Spinacia oleracea*, *Oryza sativa* and *Beta vulgaris* in Phytozome V13 (https://phytozome-next.jgi.doe.gov), and *C. pallidicaule* and *C. suecicum* in *Chenopodium* database (https://www.cbrc.kaust.edu.sa/chenopodiumdb/), using FT (AT1G65480.1), TFL (AT5G03840.1) and MFT (AT1G18100.1) protein sequences of Arabidopsis as the queries. Then those homologs were aligned with the Hidden Markov Model (HMM) profile of PEBP domain (PF01161) using Pfam search (E-value<1.0e-10) (Finn et al., 2016) to ensure those homologs habor a PEBP domain. Multiple sequence alignment of PEBP sequences from various species was performed using CLASTALW (Thompson et al., 2002). Phylogenetic tree was constructed using MEGA 11.0 (Tamura et al., 2021) based on the Neighbor-Joining method (Dohm et al., 2013) with a bootstrap value of 1000.

Chromosomal location, gene structure and conserved motif analysis

The General Feature Format (GFF) file and chromosome-scale genome sequence of quinoa were downloaded from Phytozome V13 database (https://phytozome-next.jgi.doe.gov/info/Cquinoa_v1_0). Based on these two files, the physical location on chromosomes, and intron and exon structures of PEBP genes were determined and visualized by using the two programs of Gene Location Visualize and Gene Structure View in TBtools (Chen et al., 2020). Multiple Em for Motif Elicitation (MEME) program (https://meme-suite.org/meme/tools/meme) was used to identify the conserved motifs in PEBP proteins setting the maximum motif count of 8. The motif analysis results were illustrated using the Gene Structure View program in TBtools (Chen et al., 2020).

Collinearity and gene duplication events analysis between *PEBP* genes of different species

The GFF file and chromosome-scale genome sequence of sugar beet (*Beta vulgaris*) were downloaded from Phytozome V13 database (https://phytozome-next.jgi.doe.gov/info/

Bvulgaris_EL10_1_0). The GFF files and chromosome-scale genomes of quinoa and sugar beet were input into the Multiple collinear scanning tool (MCScanX) in TBtools (Chen et al., 2020) and the collinearity between quinoa and sugar beet genomes was analyzed. Then the collinearities between *PEBP* genes from quinoa and sugar beet were determined. Meanwhile, the duplication events among quinoa *PEBP* genes were analyzed. The results were visualized by using the Advanced Circos program in TBtools (Chen et al., 2020).

Distribution of *cis*-acting element in quinoa *PEBP* gene promoters

The 2000bp sequences upstream of the translation start site of *PEBP* genes were recognized as promoter regions. The promoter sequences were uploaded to the Plant *Cis*-Acting Regulatory Elements (PlantCARE) database to search *cis*-acting elements. The physical distribution of various *cis*-acting elements was displayed using Simple BioSequence Viewer in TBtools (Chen et al., 2020).

Expression analysis of PEBPs in various tissues and biological events

For tissue-specific expression pattern analysis, we used the RNA-seq data from Zou et al. (Zou et al., 2017). Raw data of various quinoa tissues, including 1-week-old seedling, stem, leaf, inflorescence from 6-week-old plants and dry seed, was downloaded from Sequence Read Archive (https://www.ncbi.nlm.nih.gov/sra) of the BioProject PRJNA394651. The fragments per kilobase of transcript per million fragments mapped (FPKM) value of each gene was calculated with the previously described methods (Wu et al., 2019; Wu et al., 2020; Wu et al., 2021). For gene expression analysis in quinoa inflorescences at six developmental stages, we investigated the FPKM values of *PEBP* genes in our published transcriptome data (Wu et al., 2019). YP1, YP2, YP3 and YP4 represent the young non-branching panicles, whereas P1 and P2 stand for the panicles of elder branching stages. Raw data generated from six-stage inflorescence samples was downloaded from SRA (BioProject PRJNA511332), and was further processed with the bioinformatic pipeline methods described before (Wu et al., 2020). To know gene expression changes of *PEBP* genes during seed germination, we used the published transcriptome data of our laboratory (Wu et al., 2020). We compared the expression levels (FPKM value) of *PEBP* genes in control and Abscisic acid (ABA)-treated seeds 5h and 15h after imbibition. To investigate *PEBP* gene expression changes in response to night-break (NB) treatment, we used our published RNA-seq data generated from leaf samples treated by short-day and NB conditions (Wu et al., 2021). To investigate the *PEBP* gene diurnal expression pattern changes in response to photoperiod, two-week-old quinoa seedlings were cultivated under short-day (8h/16h, light/dark) and long-day (16h/8h, light/dark) conditions for sampling and RNA-seq analysis. The top two fully-expanded leaves of 3~5 plants were harvested at the time point of 17:00, 20:00, 23:00, 02:00, 05:00, 08:00, 11:00 and 14:00, with two biological replicates. All the leaf samples were subjected to RNA extraction, high-throughput sequencing and data analysis as previously described (Wu et al., 2019; Wu et al., 2021).

Prediction of the co-expressed transcription factors with *PEBP*s

The expression profiles of 32 samples covering 16 time points of SD and LD were subjected to weighted gene co-expression network analysis (WGCNA) by applying R package, with the parameters set as following: gene expression threshold: FPKM≥1.0, power=1, minimum module size=30, minimum height for merging modules=0.25. As a result, 3972 genes were sorted into 6 co-expressed modules. *CqFTL5* and *CqFTL8* with were clustered into the blue module containing 934 co-expressed genes. Then all the 934 protein sequences in blue module were uploaded to PlantTFDB v5.0 website (Jin et al., 2017) (http://planttfdb.cbi.pku.edu.cn) for TF prediction.

Results and discussion

Identification, phylogenetic relationship and chromosomal location analysis of *PEBP*s in quinoa

By using BLASTP and Pfam search methods, 23 *PEBP* homologous genes were identified in quinoa genome. The shortest quinoa *PEBP* gene (AUR62033889, named *CqFTL3*), encoding for 88 amino acid residues (Table 1), was predicted to harbor an incomplete PEBP domain. The longest *PEBP* gene (AUR62013052, named *CqFTL2*), encoding for 339 amino acid residues (Table 1), was predicted to harbor two PEBP domains. These results are in line with the points in previous study (Jarvis et al., 2017; Štorchová, 2020), indicating that *CqFTL3* may be a pseudogene and *CqFTL2* may experience tandem duplication.

To evaluate the phylogenetic relationships between *PEBP* genes of quinoa, and closet relatives of its diploid ancestors, *C. pallidicaule* (AA) and *C. suecicum* (BB), and its relatives in Amaranthaceae family, spinach (*Spinacia oleracea*) and sugar beet (*Beta vulgaris*), and model dicot and monocot plants, Arabidopsis (*Arabidopsis thaliana*) and rice (*Oryza sativa*), a Neighbor-Joining phylogenetic tree containing 83 sequences (Supplementary File 1) was inferred using MEGA 11.0 (Tamura et al., 2021). As displayed in Figure 1, 23 quinoa *PEBP* genes were sorted into three major subfamilies. Quinoa contains 5 *MFT* clade members, named *CqMFT1* to *CqMFT5*, 11 *FT-like* (*FTL*) clade members, named *CqFTL1* to

TABLE 1 Summary of *PEBP* gene family in *C. quinoa* and *B. vulgaris*.

Subfamily	Sub-genome	Gene ID	Gene name	Chr	Start-end position (+/- strand)	CDS (bp)	Protein (aa)	Ortholog in *B. vulgaris*	Ortholog ID	Orthologs in quinoa	Ratio (A: B)	Ratio (Cq: Bv)
MFT	A	AUR62029959	*CqMFT1*	Chr08	39662874-39666112 (+strand)	519	173				3:2	5:2
	A	AUR62012495	*CqMFT2*	Chr02	4599298-4601543 (-strand)	537	179					
	B	AUR62014699	*CqMFT3*	Chr01	29212627-29215543 (-strand)	537	179					
	B	AUR62014698	*CqMFT4*	Chr01	29210009-29211182 (-strand)	537	179	*BvMFT1*	EL10Ac8g20548.1	*CqMFT5*		
	A	AUR62012496	*CqMFT5*	Chr02	4594321-4597301 (-strand)	483	161	*BvMFT1*	EL10Ac8g20548.1	*CqMFT4*		
FTL	A	AUR62010060	*CqFTL1*	Chr15	4930835-4933952 (-strand)	540	180	*BvFT1*	EL10Ac9g21401.1	*CqFTL2*	5:6	11:3
	B	AUR62013052	*CqFTL2*	Chr17	79266951-79277600 (+strand)	1017	339	*BvFT1*	EL10Ac9g21401.1	*CqFTL1*		
	A	AUR62033889	*CqFTL3*	Chr10	48544986-48545464 (+strand)	264	88					
	A	AUR62000271	*CqFTL4*	Chr12	3192361-3196369 (+strand)	594	198	*BvFT2*	EL10Ac4g10025.1	*CqFTL5*		
	B	AUR62006619	*CqFTL5*	Chr05	77596526-77601590 (-strand)	528	176	*BvFT2*	EL10Ac4g10025.1	*CqFTL4*		
	A	AUR62026245	*CqFTL6*	Chr14	23254495-23258984 (+strand)	525	175	*BvFTL3*	EL10Ac6g13314.1	*CqFTL9*		
	B	AUR62026433	*CqFTL7*	Chr06	69532880-69534843 (+strand)	522	174					
	A	AUR62026237	*CqFTL8*	Chr14	22669640-22670374 (+strand)	315	105					
	B	AUR62026437	*CqFTL9*	Chr06	68781713-68782414 (+strand)	315	105			*CqFTL6*		
	B	AUR62026436	*CqFTL10*	Chr06	69126739-69127582 (+strand)	339	113					
	B	AUR62026435	*CqFTL11*	Chr06	69423321-69434601 (+strand)	792	264					

(Continued)

10.3389/fpls.2022.1119049

TABLE 1 Continued

Subfamily	Sub-genome	Gene ID	Gene name	Chr	Start-end position (+/- strand)	CDS (bp)	Protein (aa)	Ortholog in *B. vulgaris*	Ortholog ID	Orthologs in quinoa	Ratio (A: B)	Ratio (Cq: Bv)
TFL	B	AUR62009771	*CqTFL1*	Chr18	26895130-26896703 (-strand)	504	168	*BvTFL1, BvTFL2*	EL10Ac7g15814.1, EL10Ac7g16930.1	*CqTFL3*	4:2	7:3
	A	AUR62039217	*CqTFL2*	Chr07	111421421-111422986 (+strand)	513	171					
	A	AUR62039216	*CqTFL3*	Chr07	111401972-111409400 (+strand)	792	264	*BvTFL1*	EL10Ac7g15814.1	*CqTFL1*		
	A	AUR62033497	*CqTFL4*	Chr07	8662902-8665663 (+strand)	519	173	*BvTFL2*	EL10Ac7g16930.1	*CqTFL5*		
	B	AUR62028545	*CqTFL5*	Chr17	43906953-43909946 (-strand)	519	173	*BvTFL2*	EL10Ac7g16930.1	*CqTFL4*		
	NA	AUR62021217	*CqTFL6*	Chr00	35800827-35802318 (+strand)	522	174	*BvBFT1*	EL10Ac3g05715.1	*CqTFL7*		
	A	AUR62016010	*CqTFL7*	Chr07	77891902-77893406 (-strand)	522	174	*BvBFT1*	EL10Ac3g05715.1	*CqTFL6*		

10.3389/fpls.2022.1119049

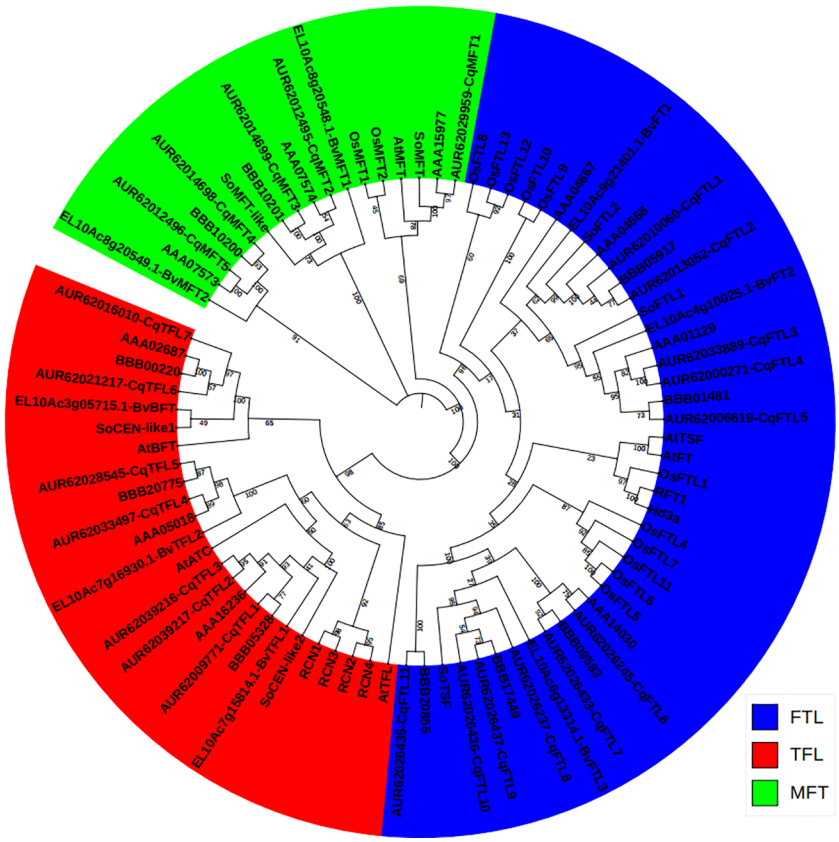


FIGURE 1
Phylogenetic relationship between the *PEBP* genes from *Chenopodium quinoa*, *C. pallidicaule*, *C. suecicum*, *Beta vulgaris*, *Spinacia oleracea*, *Arabidopsis thaliana* and *Oryza saltiva*. The PEBP protein sequences were downloaded from Phytozome V13 database. *MFT*, *FT* and *TFL* subclades are colored in green, blue and red, respectively. A total of 83 protein sequences were aligned using CLASTALW in MEGA 11.0. The phylogenetic tree was constructed by MEGA 11.0 using the Neighbor-Joining method with a bootstrap of 1000. The tree is unrooted, bootstrap values are indicated on branches.

CqFTL11, 7 *TFL1-like* (*TFL*) clade members, named *CqTFL1* to *CqTFL7* (Figure 1). *C. pallidicaule* contains 3 *MFT*s, 4 *FTL*s and 3 *TFL*s, and *C. suecicum* contains 2 *MFT*s, 5 *FTL*s and 3 *TFL*s (Figure 1).

Then, we drew a chromosomal location map of *PEBP*s. As illustrated in Figure 2, 23 *PEBP* genes are distributed on 12 chromosomes with the exception that *CqTFL6* is anchored on Chr00. Sub-genome A and sub-genome B have nearly equal number of *PEBP*s. Chr06 of sub-genome B and Chr07 of sub-genome A contain the largest number of *CqFTL*s and *CqTFL*s, respectively.

Gene structure and conserved motifs of *PEBP* genes

Diagram of *PEBP* gene structures shows that, out of 23 *PEBP*s, most (17 of 23) harbor 4 exons and 3 introns (Figure 3). *CqFTL2* contains the most (8 exons and 7 introns), whereas *CqFTL3* contains the least (2 exons and 1 intron) number of exons and introns (Figure 3). MEME program was used to identify the conserved motifs in quinoa PEBP proteins. A total of 8 conserved motifs were identified in PEBP proteins (Figure 3). Motif composition schematic analysis indicates that motif 1, 2, 3, 4 and 5 were constitutively occurred in most of the *PEBP* genes (Figure 3). Yet, several PEBP proteins lack some specific motifs. For example, CqFTL2 lacks motif 1, 2 and 3; CqFTL8, CqFTL9 and CqFTL10 lack motifs 2 and 3; CqFTL11 lacks motif 4; CqMFT1 lacks motif 5; CqMFT4 lacks motifs 2, 4, and 5 (Figure 3). In addition, we found that motif 6 was occurred in most of the *FTL* and *TFL* subfamilies, whereas was absent in *MFT* subfamily (Figure 3). Motif 7 instead of motif 8 was possessed by most of the *FTL* subfamily members, while motif 8 was exclusively occurred in majority of *TFL* and *MFT* subfamilies (Figure 3). As PEBP proteins usually interact with other proteins (Taoka et al., 2013; Tsuji et al., 2013), such as FD and 14-3-3 proteins, to exert their roles, we speculated that these distinct motif compositions of quinoa PEBPs may result in varied protein-protein interactions and function diversities.

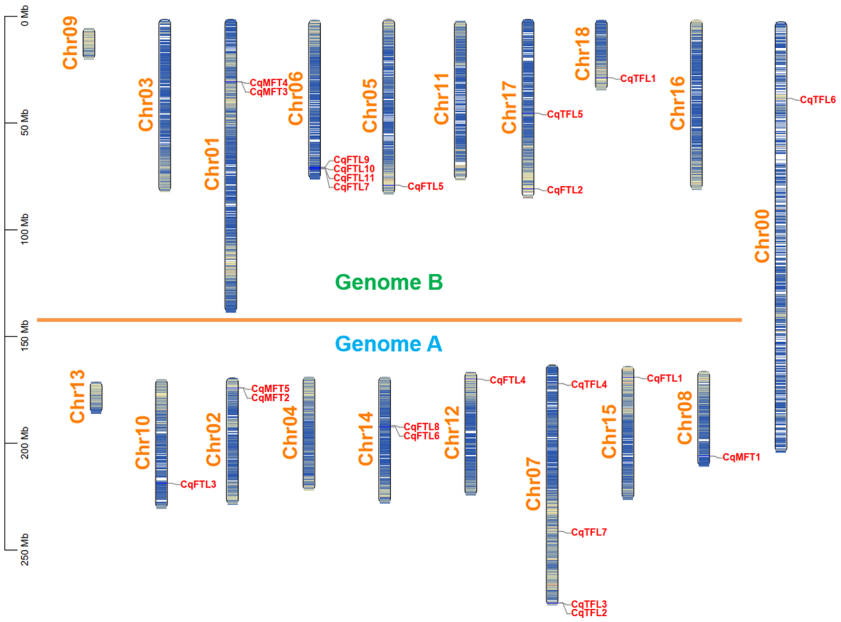


FIGURE 2
Physical location of *PEBP* genes on quinoa chromosomes. Chromosome segments colored in white and blue indicate high and low gene densities, respectively. All the quinoa chromosomes are divided into sub-genome A and B with the exception of Chr00.

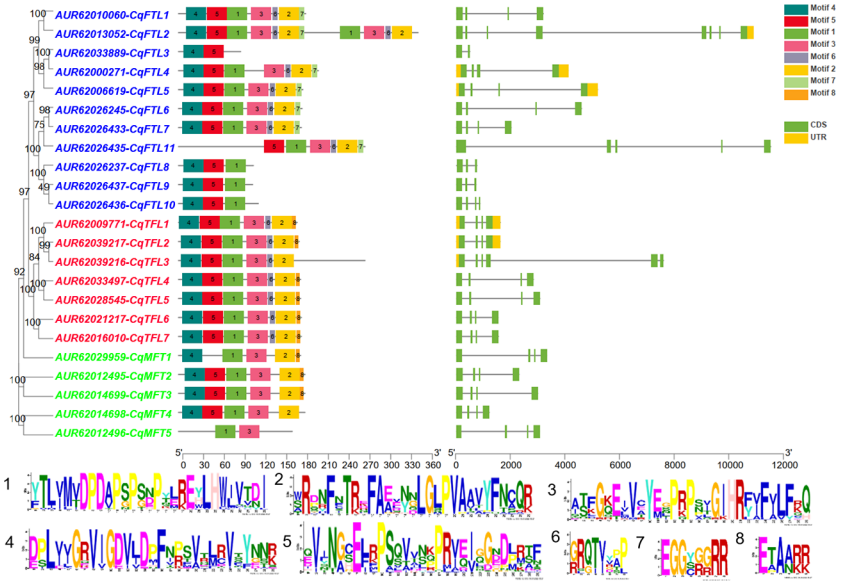


FIGURE 3
Diagram shows the conserved amino acid motifs and gene structure of *PEBP* genes in quinoa. The conserved motifs were identified by MEME. A total of 8 motifs were discovered among all *PEBP* genes. Different motifs are indicated by the colored boxes. Motif 1 to motif 5 are appeared in most PEBP proteins. Amino acid sequences of the conserved motifs are listed below. The identities of *FT*, *TFL* and *MFT* subclade genes are colored in blue, red and green in phylogenetic tree at the left side. Exons, introns and untranslated regions (UTRs) of *PEBP* genes are denoted by green, yellow boxes and dark lines, respectively.

Multiple sequence alignment suggests that the conserved amino acid residue tyrosine (Y) or histidine (H) in motif 1 is a key site distinguishing *FTL* and *TFL* clades (Figure 4), in consistent with the findings in PEBP families of other plants (Wang et al., 2015). Besides, as the occurrences of motif 7 and motif 8 in *FTL* and *TFL*/*MFT* subfamilies are mutually exclusive (Figure 4), we speculated that in these two motifs there may be specific amino acid residues distinguishing *FTL* and *TFL*/*MFT* clades. As expected, we found that, the third amino acid residue in motif 7 and 8 was Glycine (G) in *FTL* genes, whereas it was Alanine (A) in *TFL* and *MFT* genes (Figure 4). Thus, this key site (G/A) may provide as a novel site for investigation of the converse functions of *FTL* and *TFL*/*MFT* clades.

Gene duplication, collinearity and evolutionary history of *PEBP*s

Diploid sugar beet contains 8 *PEBP* genes in the genome. As a close relative of sugar beet, theoretically tetraploid quinoa (AABB) should has doubled number of *PEBP* homologs. However, we identified 23 *PEBP*s, nearly 3 times of that in *B. vulgaris* (Table 1). We wondered what caused the disproportional *PEBP* family expansion. Gene number ratios between quinoa and sugar beet in *MFT* and *TFL* clades are 5:2 and 7:3, respectively (Table 1). Surprisingly, gene number ratio in *FTL* clade is 11:3 (Table 1), remarkably higher than the plausible ratio 2:1. We analyzed the collinearities of *PEBP* genes between sub-genomes, and found that a total of 7 orthologous gene pairs from sub-genome A and B, including *CqMFT4*/*CqMFT5*, *CqFTL1*/*CqFTL2*, *CqFTL4*/*CqFTL5*, *CqFTL6*/*CqFTL9*, *CqTFL1*/*CqTFL3*, *CqTFL4*/*CqTFL5* and *CqTFL6*/*CqTFL7*, displayed collinear relationships (Table 1; Figure 5A). In addition, *CqTFL1* and *CqTFL5* also displayed collinearity (Figure 5A), indicating inner sub-genome duplication. Collinearity analysis between quinoa and sugar beet genomes showed that, 13 out of 23 *PEBP* genes of quinoa were orthologous to 7 out of 8 *PEBP*s of sugar beet (Table 1; Figure 5B). Due to the allotetraploidy of quinoa genome, most of the sugar beet *PEBP*s have two sister genes, from quinoa sub-genome A and sub-genome B, respectively. There are 7 orthologous clusters, including *CqMFT4*/*CqMFT5-BvMFT1*, *CqFTL1*/*CqFTL2-BvFT1*, *CqFTL4*/*CqFTL5-BvFT2*, *CqFTL6-BvFTL3*, *CqTFL1*/*CqTFL3-BvTFL1*, *CqTFL1*/*CqTFL4*/*CqTFL5-BvTFL2* and *CqTFL6*/*CqTFL7-BvBFT1* (Table 1; Figure 5B). Most of these orthologous groups are in line with the gene pairs from sub-genome A and sub-genome B. There are 3 copies of quinoa *TFL* gene in the *CqTFL1*/*CqTFL4*/*CqTFL5-BvTFL2* cluster (Figure 5B), probably rising from the segmental duplication events between Chr17 and Chr18 in sub-genome B (Figure 5A).

The rest of *PEBP*s, including *CqMFT1*, *CqMFT2*, *CqMFT3*, *CqFTL3*, *CqFTL7*, *CqFTL8*, *CqFTL10*, *CqFTL11* and *CqTFL2*, lack syntenic regions. We analyzed the gene physical locations, and found some of them were distributed in close vicinity to the other *PEBP* genes (Table 1; Figure 2). Among those genes, 4 tandem repeats were found, including *CqMFT3*/*CqMFT4* on Chr01, *CqMFT2*/*CqMFT5* on Chr02, *CqFTL9*/*CqFTL10*/*CqFTL11*/*CqFTL7* on Chr06 and *CqTFL2*/*CqTFL3* on Chr07 (Table 1; Figure 2). Notably, all these tandem repeats are located on distal telomeric ends of chromosomes (Figure 5A), indicating possible higher frequency of tandem duplications in distal telomeric segments. In general, these evidences indicated that tandem duplication is the major mechanism underlying *PEBP* family expansion.

Then, we analyzed the rise of gene duplication of *PEBP*s by combining phylogenetic analysis with collinearity analysis. According to the genome map of sugar beet, *BvMFT1* and

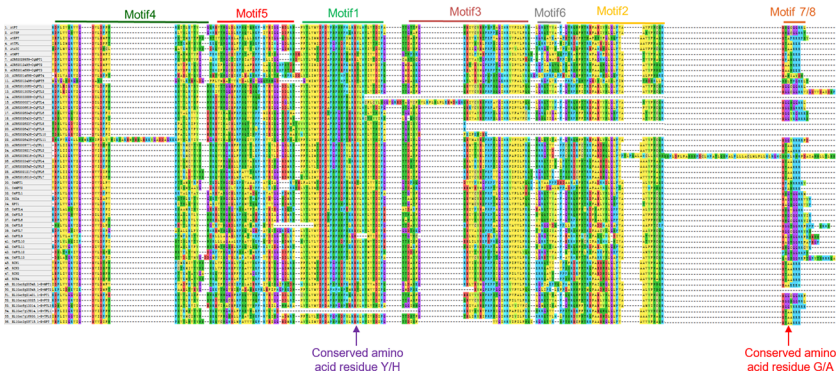


FIGURE 4
Multiple sequence alignment of the PEBP proteins from different plants. Sequences numbered with 1 to 6, 7 to 29, 30 to 48 and 49 to 56 stand for *PEBP* genes from Arabidopsis, rice, quinoa and sugar beet, respectively. The conserved motif 1 to motif 5 are indicated at the top of aligned sequences. Red arrow indicates the key conserved amino acid residue that determines *FT* and *TFL* subclade functions.

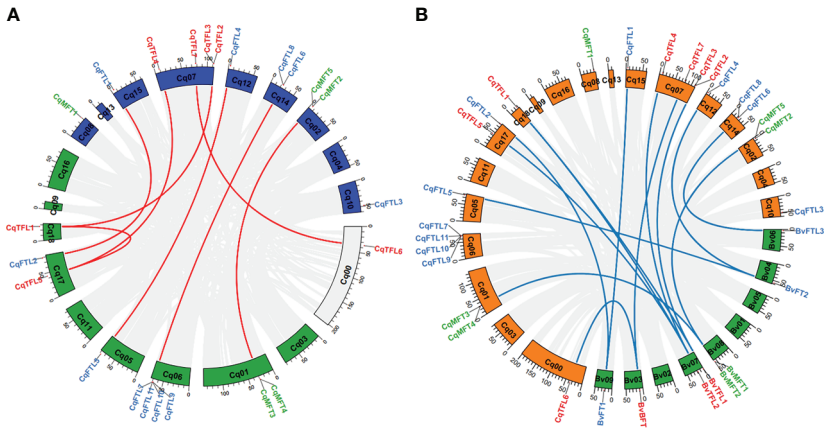


FIGURE 5
Circos plots shows the syntenic relationships between *PEBP*s from sub-genome A and B of quinoa and between *PEBP*s of quinoa and sugar beet. **(A)** Schematic representation of the syntenic relationships between homoeologous *PEBP*s of sub-genome A (blue) and B (green). Gray lines connect the homoeologous gene pairs of different chromosomes. Homoeologous *PEBP*s in quinoa are linked by red lines. **(B)** Syntenic relationships between the homoeologous *PEBP*s of quinoa and sugar beet. Chromosomes of quinoa and sugar beet are colored in green and orange, respectively. Gray lines connect the homoeologous gene pairs of quinoa and sugar beet. Blue lines connect the homoeologous *PEBP* gene pairs of quinoa and sugar beet.

BvMFT2 are in tandem repeat location. Meanwhile, we found the orthologs, in the relatives of sub-genome A (AAA07574/AAA07573 in *C. pallidicaule*) and sub-genome B (BBB10201/BBB10201 in *C. suecicum*), are also in tandem repeat locations. Thus, we deduced that the tandem duplications of *CqMFT3/CqMFT4* and *CqMFT2/CqMFT5* happened before the divergence of the ancestor of *Chenopodium* from sugar beet, far before the tetraploidization of quinoa. Though *CqMFT1* was in non-synteny with sugar beet and *C. suecicum*, however, it is sister to the *MFT* genes of Arabidopsis, and *C. pallidicaule* and spinach. Phylogenetic analysis indicated that the tandem duplication of *CqTFL2/CqTFL3* was happened after the ancestor of sub-genome A diverged from *C. pallidicaule*. *CqFTL3*, a sister of *CqFTL4* in the phylogenetic tree, but is in non-syntenic region and lacks 4 motifs of *CqFTL4*, indicating that gene amplification may occur through transposable elements. *CqFTL8*, despite located nearby *CqFTL6*, but was not detected as a tandem repeat with *CqFTL6* (Figure 5A). Meanwhile, *CqFTL8* is in non-syntenic region and lacks motif 3, 6, 2, 7 compared with the other *CqFTL* genes (Figure 3), indicating it might arise from small-scale transposition. Synteny was found in *CqFTL9* of sub-genome A and *CqFTL6* of sub-genome B (Figure 5A), whereas only *CqFTL6*, rather than *CqFTL9*, has synteny with *BvFTL3* of sugar beet (Figure 5B), suggesting the tandem duplication in *CqFTL9/CqFTL10/CqFTL11/CqFTL7* possibly was generated after the divergence of the ancestor of sub-genome B from sugar beet.

Distribution of *cis*-acting elements in *PEBP* promoters

Cis-acting elements in gene promotor have important regulatory roles mediating transcriptional activation and repression, and various *cis*-acting elements controlling specific progresses have been identified. To predict and understand the versatile functions of *PEBP* genes, we explored the *cis*-acting elements in promotors. Two kilo base pairs upstream of the translational start site of *PEBP* genes were submitted to PlantCARE database for *cis*-acting elements prediction. As displayed in Figure 6, various *cis*-acting elements, related to light, phytohormones, cold, drought and circadian clock responses, were identified. Light responsive elements (LREs) take a great proportion among various *cis*-acting elements in *PEBP* promoters, suggesting that *PEBP*s may be tightly associated with light biological events. A total of 8 kinds of LREs were identified. LRE_Box4 and LRE_G-box were the two major elements distributed in all *PEBP* promoters. Some *cis*-acting elements showed gene-specific distribution patterns. *CqFTL8* promoter harbors the most abundant LRE_Box4. *CqFTL11* promoter harbors more LRE_Sp1. LRE_G-box rather than LRE_Box4 was more abundant in the promoters of *MFT* clade. More low-temperature responsive elements (LTRs) were enriched in *CqMFT1* promoter, indicating the possible roles of *CqMFT1* in sensing cold stress. Circadian regulatory elements (Circadian) were only distributed in the promoters of *CqFTL7*, *CqFTL11*, *CqMFT2*, *CqTFL2* and *CqTFL3*. Abscisic acid responsive

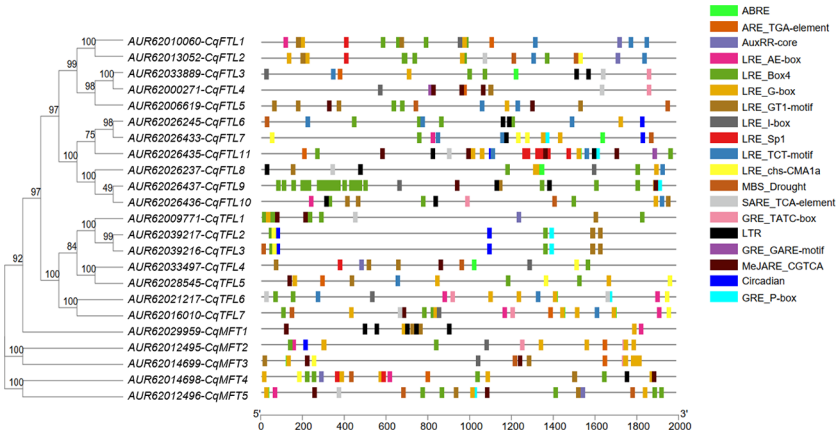


FIGURE 6
Illustration shows the physical locations of various *cis*-acting elements in the promoter regions of *PEBP*s. Promoter sequences of quinoa *PEBP*s were submitted to PlantCARE database to identify various *cis*-acting elements. Boxes with different colors indicate various *cis*-acting elements located in the promoter regions of *PEBP*s. Abbreviations and sequences: ABRE, abscisic acid responsive element, ABRE: ACGTG; ARE, auxin-responsive element, ARE_TGA-element: AACGAC; AuxRR, auxin responsive element, AuxRR-core: GGTCCAT; LRE, light responsive element, LRE_AE-box: AGAAACAA, LRE_Box 4: ATTAAT, LRE_G-box: CACGTG, LRE_GT1-motif: GGTTAA, LRE_I-box: AGATAAGG, LRE_Sp1: GGGCGG, LRE_TCT-motif: TCTTAC, LRE_chs-CMA1a: TTACTTAA; MBS, MYB binding site, MBS_Drought: CAACTG; SARE, salicylic acid responsive element, SARE_TCA-element: CCATCTTTTT; GRE, gibberellin-responsive element, GRE_TATC-box: TATCCCA, GRE_GARE-motif: TCTGTTG, GRE_P-box: CCTTTTG; LTR, low-temperature responsive element, CCGAAA; MeJARE, MeJA-responsive element, MeJARE_CGTCA: CGTCA; Circadian, circadian regulatory element, Circadian: CAAAGATATC.

elements (ABREs) were abundantly located in promoters of *CqMFT2* and *CqMFT3*, indicating these two gene may be involved in ABA signaling. In addition, we noticed distributions of the *cis*-acting elements in promoters of some gene pairs were also in synteny. The distributions of ABRE, AuxRR-core, LRE_Box4, LRE_G-box, LRE_GT1-motif, LRE_Sp1 and LRE_TCT-motif in *CqFTL1* promoter were in synteny with that in *CqFTL2* promoter. This is in line with the collinearity between these two genes. The distributions of the *cis*-acting elements including Circadian, GRE_P-box, LRE_Box4, LRE_chs-CMA1a and LRE_GT1-motif in *CqTFL2* promoter were in highly collinear relationship with that in *CqTFL3* promoter. This may be caused by the tandem duplication of *CqTFL2*/*CqTFL3*. The specific distributions of *cis*-acting elements in *PEBP* promoters suggest their possible roles in different biological events.

Tissue-specific expression analysis of *PEBP*s

To characterize the expression patterns of *PEBP*s in various quinoa organs, we analyzed the RNA-seq data and compared the FPKM values of 23 *PEBP* genes in seedling, leaf, stem, inflorescence and seed. Out of 23 *PEBP* genes, 18 were expressed in at least one organ tissue (Figure 7; Supplementary File 2). Generally, expressions of *PEBP*s in various tissues were clade-specific. *FTL* clade genes were more enriched in leaf, stem and inflorescence, and *TFL* clade genes were abundant in seedling, stem and inflorescence, whereas expressions of the *MFT* clade genes were relatively higher in seed (Figure 7). It seems like that expressions of the *MFT* clade genes are more conserved, and all the *MFT* gene pairs were expressed in similar patterns (Figure 7). By contrast, expressions of the orthologous gene pairs in *FTL* and *TFL* clades were likely differing (Figure 7). For *CqFTL1*/*CqFTL2*, *CqFTL1* was highly enriched in inflorescence, whereas *CqFTL2* was ubiquitously expressed in all tissues (Figure 7). For *CqFTL4*/*CqFTL5*, *CqFTL4* was expressed mainly in leaf, and *CqFTL5* was abundant in leaf and stem (Figure 7). Another branch containing *CqFTL7* and *CqFTL8* was expressed specifically in leaf (Figure 7). For *CqTFL1*/*CqTFL3*, higher expressions of *CqTFL1* were detected in both stem and inflorescence, and *CqTFL3* was specifically abundant in stem (Figure 7). The duplicated gene of *CqTFL2*, *CqTFL3*, was also expressed mainly in stem (Figure 7). For *CqTFL6*/*CqTFL7*, *CqTFL6* was specifically expressed in seedling, whereas *CqTFL7* transcripts were abundant in seedling and inflorescence (Figure 7). *CqTFL4*/*CqTFL5* harbored similar expression patterns, both were specifically expressed in seedling. The differing expression patterns in gene pairs indicate the possible diversified roles of orthologs from sub-genome A and sub-genome B.

The *PEBP*s involved in inflorescence branching

Branching of inflorescence is a key factor influencing plant architecture and yield. Numerous evidences have suggested the important roles of *PEBP* genes in inflorescence branching and yield

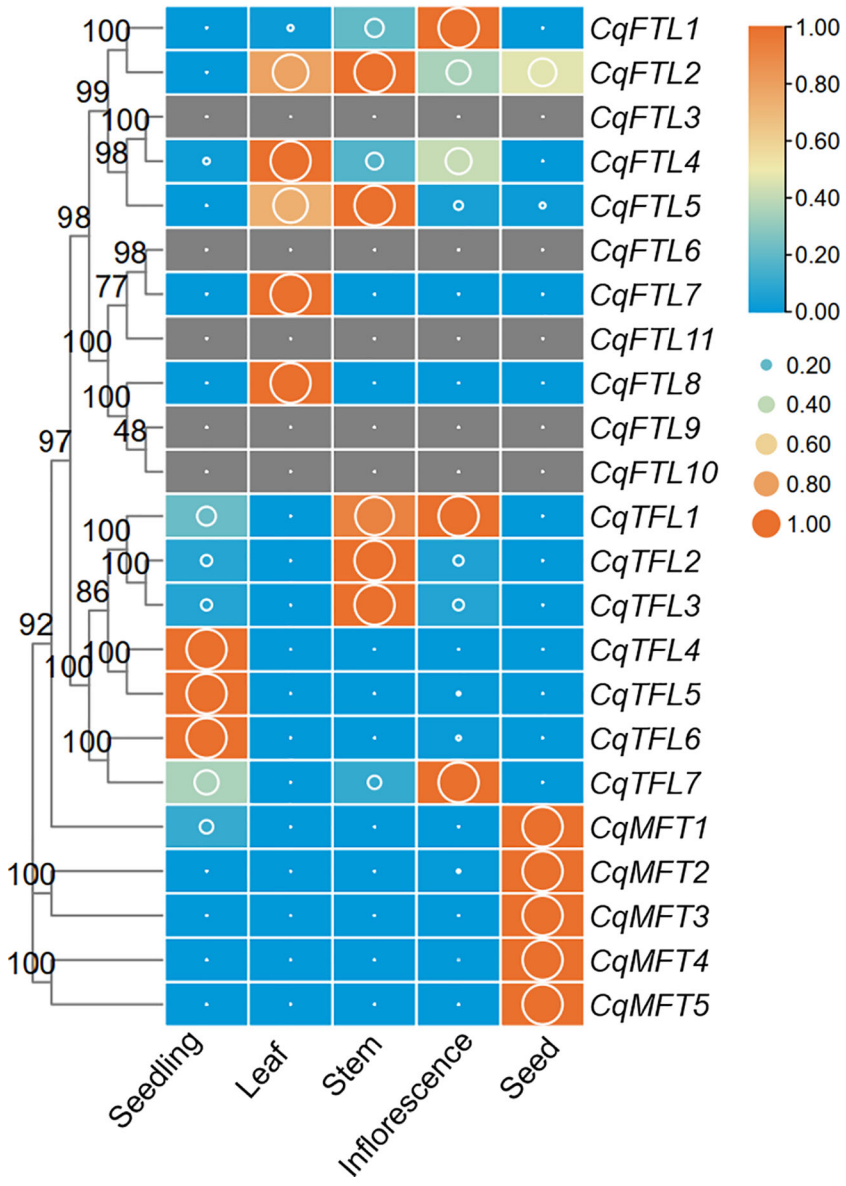


FIGURE 7
Expression profiles of *PEBP* genes in 5 quinoa tissues. The expression value is denoted by FPKM values (cutoff=0.1) generated by using the RNA-seq data from Zou et al. (Zou et al., 2017). Each expression value was generated from three replicates. Orange boxes with larger-size bubbles indicate higher expression levels. Expressions in various tissues of the same gene were normalized to 0 to 1 by row.

control (Teo et al., 2014; Benlloch et al., 2015). To investigate which quinoa *PEBP* members are potential regulator for panicle branching, we analyzed the expression changes of *PEBP*s across 6 developmental stages, before and after panicle branching. Expressions of 19 *PEBP*s were detected in at least one stage (Figure 8; Supplementary File 2). Differing from the diversified expressions in tissues above, most of the orthologous gene pairs in *FTL* and *TFL* clades showed similar expression patterns. For *FTL* clade, with the exception that *CqFTL3* and *CqFTL4* were expressed ubiquitously at non-branching stages (YP1 to YP4) (Figure 8), expressions of *CqFTL1*/*CqFTL2*, *CqFTL5*, *CqFTL6*/*CqFTL9* and *CqFTL8*, ascended with the development of inflorescence and were abundant at branching stages (P1 and P2) (Figure 8). For *TFL* clade, expression levels of gene pairs *CqTFL1*/*CqTFL3*, *CqTFL4*/*CqTFL5*, and the duplicated gene of *CqTFL3*, *CqTFL2* were relatively higher at non-branching stages (YP1 and YP2) while descended with the development of inflorescence (Figure 8), whereas the gene pairs *CqTFL6*/*CqTFL7* were relatively abundant in branching stages (P2) (Figure 8). Most of the *MFT* clade genes were ubiquitously expressed across all stages (Figure 8). The specific

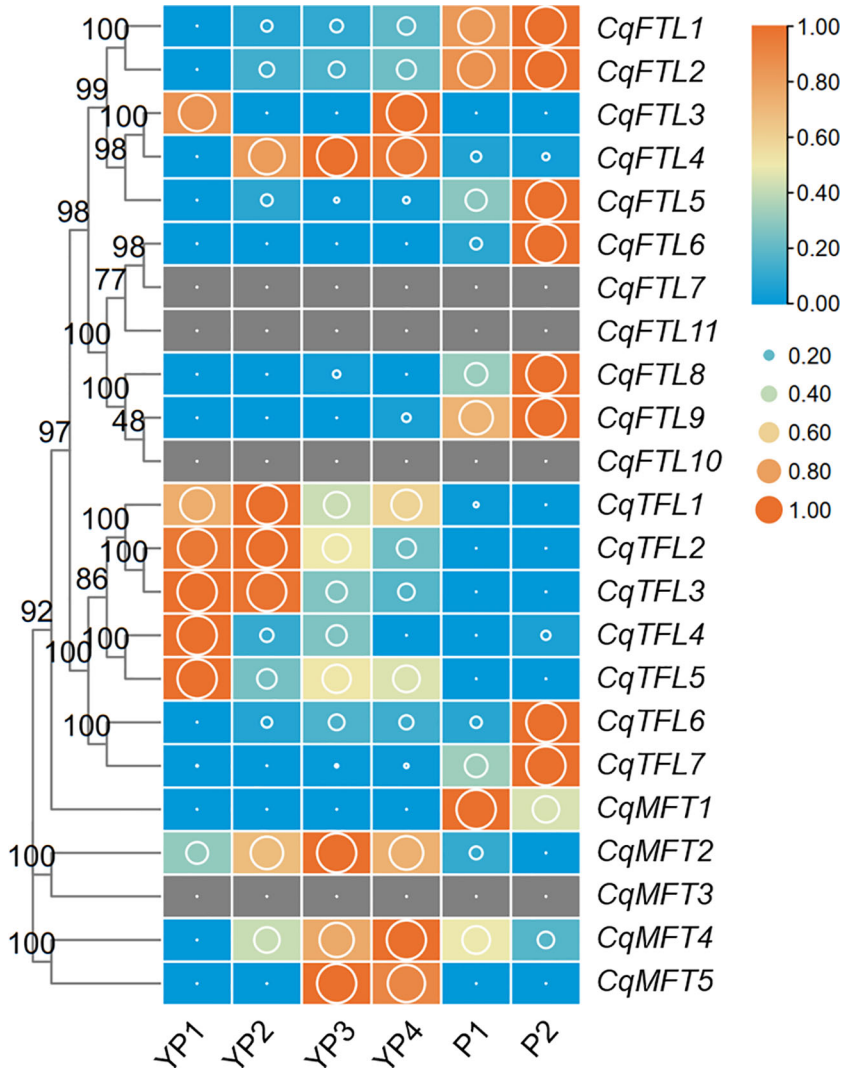


FIGURE 8
Expression of *PEBP* genes during quinoa inflorescence development. RNA-seq data of inflorescences sampled from 6 developmental stages was used to investigate the expression changes of *PEBP*s (Wu et al., 2019). Expression levels are denoted by FPKM values (cutoff=0.1). YP1 to YP4 indicates non-branching young panicles at 4 earlier developmental stages. P1 and P2 indicate branching panicles at 2 later developmental stages. Each expression value was generated from three replicates. Orange boxes with larger-size bubbles indicate higher expression levels. Expression levels at 6 stages of the same gene were normalized to 0 to 1 by row.

expressions of *CqTFL1*/*CqTFL2*/*CqTFL3* and *CqTFL4*/*CqTFL5* at non-branching stages, and the enrichment of *CqFTL1*/*CqFTL2*, *CqFTL5*, *CqFTL6*/*CqFTL9*, *CqFTL8* and *CqTFL6*/*CqTFL7* at branching stages, suggest that those *PEBP*s may participate in panicle architecture regulation.

The principal *PEBP*s involved in seed germination

Pre-harvest sprouting (PHS) is a knotty problem that influences the yield and nutritional qualities of quinoa (Wu et al., 2020). Understanding the mechanisms underlying seed germination will facilitate breeding PHS-resistant elites of quinoa. We investigated the expressions of *PEBP*s during seed germination before and after ABA treatment. A higher proportion of *MFT* clade genes (4 out of 5, 80%), whereas relative lower percentages of *FTL* clade (6 out of 11, 55%) and *TFL* clade genes (1 out of 7, 15%) were detected during seed germination (Figure 9; Supplementary File 2). We found *CqMFT2*, *CqMFT3* and *CqMFT4* were largely repressed as germination went on (Figure 9). Further, we found that the transcriptional changes of *CqMFT2*, *CqMFT3* and *CqMFT4* during germination were attenuated when treated by ABA

(Figure 9). This is in agreement with the enrichment of the *cis*-acting element ABRE in *MFT* gene promoters, indicating those three genes may respond to ABA to regulate seed germination.

Responses of *PEBP* genes to night break

Night break (NB) has a repressive effect on the flowering of short-day plant (SDP) mostly by repressing florigen-encoding genes (Ishikawa et al., 2005). Like that in other SDPs, in our previous study we also noticed the repressive effects of NB on quinoa flowering. To know which *PEBP* genes may be involved in NB responses, we analyzed the transcriptome data of quinoa leaf samples collected from SD and NB conditions. As illustrated in Figure 10, *FT* clade genes were more active than *TFL* and *MFT* clade genes. A total of 9 *PEBP* genes were detected, of which 8 were *FTL* clade genes (Figure 10; Supplementary File 2). Obviously, *CqFTL3*, *CqFTL5*, *CqFTL7*, *CqFTL8* and *CqFTL9*

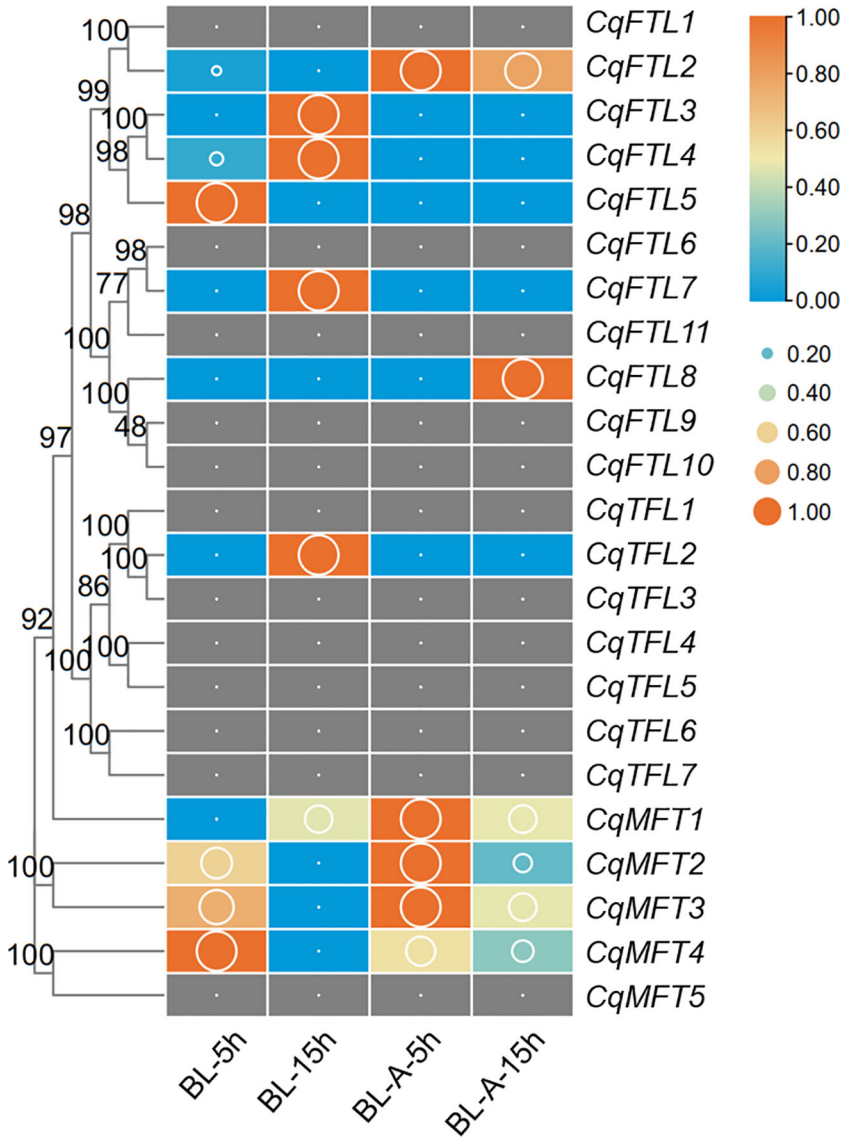


FIGURE 9
Expression changes of *PEBP*s during seed germination. RNA-seq data of control and ABA-treated seeds at 5h and 15h after imbibition was used to investigate the expression levels denoted by FPKM values (cutoff=0.1) (Wu et al., 2020). BL-A-5h/15h and BL-5h/15h indicate the seeds of cultivar "BL" with or without ABA treatment. Each expression value was generated from three replicates. Orange boxes with larger-size bubbles indicate higher expression levels. Expression values in various seed samples of the same gene were normalized to 0 to 1 by row.

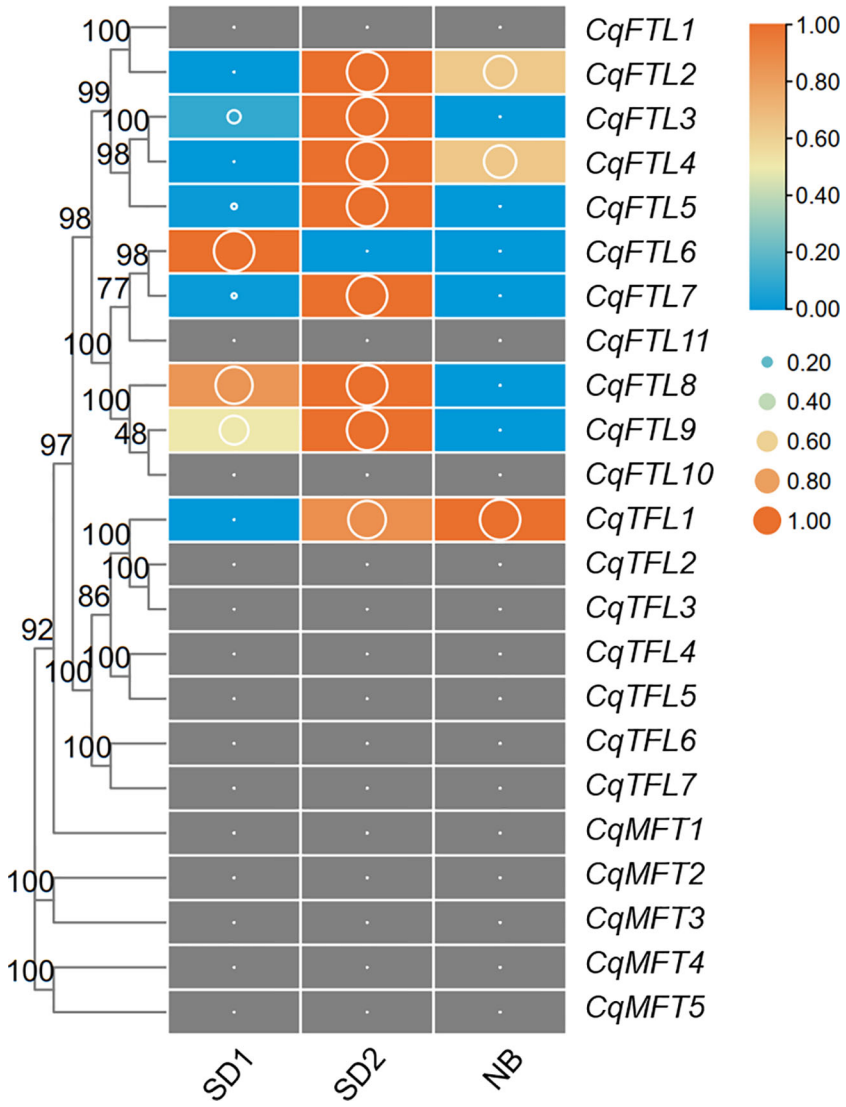


FIGURE 10
Investigation on the transcriptional changes of *PEBP*s in response to night break. Our previously published transcriptome data was used to. SD1 and SD2 indicate leaf samples collected from 14-day- and 26-day-old plants under short-day conditions. NB indicates the leaf sample of SD2 plants with 2d night-break treatment. Each expression value was generated from three replicates. Orange boxes with larger-size bubbles indicate higher expression levels. Expression values (cutoff=0.1) in various leaf samples of the same gene were normalized to 0 to 1 by row.

were expressed at higher levels under SD, whereas were rapidly down-regulated by NB treatment (Figure 10). The opposed expression patterns of the 5 *FTL* genes under SD and NB indicate their possible positive effects in quinoa flowering. In addition, we noticed the differing responses to NB between orthologous gene pairs. For example, in *CqFTL4*/*CqFTL5*, *CqFTL5* was sensitive to NB, whereas *CqFTL4* was only slightly repressed by NB (less than 2 folds). In *CqFTL6*/*CqFTL9*, *CqFTL9* was largely repressed by NB, while *CqFTL6* were down-regulated regardless of light conditions.

Diurnal expression analysis of *PEBP*s

As quinoa belongs to SDPs, SD induces whereas long day (LD) represses quinoa flowering. To evaluate the relationships between *PEBP*s and day lengths, we performed a comprehensive analysis on *PEBP* expressions over 24 h under SD and LD. The quinoa plants were grown in growth chamber supplied with SD or LD. Before bolting stage, the top-fully-expanded leaves of two-week-old seedlings were collected for RNA-seq. We detected the expressions of 9 *FTL* clade genes, but only 2 *TFL*

and 1 *MFT* clade genes were detected in leaves at this stage (Figures 11A-L; Supplementary File 2). The 24 h expression profile showed that, the expressions of several quinoa *PEBP* genes were diurnally rhythmic and differed under different day lengths (Figures 11A, E, F, H, L). Clearly, expression patterns between the orthologous gene pairs *CqFTL1*/*CqFTL2* (Figures 11A, B), *CqFTL4*/*CqFTL5* (Figures 11D, E), *CqFTL6*/*CqFTL9* (Figures 11F, I) were different. Expressions of *CqFTL1*, *CqFTL4*, *CqFTL5* and *CqFTL8* were considerably higher under SD than that under LD (Figures 11A, D, E, H). By contrast, *CqFTL6* was induced by LD rather than SD (Figure 11F). Expression levels of *CqFTL7* and *CqFTL9* were slightly higher under SD than that under LD (Figures 11G, I). Expression levels of *CqFTL3* were comparable between under SD and under LD (Figure 11C). *CqFTL1* and *CqFTL8* expressions were peaked at dawn time, and *CqFTL5* and *CqFTL9* were abundantly expressed at 2 h after dawn break (Figures 11A, E, H, I). *CqFTL2* was highly expressed throughout the day under both SD and LD (Figure 11B). The expressions of *CqTFL1* and *CqTFL2* were comparable between under SD and LD (Figures 11J, K). *CqMFT2* was induced by SD and peaked at the end of day (Figures 11L). Together, these results indicated that *CqFTL1*, *CqFTL4*, *CqFTL5* and *CqFTL8* may act as SD-type flowering inducer, whereas *CqFTL6* may act as LD-type flowering inducer. Recently, Patiranage et al. also investigated the diurnal expression patterns of *FTL* genes, but only limited to 6 *FTL* genes of quinoa at bolting stage (Patiranage et al., 2021). As florigen encoding *FT-like* genes are highly induced in leaves before floral transition, we chose to investigate the diurnal expressions at vegetative stage before bolting. The classical florigen genes are likely to diurnally expressed and are sensitive to night break (Ishikawa et al., 2005; Meng et al., 2011). By combining the responses to night break and photoperiodic expressions of quinoa *FTL* genes, we speculated

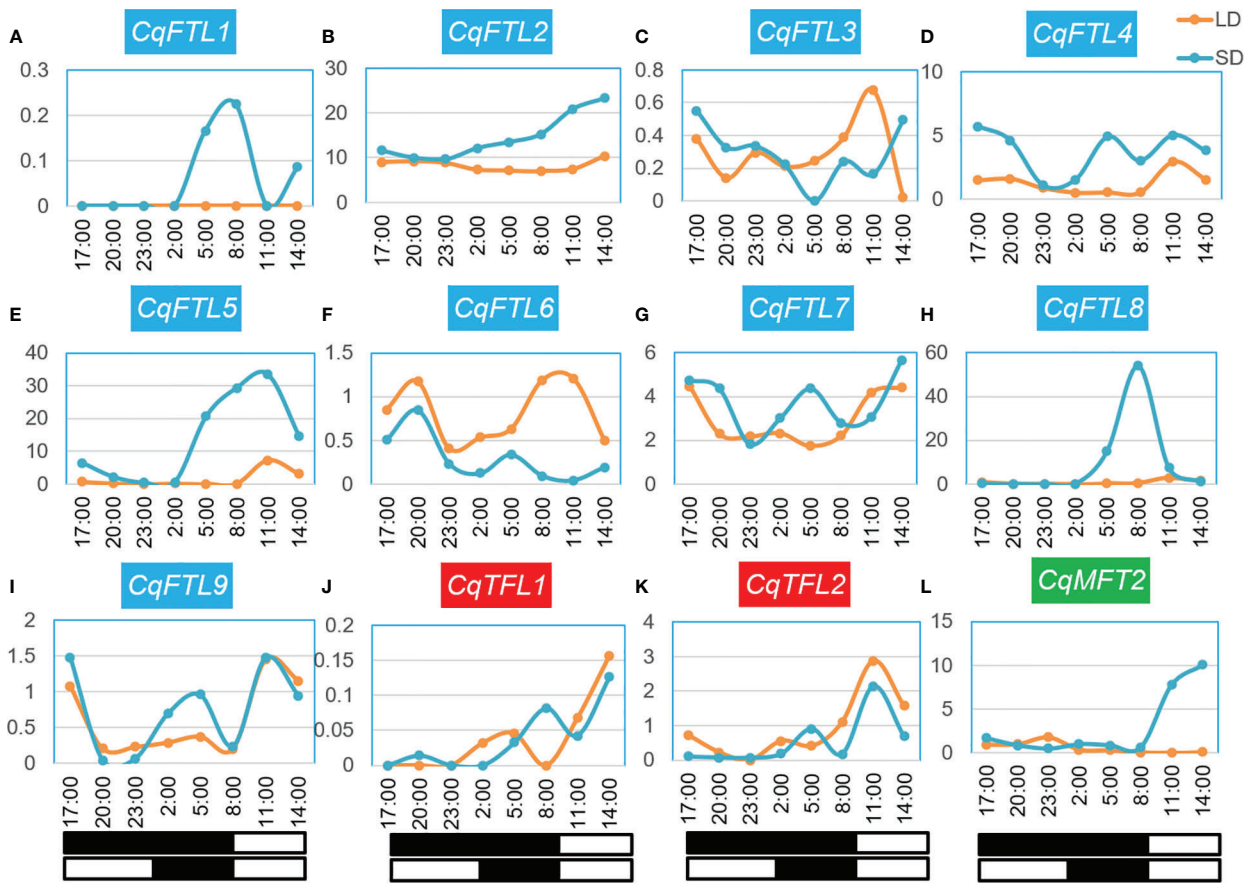


FIGURE 11
Diurnal expression patterns of *PEBP*s under short-day and long-day conditions. **(A-L)** Expression patterns of *CqFTL1*, *CqFTL2*, *CqFTL3*, *CqFTL4*, *CqFTL5*, *CqFTL6*, *CqFTL7*, *CqFTL8*, *CqFTL9*, *CqTFL1*, *CqTFL2* and *CqMFT2* under short-day and long-day conditions. Two-week-old plants grown under short-day and long-day conditions were used to collect the leaf samples, respectively, at 17:00, 20:00, 23:00, 02:00, 05:00, 08:00, 11:00 and 14:00. RNA-seq analysis was performed to investigate the expression patterns of *PEBP*s. White and black boxes indicate day and night time, respectively. Each expression value was generated from two replicates. Orange and blue curves stand for the expression levels of *PEBP*s under long-day and short-day conditions.

that *CqFTL5* and *CqFTL8* are the major florigen-encoding genes in quinoa.

Prediction of the putative TFs upstream of *CqFTL5* and *CqFTL8* by large-scale co-expression network analysis

As floral integrators, *FT* genes usually are regulated by a complex network consisting of a lot of transcription factors (TFs) (Tsuji et al., 2013). To know which TFs may lay in the signal cascades upstream of *FT* genes in quinoa, WGCNA was performed to identify the co-expressed genes. The expression profiles of 32 samples covering 16 time points of SD and LD were used for WGCNA. A total of 6 co-expression modules were obtained. We found only *CqFTL5* and *CqFTL8* were predicted to occur in those modules. *CqFTL5* and *CqFTL8* were clustered into the same module-blue module, which contains 934 co-expressed genes. Then those genes were submitted to PlantTFDB v5.0 for TF prediction. A total of 64 TFs, 6.85% of blue module genes, were predicted to be co-expressed with *CqFTL5* and *CqFTL8*. A heatmap was drawn according to their expression profiles (Figure 12). Clearly, those TFs were clustered into two major clusters. One cluster displayed LD-repressive and SD-inducible expression patterns, consistent with that of *CqFTL5* and *CqFTL8*, whereas, the other cluster displayed opposed expression patterns (Figure 12). Thus, those TFs of the two clusters were predicted to be putative promoters and repressors of *CqFTL5* and *CqFTL8*. Notably, compared with the overall TF percentage of blue module genes, remarkable higher rates of family members were found in CO-like (CONSTANS), DBB (Double B-box), GeBP, HSF (Heat Stress Transcription Factors) and bZIP families (Table 2). *DBB* encodes for a double B-box zinc finger protein, and is regulated by circadian rhythm and

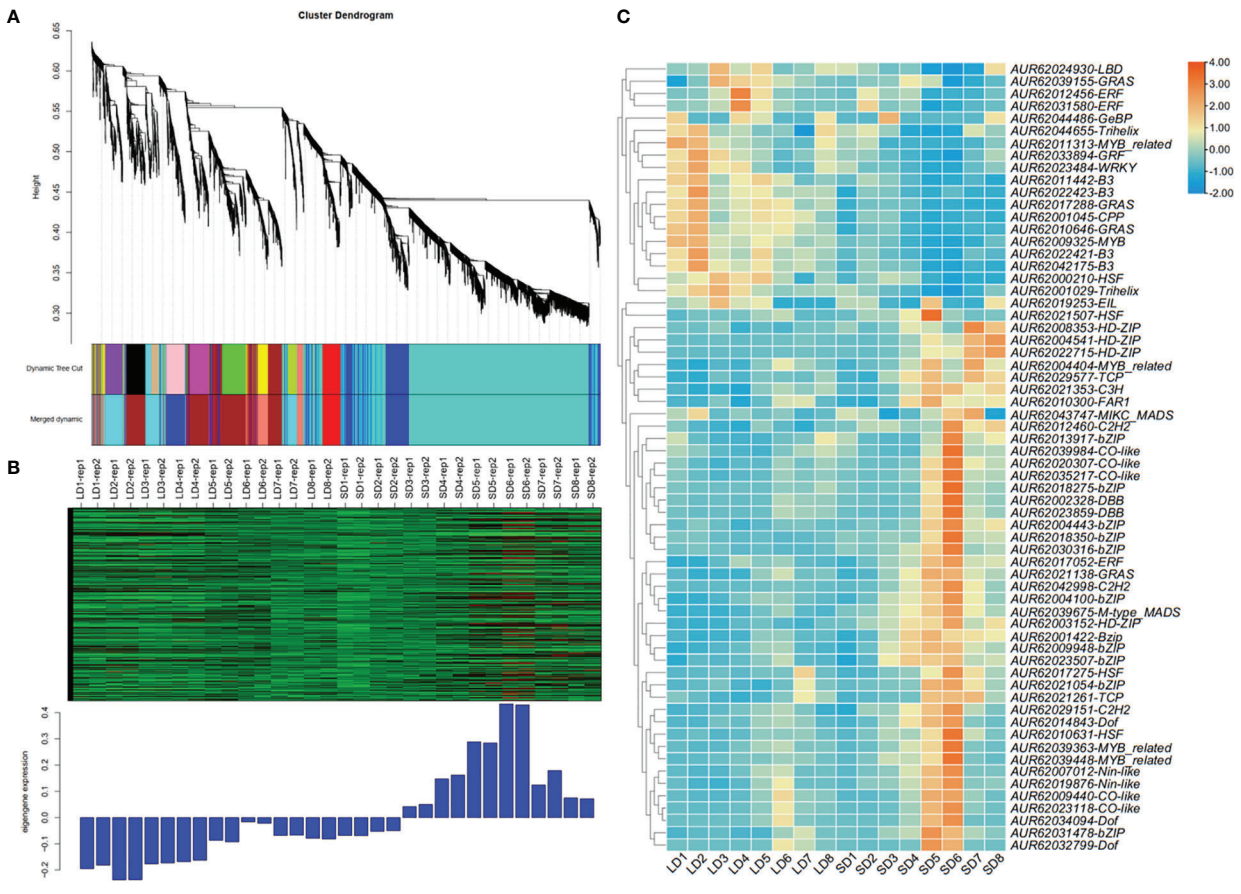


FIGURE 12
Weighted gene co-expression network analysis and transcription factor prediction methods were used to identify the co-expressed TFs with *CqFTL5* and *CqFTL8*. **(A)** All the expressed genes were sorted into 6 co-expressed modules by applying weighted gene co-expression network analysis. **(B)** *CqFTL5* and *CqFTL8* were clustered into blue module containing 934 co-expressed genes. SD1 to SD8, or LD1 to LD8, refers to the samples diurnally collected at 17:00, 20:00, 23:00, 02:00, 05:00, 08:00, 11:00 and 14:00. **(C)** Heatmap shows the expression levels of 64 TFs identified from blue module.

participates in light signal transduction during photomorphogenesis. Previous studies indicated that *CO-like* encodes for a B-box zinc finger protein, and functions as an important mediator of circadian clock to control *FT* during floral transition (Putterill et al., 1995; Izawa et al., 2002). Thus, like that in other plants, the *CO-FT* signal pathway is presumably conserved in quinoa flowering. The central roles of MADS-box genes during floral transition have been extensively investigated (Kim et al., 2008; Ryu et al., 2009). We identified that two MADS-box genes were co-expressed with *CqFTL5* and *CqFTL8*. Together, we speculated that these TFs may modulate the transcription levels of *CqFTL5* and *CqFTL8* in leaves to control quinoa flowering time.

Conclusions

In this study, we identified 23 *PEBP* family members in quinoa. We investigated the relationships of *PEBP* genes in sugar beet and relatives of quinoa diploid ancestors, and identified 7 orthologous *PEBP* gene pairs. Evolution analysis indicated the reasons for *PEBP* duplication events varied, and tandem duplication is the major driving force for *PEBP* family expansion. Then, we identified the pivotal *PEBP* genes for inflorescence branching, seed germination and flowering time regulation, and predicted 64 putative transcription factors upstream of *CqFTL5* and *CqFTL8* by performing co-expression network analysis.

TABLE 2 The transcription factors co-expressed with *CqFTL5* and *CqFTL8* in blue module.

TF family	TF number in blue module	Number of TF family members	Percentage
B3	4	107	3.74%
bZIP	11	89	12.36%
C2H2	3	98	3.06%
C3H	1	79	1.27%
CO-like	5	14	35.71%
CPP	1	10	10.00%
DBB	2	8	25.00%
Dof	3	38	7.89%
EIL	1	9	11.11%
ERF	3	123	2.44%
FAR1	1	136	0.74%
GeBP	1	5	20.00%
GRAS	4	54	7.41%
GRF	1	11	9.09%
HD-ZIP	4	46	8.70%
HSF	4	30	13.33%
LBD	1	54	1.85%
MIKC_MADS	1	50	2.00%
M-type_MADS	1	53	1.89%
MYB	1	107	0.93%
MYB_related	4	112	3.57%
Nin-like	2	20	10.00%
TCP	2	31	6.45%
Trihelix	2	48	4.17%
WRKY	1	90	1.11%

Data availability statement

The data presented in the study are deposited in the NCBI SRA repository (http://trace.ncbi.nlm.nih.gov/Traces/sra), accession number PRJNA824606, PRJNA825321, PRJNA824547, PRJNA824641, PRJNA824640, PRJNA824668, PRJNA824959, PRJNA824960, PRJNA824961, PRJNA824962 and PRJNA824963.

Author contributions

QW conceived and designed this study. QW performed most of the bioinformatics analysis and wrote the manuscript. XB participated in phylogenetic analysis and evolution analysis. MN, LL and YL cultivated the seedlings and harvested the samples for RNA sequencing. YF, CL, XY and LZ participated in expression analysis. All authors contributed to the article and approved the submitted version.

Funding

This work is supported by Natural Science Foundation of Sichuan Province (Grant 2022NSFSC1773, 2022NSFSC1725), National Natural Science Foundation of China (Grant 31701493), China Agriculture Research System (Grant CARS-07-B-1), the project of the Key Laboratory of Coarse Cereal Porcessing, Ministy of Agriculuture and Rural Affairs (Grant 2021CC004, 2019CC14), and the project of Irradiation Preservation Technology Key Laboratory of Sichuan Province, Sichuan Institute of Atomic Energy (Grant FZBC2022004).

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.2022.1119049/full#supplementary-material

SUPPLEMENTARY FILE 1
Protein sequences of PEBPs from Chenopodium quinoa, C. pallidicaule, C. suecicum, Beta vulgaris, Spinacia oleracea, Arabidopsis thaliana and Oryza saltiva.

SUPPLEMENTARY FILE 2
Expression data of *PEBP* genes in various biological events.

SUPPLEMENTARY FILE 3
Expression profiles of the TFs in blue module.

References

Benlloch, R., Berbel, A., Ali, L., Gohari, G., Millan, T., and Madueno, F. (2015). Genetic control of inflorescence architecture in legumes. *Front. Plant Sci.* 6, 543. doi: 10.3389/fpls.2015.00543

Bi, X., van Esse, W., Mulki, M. A., Kirschner, G., Zhong, J., Simon, R., et al. (2019). CENTRORADIALIS interacts with FLOWERING LOCUS T-like genes to control floret development and grain number. *Plant Physiol.* 180, 1013–1030. doi: 10.1104/pp.18.01454

Chab, D., Kolar, J., Olson, M. S., and Storchova, H. (2008). Two flowering locus T (FT) homologs in chenopodium rubrum differ in expression patterns. *Planta* 228, 929–940. doi: 10.1007/s00425-008-0792-3

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

Comadran, J., Kilian, B., Russell, J., Ramsay, L., Stein, N., Ganal, M., et al. (2012). Natural variation in a homolog of antirrhinum CENTRORADIALIS contributed to spring growth habit and environmental adaptation in cultivated barley. *Nat. Genet.* 44, 1388–1392. doi: 10.1038/ng.2447

Danilevskaya, O. N., Meng, X., and Ananiev, E. V. (2010). Concerted modification of flowering time and inflorescence architecture by ectopic expression of TFL1-like genes in maize. *Plant Physiol.* 153, 238–251. doi: 10.1104/pp.110.154211

Dohm, J. C., Minoche, A. E., Holtgräwe, D., Capella-Gutiérrez, S., Zakrzewski, F., Tafer, H., et al. (2013). The genome of the recently domesticated crop plant sugar beet (Beta vulgaris). *Nature* 505, 546. doi: 10.1038/nature12817

Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., et al. (2016). The pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* 44, D279–D285. doi: 10.1093/nar/gkv1344

Goretti, D., Silvestre, M., Collani, S., Langenecker, T., Méndez, C., Madueno, F., et al. (2020). TERMINAL FLOWER 1 functions as a mobile transcriptional cofactor in the shoot apical meristem. *Plant Physiol.* 182(4), 2061–2095. doi: 10.1104/pp.19.00867

Hanano, S., and Goto, K. (2011). Arabidopsis TERMINAL FLOWER1 is involved in the regulation of flowering time and inflorescence development through transcriptional repression. *Plant Cell* 23, 3172–3184. doi: 10.1105/tpc.111.088641

Holt, A. L., van Haperen, J. M. A., Groot, E. P., and Laux, T. (2014). Signaling in shoot and flower meristems of arabidopsis thaliana. *Curr. Opin. Plant Biol.* 17, 96–102. doi: 10.1016/j.pbi.2013.11.011

Ishikawa, R., Tamaki, S., Yokoi, S., Inagaki, N., Shinomura, T., Takano, M., et al. (2005). Suppression of the floral activator Hd3a is the principal cause of the night break effect in rice. *Plant Cell* 17, 3326–3336. doi: 10.1105/tpc.105.037028

Izawa, T., Oikawa, T., Sugiyama, N., Tanisaka, T., Yano, M., and Shimamoto, K. (2002). Phytochrome mediates the external light signal to repress FT orthologs in photoperiodic flowering of rice. *Genes Dev.* 16, 2006–2020. doi: 10.1101/gad.999202

Jarvis, D. E., Ho, Y. S., Lightfoot, D. J., Schmockel, S. M., Li, B., Borm, T. J., et al. (2017). The genome of chenopodium quinoa. *Nature* 542, 307–312. doi: 10.1038/nature21370

Jiang, K., Liberatore, K. L., Park, S. J., Alvarez, J. P., and Lippman, Z. B. (2013). Tomato yield heterosis is triggered by a dosage sensitivity of the florigen pathway that fine-tunes shoot architecture. *PloS Genet.* 9, e1004043. doi: 10.1371/journal.pgen.1004043

Jin, J., Tian, F., Yang, D. C., Meng, Y. Q., Kong, L., Luo, J., et al. (2017). 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.

Kaneko-Suzuki, M., Kurihara-Ishikawa, R., Okushita-Terakawa, C., Kojima, C., Nagano-Fujiwara, M., Ohki, I., et al. (2018). TFL1-like proteins in rice antagonize rice FT-like protein in inflorescence development by competition for complex formation with 14-3-3 and FD. *Plant Cell Physiol.* 59, 458–468. doi: 10.1093/pcp/pcy021

Karlgren, A., Gyllenstrand, N., Kallman, T., Sundstrom, J. F., Moore, D., Lascoux, M., et al. (2011). Evolution of the PEBP gene family in plants: functional diversification in seed plant evolution. *Plant Physiol.* 156, 1967–1977. doi: 10.1104/pp.111.176206

Kim, S. L., Lee, S., Kim, H. J., Nam, H. G., and An, G. (2008). OsMADS51 is a short-day flowering promoter that functions upstream of Ehd1, OsMADS14, and Hd3a. *Plant Physiol.* 147, 438–438. doi: 10.1104/pp.107.103291

Komiya, R., Ikegami, A., Tamaki, S., Yokoi, S., and Shimamoto, K. (2008). Hd3a and RFT1 are essential for flowering in rice. *Development* 135, 767–774. doi: 10.1242/dev.008631

Komiya, R., Yokoi, S., and Shimamoto, K. (2009). A gene network for long-day flowering activates RFT1 encoding a mobile flowering signal in rice. *Development* 136, 3443–3450. doi: 10.1242/dev.040170

Krieger, U., Lippman, Z. B., and Zamir, D. (2010). The flowering gene SINGLE FLOWER TRUSS drives heterosis for yield in tomato. *Nat. Genet.* 42, 459–463. doi: 10.1038/ng.550

Kyozuka, J., Tokunaga, H., and Yoshida, A. (2014). Control of grass inflorescence form by the fine-tuning of meristem phase change. *Curr. Opin. Plant Biol.* 17, 110–115. doi: 10.1016/j.pbi.2013.11.010

Liu, C., Teo, Z. W., Bi, Y., Song, S., Xi, W., Yang, X., et al. (2013). A conserved genetic pathway determines inflorescence architecture in arabidopsis and rice. *Dev. Cell* 24, 612–622. doi: 10.1016/j.devcel.2013.02.013

Lopez-Marques, R. L., Norrevang, A. F., Ache, P., Moog, M., Visintainer, D., Wendt, T., et al. (2020). Prospects for the accelerated improvement of the resilient crop quinoa. *J. Exp. Bot.* 71, 5333–5347. doi: 10.1093/jxb/eraa285

Meng, X., Muszynski, M. G., and Danilevskaya, O. N. (2011). The FT-like ZCN8 gene functions as a floral activator and is involved in photoperiod sensitivity in maize. *Plant Cell* 23, 942–960. doi: 10.1105/tpc.110.081406

Mulki, M. A., Bi, X., and von Korff, M. (2018). FLOWERING LOCUS T3 controls spikelet initiation but not floral development. *Plant Physiol.* 178, 1170–1186. doi: 10.1104/pp.18.00236

Nakagawa, M., Shimamoto, K., and Kyozuka, J. (2002). Overexpression of RCN1 and RCN2, rice TERMINAL FLOWER 1/CENTRORADIALIS homologs, confers delay of phase transition and altered panicle morphology in rice. *Plant j.: Cell Mol. Biol.* 29, 743–750. doi: 10.1046/j.1365-313X.2002.01255.x

Nakamura, S., Abe, F., Kawahigashi, H., Nakazono, K., Tagiri, A., Matsumoto, T., et al. (2011). A wheat homolog of MOTHER OF FT AND TFL1 acts in the regulation of germination. *Plant Cell* 23, 3215–3229. doi: 10.1105/tpc.111.088492

Navarro, C., Abelenda, J. A., Cruz-Oro, E., Cuellar, C. A., Tamaki, S., Silva, J., et al. (2011). Control of flowering and storage organ formation in potato by FLOWERING LOCUS T. *Nature* 478, 119–122. doi: 10.1038/nature10431

Patiranage, D. S. R., Asare, E., Maldonado-Taipe, N., Rey, E., Emrani, N., Tester, M., et al. (2021). Haplotype variations of major flowering time genes in quinoa unveil their role in the adaptation to different environmental conditions. *Plant. Cell Environ* 44, 2565–2579. doi: 10.1111/pce.14071

Pin, P. A., Benlloch, R., Bonnet, D., Wremerth-Weich, E., Kraft, T., Gielen, J. J., et al. (2010). An antagonistic pair of FT homologs mediates the control of flowering time in sugar beet. *Science* 330, 1397–1400. doi: 10.1126/science.1197004

Putterill, J., Robson, F., Lee, K., Simon, R., and Coupland, G. (1995). The CONSTANS gene of arabidopsis promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. *Cell* 80, 847–857. doi: 10.1016/0092-8674(95)90288-0

Ryu, C.-H., Lee, S., Cho, L.-H., Kim, S. L., Lee, Y.-S., Choi, S. C., et al. (2009). OsMADS50 and OsMADS56 function antagonistically in regulating long day (LD)-dependent flowering in rice. *Plant. Cell Environ.* 32, 1412–1427. doi: 10.1111/j.1365-3040.2009.02008.x

Song, Y. H., Shim, J. S., Kinmonth-Schultz, H. A., and Imaizumi, T. (2015). Photoperiodic flowering: time measurement mechanisms in leaves. *Annu. Rev. Plant Biol.* 66, 441–464. doi: 10.1146/annurev-arplant-043014-115555

Song, S., Wang, G., Wu, H., Fan, X., Liang, L., Zhao, H., et al. (2020). OsMFT2 is involved in the regulation of ABA signaling-mediated seed germination through interacting with OsbZIP23/66/72 in rice. *Plant j.: Cell Mol. Biol.* 103, 532–546. doi: 10.1111/tpj.14748

Štorchová, H. (2020). "The evolution of the FLOWERING LOCUS T-like (FTL) genes in the goosefoot subfamily chenopodioideae," in *Evolutionary biology–a transdisciplinary approach*. Ed. P. Pontarotti (Cham: Springer International Publishing), 325–335.

Storchova, H., Hubackova, H., Abeyawardana, O. A. J., Walterova, J., Vondrakova, Z., Eliasova, K., et al. (2019). Chenopodium ficifolium flowers under long days without upregulation of FLOWERING LOCUS T (FT) homologs. *Planta* 250, 2111–2125. doi: 10.1007/s00425-019-03285-1

Takeshima, R., Nan, H., Harigai, K., Dong, L., Zhu, J., Lu, S., et al. (2019). Functional divergence between soybean FLOWERING LOCUS T orthologues FT2a and FT5a in post-flowering stem growth. *J. Exp. Bot.* 70, 3941–3953. doi: 10.1093/jxb/erz199

Tamaki, S., Matsuo, S., Wong, H. L., Yokoi, S., and Shimamoto, K. (2007). Hd3a protein is a mobile flowering signal in rice. *Science* 316, 1033–1036. doi: 10.1126/science.1141753

Tamura, K., Stecher, G., and Kumar, S. (2021). MEGA11 molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* 38, 3022–3027. doi: 10.1093/molbev/msab120

Taoka, K., Ohki, I., Tsuji, H., Kojima, C., and Shimamoto, K. (2013). Structure and function of florigen and the receptor complex. *Trends Plant Sci.* 18, 287–294. doi: 10.1016/j.tplants.2013.02.002

Teo, Z. W., Song, S., Wang, Y. Q., Liu, J., and Yu, H. (2014). New insights into the regulation of inflorescence architecture. *Trends Plant Sci.* 19, 158–165. doi: 10.1016/j.tplants.2013.11.001

Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2002). "Multiple sequence alignment using ClustalW and ClustalX," in *Current protocols in bioinformatics* 1, 2.3.1-2.3.22. doi: 10.1002/0471250953.bi0203s00

Tsuji, H., Tachibana, C., Tamaki, S., Taoka, K., Kyozuka, J., and Shimamoto, K. (2015). Hd3a promotes lateral branching in rice. *Plant j.: Cell Mol. Biol.* 82, 256–266. doi: 10.1111/tpj.12811

Tsuji, H., Taoka, K.-I., and Shimamoto, K. (2013). Florigen in rice: complex gene network for florigen transcription, florigen activation complex, and multiple functions. *Curr. Opin. Plant Biol.* 16, 228–235. doi: 10.1016/j.pbi.2013.01.005

Wang, Y., Yu, H., Tian, C., Sajjad, M., Gao, C., Tong, Y., et al. (2017). Transcriptome association identifies regulators of wheat spike architecture. *Plant Physiol.* 175, 746–757. doi: 10.1104/pp.17.00694

Wang, Z., Zhou, Z., Liu, Y., Liu, T., Li, Q., Ji, Y., et al. (2015). Functional evolution of phosphatidylethanolamine binding proteins in soybean and arabidopsis. *Plant Cell* 27, 323–336. doi: 10.1105/tpc.114.135103

Wickland, D. P., and Hanzawa, Y. (2015). The FLOWERING LOCUS T/TERMINAL FLOWER 1 gene family: Functional evolution and molecular mechanisms. *Mol. Plant* 8, 983–997. doi: 10.1016/j.molp.2015.01.007

Wu, Q., Bai, X., Wu, X. Y., Xiang, D. B., Wan, Y., Luo, Y. M., et al. (2020). Transcriptome profiling identifies transcription factors and key homologs involved in seed dormancy and germination regulation of chenopodium quinoa. *Plant Physiol. Biochem.* 151, 443–456. doi: 10.1016/j.plaphy.2020.03.050

Wu, Q., Bai, X., Zhao, W., Shi, X., Xiang, D., Wan, Y., et al. (2019). Investigation into the underlying regulatory mechanisms shaping inflorescence architecture in chenopodium quinoa. *BMC Genomics* 20, 658. doi: 10.1186/s12864-019-6027-0

Wu, Q., Luo, Y., Wu, X., Bai, X., Ye, X., Liu, C., et al. (2021). Identification of the specific long-noncoding RNAs involved in night-break mediated flowering retardation in chenopodium quinoa. *BMC Genomics* 22, 284. doi: 10.1186/s12864-021-07605-2

Zou, C., Chen, A., Xiao, L., Muller, H. M., Ache, P., Haberer, G., et al. (2017). A high-quality genome assembly of quinoa provides insights into the molecular basis of salt bladder-based salinity tolerance and the exceptional nutritional value. *Cell Res.* 27, 1327–1340. doi: 10.1038/cr.2017.124

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 21 February 2023
DOI 10.3389/fpls.2023.1140727

OPEN ACCESS

EDITED BY
Meng Kou,
Sweet Potato Research Institute (CAAS), China

REVIEWED BY
Lingyun Yuan,
Anhui Agricultural University, China
Avneesh Kumar,
Akal University, India
Yuan Li,
China Agricultural University, China

*CORRESPONDENCE
Mingku Zhu
mingkuzhu007@126.com
Xiaoqing Meng
xiaoqingmeng008@126.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Crop and Product Physiology, a section of the journal Frontiers in Plant Science

RECEIVED 09 January 2023
ACCEPTED 10 February 2023
PUBLISHED 21 February 2023

CITATION
Zhang C, Dong T, Yu J, Hong H, Liu S, Guo F, Ma H, Zhang J, Zhu M and Meng X (2023) Genome-wide survey and expression analysis of Dof transcription factor family in sweetpotato shed light on their promising functions in stress tolerance.
Front. Plant Sci. 14:1140727.
doi: 10.3389/fpls.2023.1140727

Genome-wide survey and expression analysis of Dof transcription factor family in sweetpotato shed light on their promising functions in stress tolerance

Chengbin Zhang[1†], Tingting Dong[1†], Jing Yu[1†], Haiting Hong[1], Siyuan Liu[1], Fen Guo[1], Hongting Ma[1], Jianling Zhang[2], Mingku Zhu[1*] and Xiaoqing Meng[1*]

[1]Institute of Integrative Plant Biology, School of Life Sciences, Jiangsu Normal University, Xuzhou, China, [2]Laboratory of Plant Germplasm Innovation and Utilization, School of Life Sciences, Liaocheng University, Liaocheng, China

DNA-binding with one finger (Dof) transcription factors play a crucial role in plant abiotic stress regulatory networks, although massive Dofs have been systematically characterized in plants, they have not been identified in the hexaploid crop sweetpotato. Herein, 43 *IbDof* genes were detected to be disproportionally dispersed across 14 of the 15 chromosomes of sweetpotato, and segmental duplications were discovered to be the major driving force for the expansion of *IbDofs*. The collinearity analysis of *IbDofs* with their related orthologs from eight plants revealed the potential evolutionary history of Dof gene family. Phylogenetic analysis displayed that IbDof proteins were assigned into nine subfamilies, and the regularity of gene structures and conserved motifs was consistent with the subgroup classification. Additionally, five chosen *IbDof* genes were shown to be substantially and variably induced under various abiotic conditions (salt, drought, heat, and cold), as well as hormone treatments (ABA and SA), according to their transcriptome data and qRT-PCR experiments. Consistently, the promoters of *IbDofs* contained a number of cis-acting elements associated with hormone and stress responses. Besides, it was noted that IbDof2 had transactivation activity in yeasts, while IbDof-11/-16/-36 did not, and protein interaction network analysis and yeast two-hybrid experiments revealed a complicated interaction connection amongst IbDofs. Collectively, these data lay a foundation for further functional explorations of *IbDof* genes, especially with regards to the possible application of multiple IbDof members in breeding the tolerant plants.

KEYWORDS

abiotic stress, Dof transcription factor, expression analysis, molecular characterization, sweetpotato

Introduction

Sweetpotato is an important root tuber crop with effective output of more than 100 million tons per annum. In terms of starch staple food, sweetpotato is second only to cassava in number in East Africa (Ssamula et al., 2020). Asia accounts not only for more than 50% of sweetpotato-producing region in the world, but also for approximately 80% of the world production (Liu, 2017). In developed nations, sweetpotato was initially grown for fresh consumption or preserved food, while the market for sweetpotato as a bio-based industrial food or value-added products is growing by degrees. Additionally, sweetpotato can also be used as suitable raw materials for industrial productions, such as for the production of bioenergy, including ethanol and butanol (Ziska et al., 2009; Jin et al., 2022). Nevertheless, adverse abiotic stresses such as salt, drought and cold severely limit crop growth, development and yield (Corrales et al., 2017; Meng et al., 2023). Genetic engineering is widely proven to be a powerful way to improve crop resistance, among them, transcription factors (TFs) are promising and efficient members to solve the problems encountered by plants in adversity, such as AP2/ERF, NAC, WRKY, MYB, GRAS and Dof TF families (Gupta et al., 2015; Erpen et al., 2018; Ichimaru et al., 2022; Zhang et al., 2022).

Typically, plant-specific Dof TFs contain a conserved Dof domain located at the N-terminal region with a highly conserved zinc finger motif, which can specifically bind to DNA sequences with a 5'-AAAG-3' core in the promoters of the target genes. Additionally, the C-terminal of Dof TFs include a highly variable transcriptional regulation region and a nuclear localization signal (Gupta et al., 2015; Waqas et al., 2020). Dof domain consists of 50–52 amino acids with four conserved cysteines ($CX_2CX_{21}CX_2C$), which can form a single zinc finger, hence it is called DNA-binding with one finger (Waqas et al., 2020). The Dof domain is considered as a bifunctional region: DNA-protein and protein-protein interactions (Krohn et al., 2002; Gupta et al., 2015). The first Dof gene, *ZmDof1*, was found in maize (Yanagisawa and Izui, 1993), and gradually expended into diverse taxonomic groups ranging from ferns and mosses to vascular plants (Kushwaha et al., 2011; Lohani et al., 2021). Presently, genome-wide isolation of Dof TFs are emerging from sundry plants, for example, 30, 30, 119, 74 and 114 Dofs were found in monocots such as *Oryza sativa* (Khan et al., 2021), *Sorghum bicolor* (Xiao et al., 2022), *Saccharum officinarum* (Gupta et al., 2014), *Musa nana* (Dong et al., 2016), and *Gossypium hirsutum* (Li et al., 2018), respectively. And there are 36, 117, 33, 34, 33 and 29 Dof members were found in eudicots like *Arabidopsis thaliana* (Lijavetzky et al., 2003), *Brassica napus* (Lohani et al., 2021), *Piper Nigrum* (Kang et al., 2016), *Solanum lycopersicum* (Cai et al., 2013), *Capsicum annuum* (Wu et al., 2016), and *Solanum melongena* (Wei et al., 2018), respectively.

Dof proteins have been substantially evidenced to partake gene expression in diverse processes such as seed maturation and germination, seedling development, light response, carbon and nitrogen metabolism, leaf axial patterning, as well as metabolic pathways involving hormones such as salicylic acid and response (Gupta et al., 2015; Ruta et al., 2020). For instance, Dof TFs VASCULAR-RELATED DOF1 (VDOF1) and VDOF2 were reported to regulate vascular cell differentiation and lignin biosynthesis in Arabidopsis (Ramachandran et al., 2020). The Arabidopsis Dof TF CDF3 was involved in nitrogen responses and improved nitrogen use efficiency in transgenic tomato (Corrales et al., 2017). Additionally, multiple studies have revealed that Dof TFs participate in the response to various environmental changes (Corrales et al., 2017; Waqas et al., 2020). For example, the transcription of tomato *SlCDF1–5* (Cycling Dof Factors) genes was significantly induced under osmotic, salt, heat, and low-temperature stresses, overexpression of *SlCDF1* or *SlCDF3* in Arabidopsis both showed improved drought and salt tolerance (Corrales et al., 2014). Rice *OsDOF15* was found to be a salt-responsive gene which participated in salt-induced inhibition of primary root growth (Qin et al., 2019). The expression of Arabidopsis *CDF3* is highly induced by drought, extreme temperature and ABA treatments. The *cdf3* mutant is more sensitive to drought and cold stress, whereas enhanced tolerance to drought, cold and osmotic stress was detected in *CDF3*-overexpressing plants (Corrales et al., 2017). Besides, *Tamarix hispida* ThDof TF functions as an upstream regulator of TheIF1A that plays a positive factor in plant salt and osmotic stress tolerance *via* the regulation of associated enzymes and ROS scavenging (Yang et al., 2017a). And *Juglans regia* JrDof3 TF also contributed to improve the high temperature stress response of JrGRAS2, which could effectively regulate the expression of *HSPs* to enhance high temperature stress tolerance (Yang et al., 2018). Similarly, grape *VaDof17d* gene was recently reported to be positively correlated with cold tolerance (Wang et al., 2021). Therefore, Dof TFs are one of the pivotal factors to enhance abiotic stress tolerance in plants by genetic engineering. Nevertheless, despite this, the specific roles of most Dof TFs in plants remain unknown, and further studies are still imperative to explore the regulatory pathways for abiotic stress tolerance through Dof TFs.

The recent advancements in plant genomic sequencing have significantly accelerated the identification of Dof genes in various plants that will expand our understandings of their potential functions in plant stress response. However, the genome-wide structure and function research of most Dofs remain to be elucidated, especially in the important hexaploid crop sweetpotato. Although the genome of hexaploid sweetpotato has been sequenced more than 5 years (Yang et al., 2017b), the identification and functional analysis of the sweetpotato Dof TF family is still blank. Here, genome-wide and transcriptome-based analysis of *IbDof* genes in sweetpotato were accomplished, including chromosomal location, motif composition, gene structure, conserved domain, collinearity analysis and protein-protein interaction. And the expression analysis of *IbDof* genes under various abiotic stresses and hormones shed light on their promising functions in stress tolerance. The information gained here will be significant for laying the groundwork for the functional characterization of multiple valuable sweetpotato IbDof TFs in response to abiotic stress.

Materials and methods

Identification of putative *IbDof* genes in sweetpotato genomes

The completed genome sequence and GFF annotation data of sweetpotato (*Ipomoea batatas*, Taizhong6) were obtained from Ipomoea Genome Hub database (https://ipomoea-genome.org) as described in our previous report (Liu et al., 2022). The nucleotide and protein sequences of 36 Dof TFs of *Arabidopsis thaliana* and 30 Dof TFs of *Oryza sativa* (Supplementary Table S1) were downloaded from TAIR (https://www.arabidopsis.org/) and Rice Genome Annotation Project (http://rice.uga.edu/), respectively. Then the acquired Dof protein sequences of rice and Arabidopsis were used as query sequences for BLASTP program (E-value ≤ 1e-5) with 500 MumofHits and 250 MumoAligns to search all the possible Dof members in sweetpotato by Tbtools (Chen et al., 2020). Ultimately, 44 non-redundant protein sequences were obtained, and the Pfam database (http://pfam.xfam.org/), online batch CD-search program (https://www.ncbi.nlm.nih.gov/cdd/Structure/cdd/wrpsb.cgi) and PROSITE database (https://prosite.expasy.org/) were used to exclude the proteins without the Dof domain. All the sequence information of 43 candidate *IbDof* genes was listed in Supplementary Table S2.

Analysis of physicochemical characteristics of IbDof proteins

The ExPASy tool (http://expasy.org/) was used to predict the specific isoelectric point and molecular weight of each IbDof protein. The phosphorylation sites of IbDofs were detected using NetPhos 3.1 (http://www.cbs.dtu.dk/services/NetPhos/). To estimate their subcellular locations, the Plant-mPLoc software (http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/) was used. Tbtools software was used to create the intron-exon architecture of the *IbDof* genes by matching their coding sequences and genomic sequences (Chen et al., 2020). In order to identify any potential cis-elements related to hormones and stresses, the 2.0 kb promoter sections of 43 *IbDof* genes were taken from the Ipomoea Genome Hub and were uploaded to the plantCARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) for detection.

Analysis of phylogenetic relationship, conserved domain, and protein interaction network

The amino acid sequences of 43 IbDofs in sweetpotato and the well-classified AtDof proteins in Arabidopsis and OsDof proteins in rice (Lijavetzky et al., 2003) were utilized to build the unrooted phylogenetic tree using the Maximum Likelihood method by MEGA-X software (Kumar et al., 2018). All the Dof protein sequences were first aligned by ClustalW with default settings, then the following parameters were used to generate the phylogenetic relationship of 43 IbDofs: the best evolutionary model of JTT + G + I + F, a bootstrap value of 1000 with partial deletions. Based on the previously defined parameters in Arabidopsis and rice, MEME 5.3.3 software (https://meme-suite.org/meme/tools/meme) was used to construct the conserved motifs with a maximum number of 15, and a length of > 6 and < 200 amino acids (Khan et al., 2021). The possible protein interaction network was carried out using STRING 11.0 (https://string-db.org/).

Chromosome position and collinearity analysis of *IbDof* genes in sweetpotato

According to the GFF annotations of the sweetpotato genomes, 43 *IbDof* genes were linked with the chromosomes. In order to conduct a synteny analysis between *IbDofs* and genes from other plant species, the genome sequence and annotation data of sweetpotato and eight representative species (including *Ipomoea triloba*, *Ipomoea trifida*, Arabidopsis, rice, tomato, pepper, cabbage, and *Brassica oleracea*), were downloaded from various databases such as Ipomoea Genome Hub, TAIR, Ensembl (http://plants.ensembl.org/index.html), and Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html). MCScanX software was used to evaluate the association of gene duplication and collinearity using default settings. Circos and TBtools softwares were used to display the obtained results, and the minimum block size setting is 30 (Krzywinski et al., 2009; Chen et al., 2020; Guo et al., 2022).

Screening of salt-responsive *IbDof* genes based on RNA-seq analysis

After treatment with 150 mM NaCl for 24 hours, the adventitious roots of two substantially different sweetpotato cultivars XuShu 22 (salt-tolerant) and XuShu 32 (salt-sensitive) were collected for transcriptome analysis. Read counts were used to measure gene expression and differentially expressed genes were screened by false discovery rate (FDR) and Log2 (fold change) as previously described (Meng et al., 2020b). Then gene functions were annotated based on the information from several databases such as the sweetpotato reference genome database, Nr, Pfam, and SwissProt. The salt-stress responsive or genotype-specific *IbDof* genes are listed in Supplementary Table S3.

Abiotic stress and hormone treatments of sweetpotato seedlings and qRT-PCR analysis

The salt-tolerant XuShu 22 seedlings were grown in 1/4 Hoagland solutions and used for all the stress and hormone treatments with three different biological replicates. The roots were immersed in 150 mM NaCl and 20% PEG6000 solutions,

respectively, for salt and dehydration stress treatment, then root samples were collected. Seedlings were incubated at 4°C and 42°C, respectively, for cold and heat conditions, then leaves samples were harvested. And the leaves of seedlings were collected after spraying with 0.1 mM ABA or 2 mM SA solutions for phytohormone tests (Meng et al., 2018). All of the samples were collected at 0, 1, 12, 24, and 48 hours after each treatment.

Total RNA was extracted using RNA Extraction Kits (TianGen, Beijing, China) following the manufacturer's instructions. PrimeScript reverse transcriptase with the gDNA Eraser (TaKaRa, Dalian, China) was used to reverse-transcribe 1 μg of each RNA sample. qRT-PCR tests were conducted using the CFX96TM Real-Time System as previously described (Bio-Rad, USA) (Meng et al., 2020a), and the sweetpotato *ARF* gene (JX177359) was employed as an internal control (Park et al., 2012). All the qRT-PCR primers are provided in Supplementary Table S4.

Analysis of transactivation activity and protein interaction of IbDof proteins in yeasts

The open reading frame sequences of *IbDof-2/-11/-16/-36* genes were independently cloned into the pGBKT7 and pGADT7 vector, respectively, using the homologous recombination method. Then Y2HGold yeasts were transformed with the pGBKT7 control, recombined pGBKT7-IbDof plasmids, or both recombined pGBKT7-IbDof and pGADT7-IbDof vectors, respectively, as previously described (Zhang et al., 2022). The yeast dilution was dropped on SD/-Trp (SDO), SD/-Trp-His-Ade (TDO) medium with or without 200 ng/mL AbA (Aureobasidin A) for the transactivation detection, and the dilutions were dropped on SD/-Trp-Leu (DDO), SD/-Trp-Leu-His-Ade (QDO) medium with or without 200 ng/mL AbA for the purpose of detecting protein interactions. All of the plates were grown at 30°C to test for transactivation or protein interaction for three days. The primers used for gene cloning and vector construction are shown in Supplementary Table S5.

Statistical analysis

The experiments were performed in three biological replicates. A cut-off value of two-fold for differential gene expression was established in consideration of the biological importance (Zhang et al., 2022). Graphs were produced using the SAS Institute's OriginPro 8 program.

Results

Identification and characterization of the Dof gene family in sweetpotato genome

In this study, the BLASTP program was used to screen all the potential sweetpotato Dofs using the known Dof proteins from Arabidopsis and rice as query sequences. One of the proteins was, however, disqualified because it lacked the Dof domain. The remaining 43 genes were named as *IbDof1* to *IbDof43* according to the location order of sweetpotato chromosomes from top to bottom. Subsequently, the amino acid (aa) length, molecular weight (Mw), and

TABLE 1 Characteristics of *IbDof* genes in *Ipomoea batatas*.

Gene name	Gene ID	Length	MW (Da)	PI	Subcellular location	No. of phosphorylation sites			
						Ser sites	Tyr sites	Thr sites	Total
IbDof1	g563.t1	269	29567.12	5.03	Nucleus	16	4	9	29
IbDof2	g909.t1	275	29902.17	7.17	Nucleus	26	4	6	36
IbDof3	g1131.t1	262	27645.86	9.18	Nucleus	26	3	7	36
IbDof4	g1308.t1	529	58136.22	6.19	Nucleus	41	2	13	56
IbDof5	g2786.t1	287	31077.49	6.19	Nucleus	18	2	13	33
IbDof6	g4774.t1	282	30986.35	6.78	Nucleus	21	7	7	35
IbDof7	g4854.t1	319	33797.46	9.2	Nucleus	27	2	13	42
IbDof8	g4893.t1	320	33861.58	9.2	Nucleus	27	3	14	44
IbDof9	g10108.t1	268	29426.78	8.22	Nucleus	17	5	10	32
IbDof10	g12187.t1	469	51472.98	9.58	Nucleus	32	4	6	42
IbDof11	g12244.t1	439	47269.37	7.11	Nucleus	48	5	15	68
IbDof12	g17872.t1	227	24647.6	9.25	Nucleus	16	1	6	23
IbDof13	g19058.t1	110	12439.99	9.66	Nucleus	7	3	2	12

(Continued)

10.3389/fpls.2023.1140727

TABLE 1 Continued

Gene name	Gene ID	Length	MW (Da)	PI	Subcellular location	No. of phosphorylation sites			
						Ser sites	Tyr sites	Thr sites	Total
IbDof14	g19324.t1	337	36307.7	8.21	Nucleus	43	5	7	55
IbDof15	g19438.t1	336	36213.6	8.21	Nucleus	41	5	5	51
IbDof16	g20269.t1	313	33258.39	9.48	Nucleus	18	4	4	26
IbDof17	g24282.t1	338	35097.92	9.39	Nucleus	34	2	12	48
IbDof18	g27012.t1	253	27431.46	8.15	Nucleus	26	3	7	36
IbDof19	g29805.t1	323	34658.06	8.44	Nucleus	36	2	5	43
IbDof20	g29939.t1	959	106122.5	6.77	Nucleus	80	9	13	102
IbDof21	g30074.t1	514	55813.84	5.28	Nucleus	40	3	19	62
IbDof22	g32116.t1	466	50705.95	5.76	Nucleus	53	4	19	76
IbDof23	g33760.t1	350	38282.31	6.93	Nucleus	36	5	11	52
IbDof24	g34515.t1	279	30234.75	6.15	Nucleus	22	3	12	37
IbDof25	g34535.t1	350	37725.58	8.9	Nucleus	33	3	9	45
IbDof26	g34611.t1	174	19098.54	8.61	Nucleus	8	2	2	12
IbDof27	g34661.t1	174	19155.6	8.61	Nucleus	7	2	2	11
IbDof28	g39250.t1	219	23949	8.84	Nucleus	17	3	8	28
IbDof29	g42238.t1	269	29776.11	4.94	Nucleus	24	2	6	32
IbDof30	g42376.t1	269	29762.09	4.94	Nucleus	24	3	6	33
IbDof31	g42798.t1	802	90431.97	7.93	Nucleus	64	10	26	100
IbDof32	g45550.t1	287	30912.28	6.18	Nucleus	18	2	13	33
IbDof33	g45578.t1	297	31961.5	6.35	Nucleus	18	2	12	32
IbDof34	g48073.t1	262	28714.99	8.64	Nucleus	23	7	3	33
IbDof35	g53151.t1	239	25725.87	9.32	Nucleus	16	3	5	24
IbDof36	g53369.t1	388	42412.24	9.33	Nucleus	30	4	17	51
IbDof37	g55838.t1	324	33387.8	9.6	Nucleus	35	3	9	47
IbDof38	g59149.t1	313	34076.93	9.33	Nucleus	40	2	9	51
IbDof39	g60025.t1	388	42282.36	9.05	Nucleus	32	3	8	43
IbDof40	g60438.t1	210	22064.56	8.16	Nucleus	12	2	3	17
IbDof41	g63457.t1	166	19305.01	9.73	Nucleus	7	5	5	17
IbDof42	g63711.t1	202	22537.16	8.45	Nucleus	19	7	6	32
IbDof43	g63743.t1	192	21594.34	9.73	Nucleus	16	3	8	27

Length, Amino acid length; MW, Molecular weight; pl, Isoelectric point

theoretical isoelectric point (pI) of all IbDof proteins were analyzed (Table 1). The amino acid lengths of IbDofs range from 110 (IbDof13) to 959 (IbDof20). As a result, their MW ranges from 12.4 kDa (IbDof13) to 106.1 kDa (IbDof20), while their pI ranges from 4.94 (IbDof29) to 9.73 (IbDof43). The prediction of subcellular localization suggests that all IbDof TFs are localized in the nucleus. Besides, the IbDofs have multiple possible phosphorylation sites, ranging from 11 (IbDof27) to 102 (IbDof20), and roughly 89% of the IbDofs contain at least 20 phosphorylation sites.

Chromosome distribution of sweetpotato *IbDof* genes

Physical locations determined using the GFF3 genome annotations of sweetpotato revealed that 43 *IbDof* genes were mapped to all the 15 chromosomes (Chr) except Chr 4 (Figure 1 and Supplementary Figure S1). Among them, the most abundant *IbDof* genes are found on Chr 1 and Chr 5, both have five members. However, only one *IbDof* gene is present on Chr 6, Chr 10 and

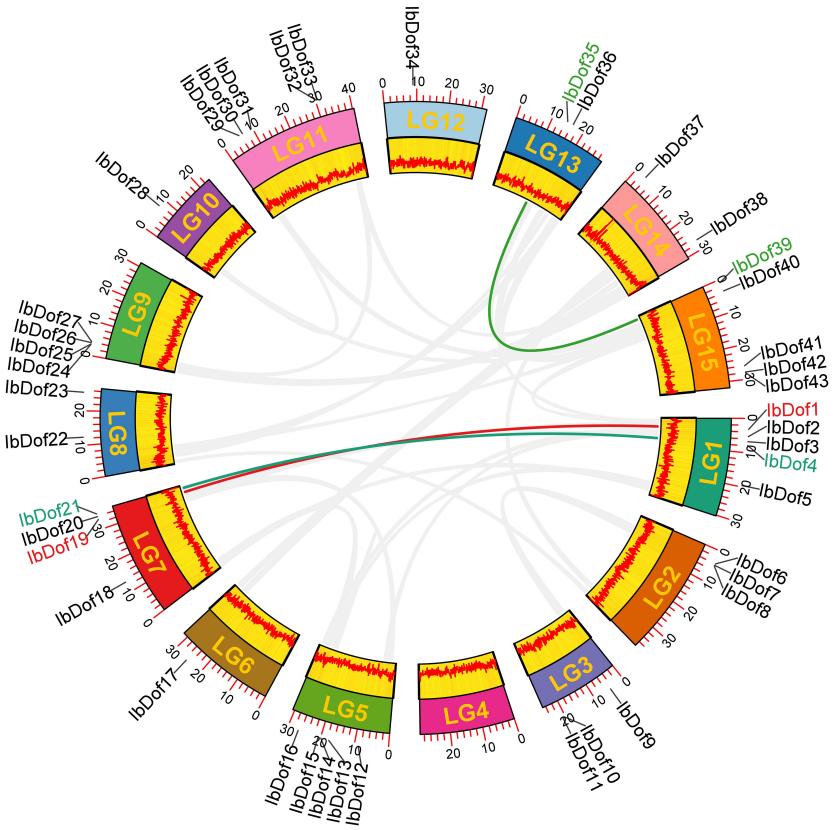


FIGURE 1
Chromosome distribution and synteny analysis of *IbDof* genes in sweetpotato chromosomes. Chromosomes LG1-LG15 are marked with colored rectangles, the location of each *IbDof* is represented by a short black line, the colored lines represent segmentally duplicated *IbDof* genes.

Chr12, and no *IbDof* gene was found on Chr 4. In the remaining chromosomes, the number of *IbDof* genes ranges from two to four. The results demonstrated that the distribution of *IbDof* gene varied greatly and was not related to chromosomal length. For instance, the short chromosome (Chr 3) has three *IbDof* genes while the long chromosome (Chr 12) only has one.

Besides, to detect the possible relationship among the 43 *IbDof* genes, the collinear analysis was conducted by the BlastP and MCScanX programs, however, no tandem repeat events were observed among the *IbDof* genes. And three gene pairs with segmental duplications were found as follows: *IbDof4-IbDof21*, *IbDof1-IbDof19*, and *IbDof35-IbDof39* (Figure 1 and Supplementary Table S6). The data suggest that segmental duplications of *IbDof* genes have a primary contribution to their expansions.

Phylogenetic relationships of IbDof proteins in sweetpotato

To examine the evolutionary connections of IbDofs in sweetpotato, a phylogenetic tree was created using the full amino acid sequences of 43 IbDofs, 36 Arabidopsis AtDofs and 30 rice OsDofs by Maximum Likelihood method using MEGA-X software. Previously, 36 AtDofs and 30 OsDofs were divided into nine subgroups (A, B1, B2, C1, C2.1, C2.2, C3, D1 and D2) (Lohani et al., 2021; Nan et al., 2021). Accordingly, 43 IbDofs were also divided into nine subgroups, and the sizes of each subgroup varied substantially as shown by the phylogenetic analysis (Figure 2). B1 and D1 subgroups had the most IbDof proteins (8 members, 18.6%), followed by A, C2.1, C2.2, and C3 subgroups (5 members, 11.6%), C1 subgroup (4 members, 9.3%), and B1 subgroup (1 member, 2.3%).

Gene structure, conserved domain and motif analysis of IbDofs

The sequence diversities of sweetpotato IbDofs were assessed through the study of exon/intron compositions and conserved domains. In line with expectations, results from Batch CD-Search revealed that all 43 IbDof proteins possess a highly conserved Dof domain (Figure 3). Besides, some of the IbDofs have the other conserved parts. For insurance, IbDof20 has two additional domains, including LINES-N superfamily and LINES-C superfamily. IbDof31 and IbDof36 comprise additional PHA03247 superfamily and MADS-SRF-like domains, respectively. Notably, the number of amino acids in the Dof domain of most IbDof members is between 50 and 60. In a similar vein, our findings demonstrated that individuals belonging to the same subgroups typically have identical

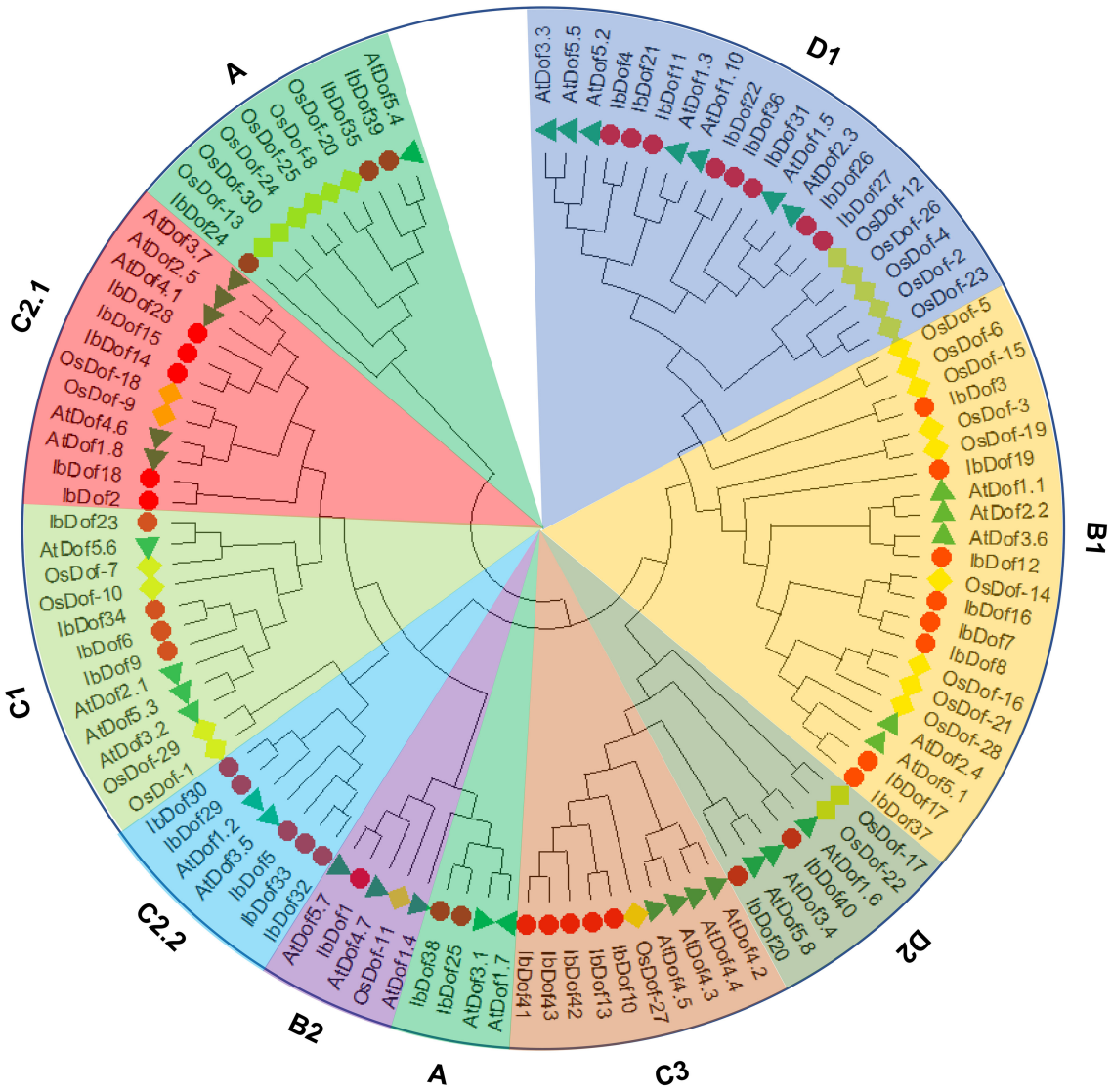


FIGURE 2
Unrooted phylogenetic tree constructed with sweetpotato IbDofs, Arabidopsis AtDofs and rice OsDofs. All the protein sequences were aligned by ClustalW and the phylogenetic relationships were derived through the Maximum Likelihood method, the best evolutionary model JTT + G + I + F calculated through MEGA X was selected with the bootstrap value of 1000. Different subgroups are named based on the reports in Arabidopsis and rice and are distinguished with different colors. The protein names are marked at the end of the branch, the red circle, green triangle and blue rhombus represent the sweetpotato IbDofs, Arabidopsis AtDofs, and rice OsDofs, respectively.

gene structures, even sharing similar exon/intron lengths (Figure 3). For example, most members of the B1 subgroup had only one intron, while multiple members from C3, D1, and D2 subgroups had the largest number of introns.

In order to better analyze the sequence diversifications of IbDofs, the compositions of conserved motif were further examined. A total of 15 different conserved motifs were identified, with the N- or C-termini containing more motifs. The IbDofs within the same subgroups have comparable motif compositions, supporting the subgroup categorization, and three IbDofs (IbDof5, IbDof32, and IbDof33) belonging to the C2.2 subfamily have the most motifs (Figure 4). Motif 1 and motif 2 represented the conserved basic Dof domain found in practically all IbDofs, except for IbDof19 that only included motif 1.

Collinearity studies of Dof genes between sweetpotato and other plants

To further infer the origin and evolutionary mechanisms of sweetpotato *IbDof* genes, the comparative syntenic relationships between 43 *IbDofs* and the related genes from eight representative species were investigated. These species included two likely diploid wild relative of sweetpotato (*Ipomoea triloba* and *Ipomoea trifida*),

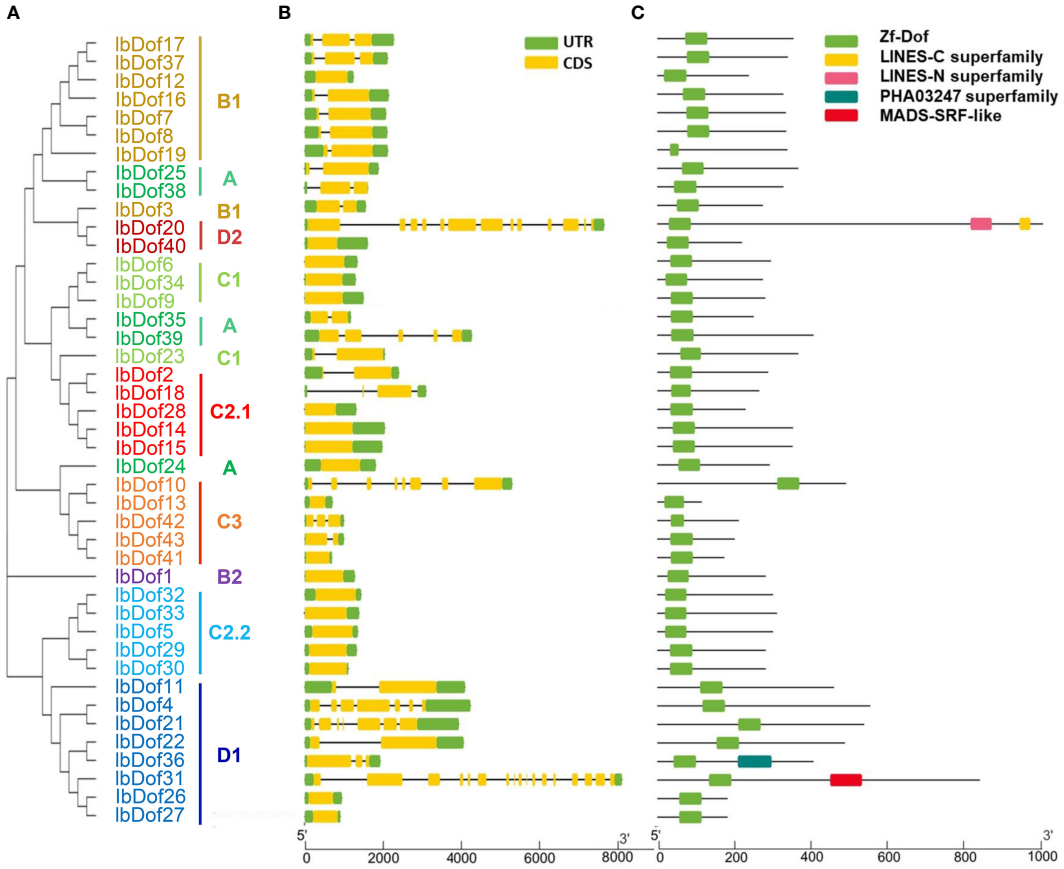


FIGURE 3
Phylogenetic relationships, gene structures and conserved domain distributions of 43 IbDof TFs in sweetpotato. **(A)** The phylogenetic tree of 43 IbDofs was constructed by MEGA X based on the consistent parameters used in Figure 2. **(B)** Gene structures of 43 *IbDof* genes. Exons and UTR are marked using yellow and green bars, respectively, black lines indicate introns. **(C)** Distributions of conserved domains detected by CD-search in the IbDof members. The colorful boxes present different conserved domains. The length of IbDofs can be estimated by the scale at the bottom.

two most representative model plants (*Arabidopsis thaliana* and *Oryza sativa*), two Solanaceae plants (*Solanum lycopersicum* and *Capsicum annuum*), two Brassica plants (*Brassica rapa* and *Brassica oleracea*). 39 *IbDof* genes (90.7%) showed syntenic connections with *Ipomoea trfida* and *Ipomoea triloba*, followed by *Solanum lycopersicum* (18), *Capsicum annuum* (8), *Arabidopsis thaliana* (3), *Brassica rapa* (2), and *Brassica oleracea* (1). However, the cereal plant *Oryza sativa* did not share any such orthologous genes with the sweetpotato *IbDofs* (Figure 5). Notably, the collinearity between *IbDofs* and genes from *Ipomoea triloba* and *Ipomoea trifida* was significantly stronger than that found in the other six species, which may be connected to the fact that both *Ipomoea triloba* and *Ipomoea trifida* are the likely diploid wild relatives of sweetpotato.

Additionally, we discovered that 23 genes from *Ipomoea triloba* and 22 genes from *Ipomoea trifida* had a collinearity connection with two or more sweetpotato *IbDof* genes. Examples of such genes are itf01g25840.t1-*IbDof-11/-21/-31*, itf04g29810.t1-*IbDof-7/-8/-16/-17*, itb12g09820.t1-*IbDof-8/-12/-16/-17*, and itb04g29250.t1-*IbDof-7/-8/-12/-16/-17*. Besides, some sweetpotato *IbDof* genes exhibits interesting collinearity with multiple detected species, for instance, *IbDof8* was found to be collinear with three transcripts, *Arabidopsis thaliana* AT1G07640.3, *Brassica rapa* Bra035667.1, and *Brassica oleracea* Bo4g166830.1, indicating that they might be derived from a common ancestor (Supplementary Table S7).

Transcriptome-wide identification of salt-responsive *IbDof* genes and their expression profiles under abiotic stress and hormone treatments

It has been found that many Dof TFs regulated different abiotic stressors (Waqas et al., 2020). In order to explore whether the *IbDof* genes identified in sweetpotato responded to abiotic stress, our previous transcriptome data were used to examine their expression levels under salt stress treatment in salt-tolerant and salt-sensitive sweetpotato cultivars. We found that a few *IbDof* genes were genotype-specific or salt stress-responsive. Then five differentially expressed *IbDof* genes, including *IbDof2*, *IbDof11*, *IbDof16*, *IbDof22*, and *IbDof36*, were selected to verify their expression profiles under various abiotic stress and hormone treatments. Taking into account the biological significance, a two-fold cut-off value of the expression difference was adopted (Zhu et al., 2015; Zhang et al., 2022).

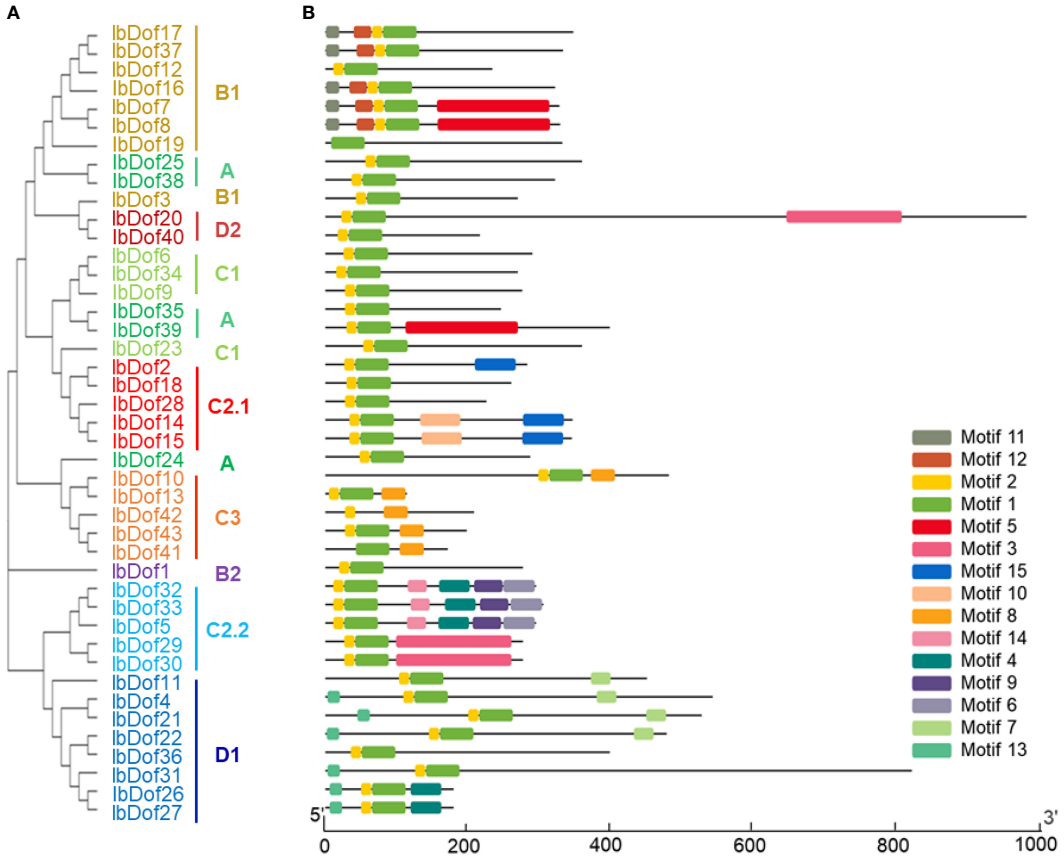


FIGURE 4
Phylogenetic relationships and conserved motif compositions in sweetpotato IbDofs. **(A)** The phylogenetic tree of 43 IbDofs was constructed by MEGA X based on the consistent parameters used in Figure 2. **(B)** Conserved motif compositions detected by MEME analysis within sweetpotato IbDofs. Boxes of different colors present different motifs, and the yellow and green boxes contained in IbDofs represent the Dof domain. The length of IbDofs can be estimated by the scale at the bottom.

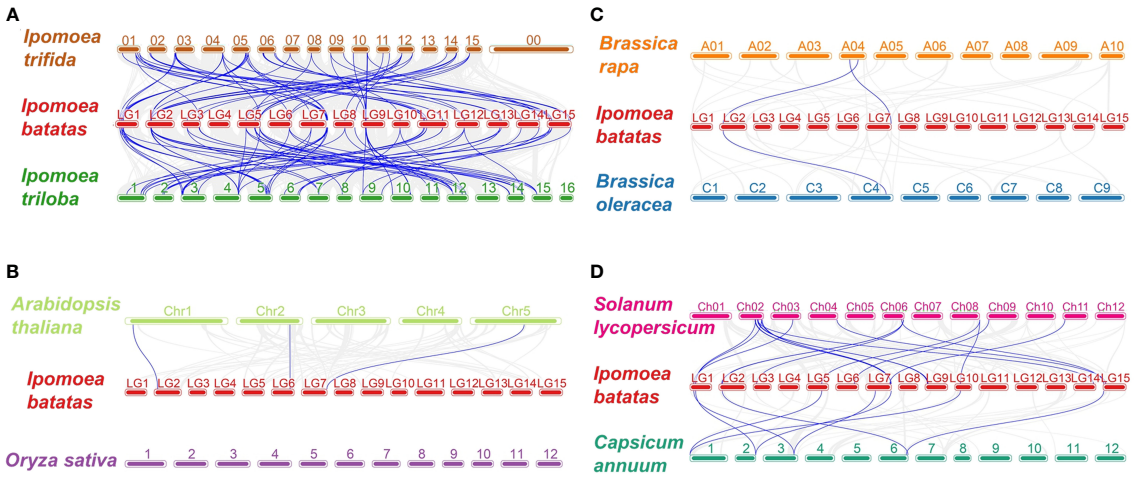


FIGURE 5
Synteny analyses of Dof genes between sweetpotato and eight representative plant species. **(A)** *Ipomoea trifida* and *Ipomoea triloba*, **(B)** *Arabidopsis thalianaa* and *Oryza sativa*, **(C)** *Brassica rapa* and *Brassica oleracea*, **(D)** *Solanum lycopersicum* and *Capsicum annuum*. The chromosomes of different plants are distinguished with differential colors. The blue lines connecting two different chromosomes highlight the syntenic Dof gene pairs within sweetpotato and other plant genomes.

10.3389/fpls.2023.1140727

Overall, the transcription level of *IbDof-2*/*-16*/*-36* was increased by more than four times under salt stress, and the expression of *IbDof16* had the most obvious enhancement with about 10-fold changes, while their expression was downregulated to varying degrees by drought stress (Figure 6A). For cold and heat stress, the mRNA level of *IbDof2* and *IbDof16* could be upregulated with 4.8-6.7-fold after cold stress, and only the transcription of *IbDof16* was enhanced with 5.7-fold after heat stress (Figure 6B). In addition, the transcription level of *IbDof2* was also significantly induced after SA treatment, about 2.7 times that of the control (Figure 6C). Therefore, sweetpotato *IbDof* genes may act as crucial and distinct regulators in response to abiotic stress, according to their considerable and varied transcriptions.

Cis-element prediction in the promoters of *IbDof* genes

In order to comprehend the potential regulatory mechanism of *IbDof* genes in response to abiotic stress and hormone treatments, the specific cis-elements in the promoter of *IbDof* genes were

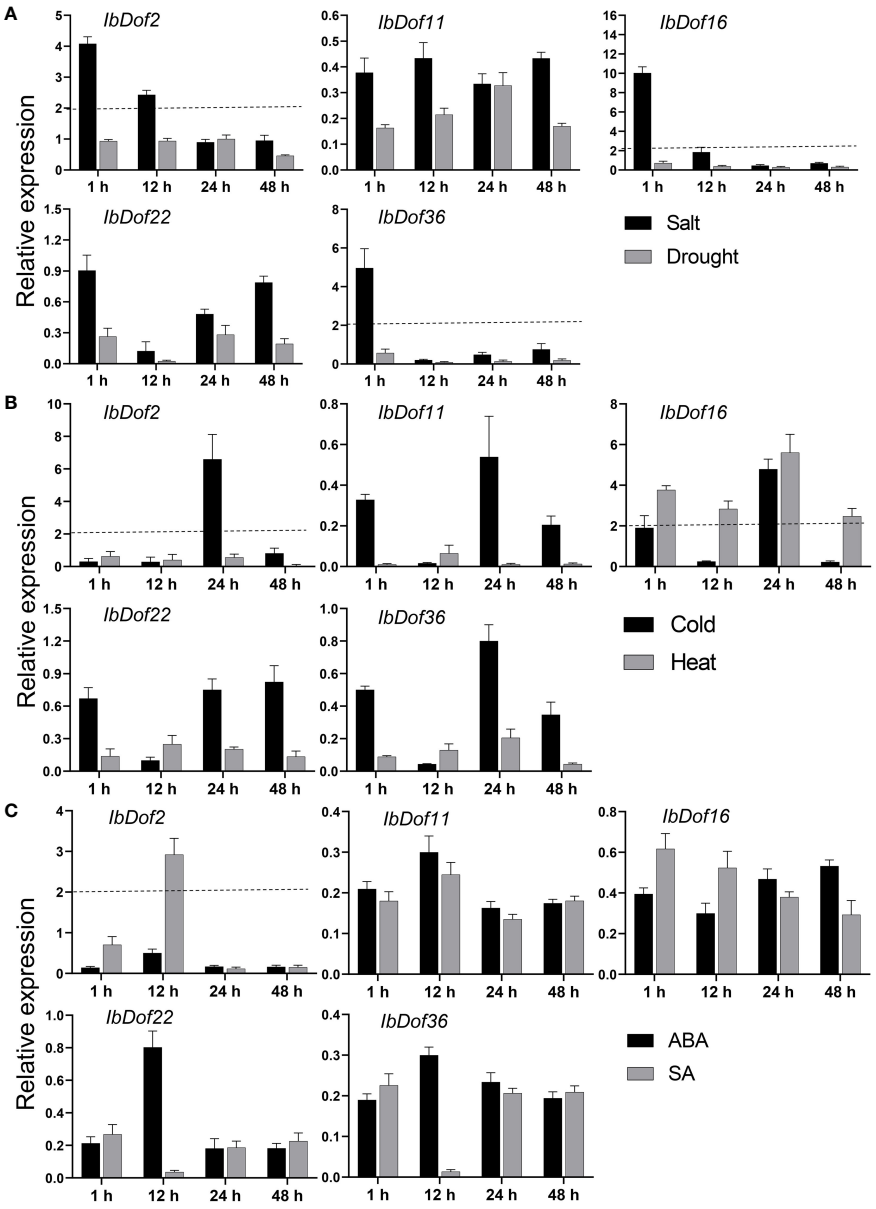


FIGURE 6
Relative expression levels of five *IbDof* genes in response to abiotic stress and hormone treatments detected by qRT-PCR. The abiotic stress treatments include salt (150 mM NaCl), drought (20% PEG6000) **(A)**, cold (4°C) and heat (42°C) **(B)**, and ABA (0.1 mM) and SA (2 mM) **(C)**. Samples were collected at 1, 12, 24, and 48 h after each treatment, the expression levels at 0 h were normalized to 1, and the Y-axis delineates the fold changes of relative expression comparing with 0 h. Bars represent the mean of three biological replicates ± SE. The two-fold threshold is presented by a dotted line.

analyzed (Figure 7). The results showed that 97% of the promoters featured numerous regulatory elements sensitive to stresses, including the drought responsive element (MBS), the defense and stress response element (TC-rich repeats), the low temperature responsive element (LTR), and the wound responsive element (WUN-motif) (Supplementary Table S8). The transcription of multiple *IbDof* genes, including *IbDof-2/-16/-36*, was significantly induced by different abiotic stresses or hormones. Accordingly, their promoters were enriched with stress-related cis-acting elements such as MBS, TC-rich repeats or LTR.

Besides, all *IbDof* promoters contained various hormone-related cis-elements, including the GA response element (GARE-motif, P-box, TATC-box), the MeJA response element (TGCCG-, TGACG- and CGTCA-motif), the auxin response element (TGA-box, AuxRR-core), and the ABA response element (ABRE). Approximately 84% of the promoters have an ABA response element. The information indicates that these cis-elements have a potential role in affecting the response of sweetpotato plants to abiotic stresses.

Interaction network analysis of the IbDof proteins in sweetpotato

The Dof domain in Dof TFs is a critical domain that can mediate protein-protein interactions (Waqas et al., 2020), indicating that IbDof TFs may also function by forming protein complexes. In order to further study the potential interaction between IbDofs, an interaction network associated with sweetpotato IbDofs was constructed based on the orthologs of Arabidopsis AtDofs by the STRING database (Figure 8). The results indicated the prospect of protein interaction between sweetpotato homologous IbDofs corresponding to the Arabidopsis AtDofs, such as TMO6 (IbDof-3/-6/-9/-24) and OBP2 (IbDof24)/OBP3 (IbDof16)/OBP4 (IbDof35 and IbDof39)/HCA2 (IbDof23). The sweetpotato IbDof interaction network indicates a potential complex association, suggesting a possible way for IbDof TFs in regulating the response of sweetpotato to environmental stresses.

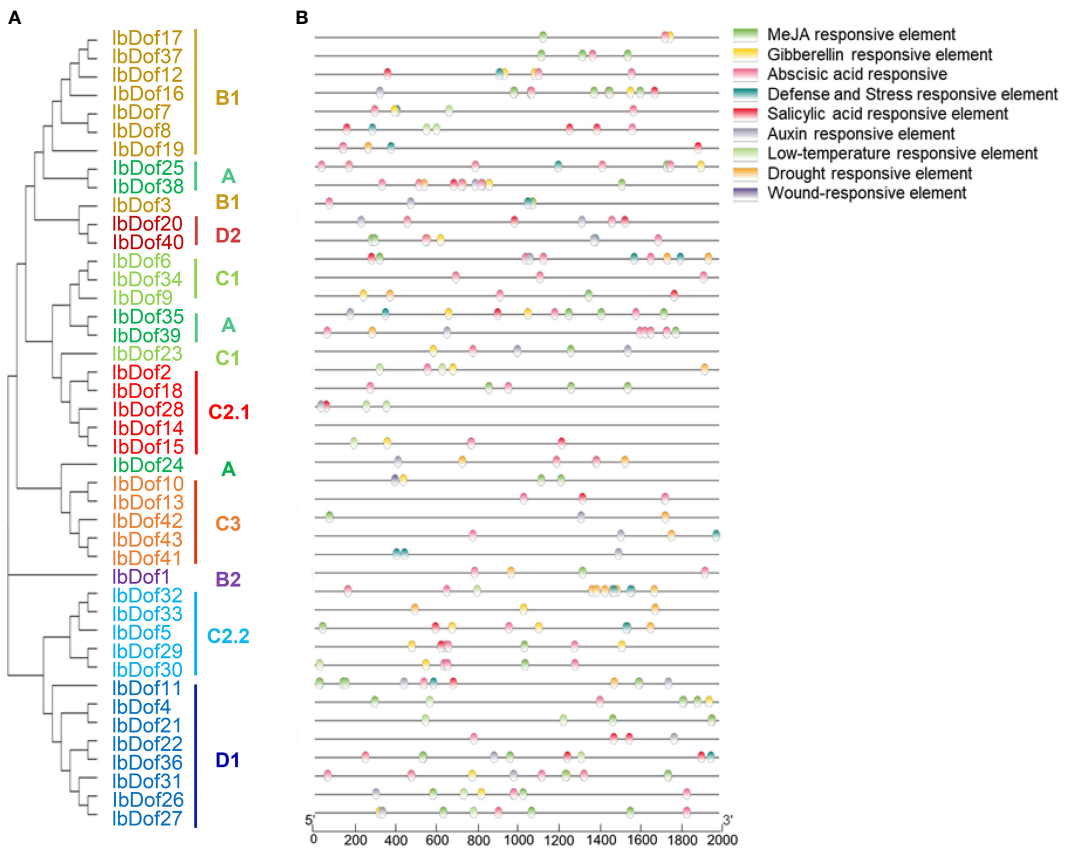


FIGURE 7
Phylogenetic clustering and predicted stress- and hormone-related cis-elements in the promoters of *IbDof* genes. **(A)** The phylogenetic tree of 43 IbDofs was constructed by MEGA X based on the consistent parameters used in Figure 2. **(B)** Predicted cis-elements in the *IbDof* promoters. 2000 bp promoter regions of each *IbDof* gene were detected by PlantCARE database. Different colored rectangles represent different cis-elements that are potentially involved in stress or hormone regulation.

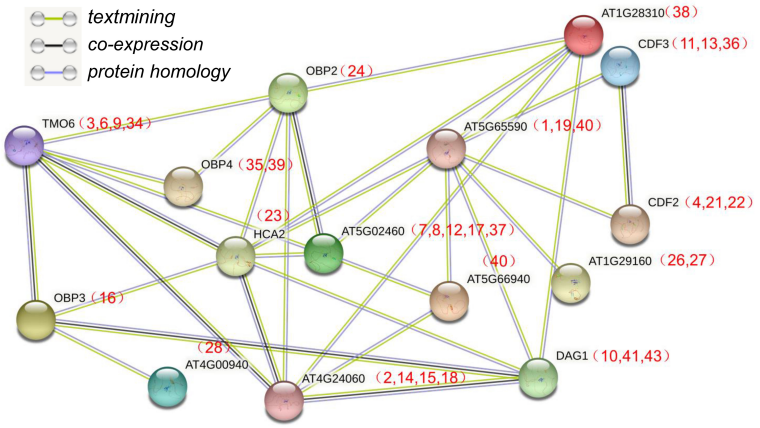


FIGURE 8
Interaction networks of IbDof proteins in sweetpotato according to the orthologues in Arabidopsis. The amino acid sequences of IbDofs were employed to search the STRING database, network node represents proteins, and edge represents protein–protein associations. The colored lines between the nodes indicate the different kinds of interactions. The numbers (*IbDof* gene number) in brackets represent the corresponding orthologues in sweetpotato.

Detection of transactivation activity and protein interaction of selected IbDofs

Considering that the transcription levels of *IbDof-2/-11/-16/-36* genes were upregulated by different abiotic stresses or hormones, they were chosen to detect possible transactivation activities in yeasts by constructing the recombinant pGBKT7 vectors. The results displayed that all the yeasts can grow normally on SDO medium. However, only the recombinant pGBKT7-*IbDof2* yeasts can survive on the TDO medium with or without AbA, while the control pGBKT7 vector and the recombinant pGBKT7-*IbDof-11/-16/-36* can not (Figure 9A). The results showed that IbDof2 TF has transactivation activity in yeasts, while IbDof-11/-16/-36 does not.

Besides, the possible interactions between any two of IbDof-2/-11/-16/-36 proteins (including the interaction with themselves) were further detected by Y2H assays, but pGBKT7-IbDof2 vector was not involved because of its self-activation activity. The results presented that the yeast transformed with both the recombinant pGADT7 vector and the recombinant or empty pGBKT7 vector could grow normally on the control DDO medium. And the results showed that IbDof16 could interact with itself in addition to IbDof2 and IbDof11, and IbDof2 could also interact with IbDof36. Other than that, no visible interaction was detected in other combinations (Figure 9B)

Discussion

Dof proteins represent one of the plant-specific TF families and play a substantial regulatory role in the resistance mechanism against various abiotic stresses in plants (Noguero et al., 2013; Waqas et al., 2020). Sweetpotato is not only a crucial food crop, but also widely used in animal feed and industrial materials. In addition, sweetpotato has a variety of natural advantages that other common crops do not have, such as wide adaptability, high yield, and strong stress resistance (Liu, 2017). However, the specifi;c roles of most Dof genes remain unknown, and the Dof TFs in sweetpotato has not been comprehensively characterized so far. The development of bioinformatics and the completion of sweetpotato genome sequencing provide a strong basis for genome-wide identification of certain gene families (Yang et al., 2017b). In this work, the Dof TFs in sweetpotato were systematically investigated, and the identification and characteristic of stress-related *IbDofs* provide worthy information for the further functional exploration of *IbDof* genes in stress tolerance.

Diverse characterizations of IbDof TFs in sweetpotato

In this study, a total of 43 *IbDof* genes were identified in the sweetpotato genomes. These genes were unevenly distributed across the chromosomes and did not correspond to chromosome size, which may be caused by the unequal gene duplications of chromosomal segments. No *IbDof* gene was observed on Chr4, probably because of fragment loss or chromosome translocation during the evolution (Zhang et al., 2022). Comparable distributions were also observed in sorghum (Xiao et al., 2022), cabbage (Ma et al., 2015) and wheat (Fang et al., 2020), demonstrating that the number of Dof members was not closely related to the size of chromosomes. Although each IbDof TF has a conserved Dof domain, there are still considerable variances in the molecular features of IbDofs, which may be due to changes in their non-conserved regions. The significant variability indicates high complexity of sweetpotato IbDof TFs. Additionally, the number of 43 *IbDof* genes is more than that in *Arabidopsis* (36) (Lijavetzky et al., 2003), rice (30) (Khan et al., 2021), tomato (34) (Cai et al., 2013), pepper (33) (Wu et al., 2016), sorghum (30) (Xiao et al., 2022), and potato (35) (Venkatesh and Park, 2015), but much lower than that in *Saccharum officinarum* (119) (Gupta et al., 2014),

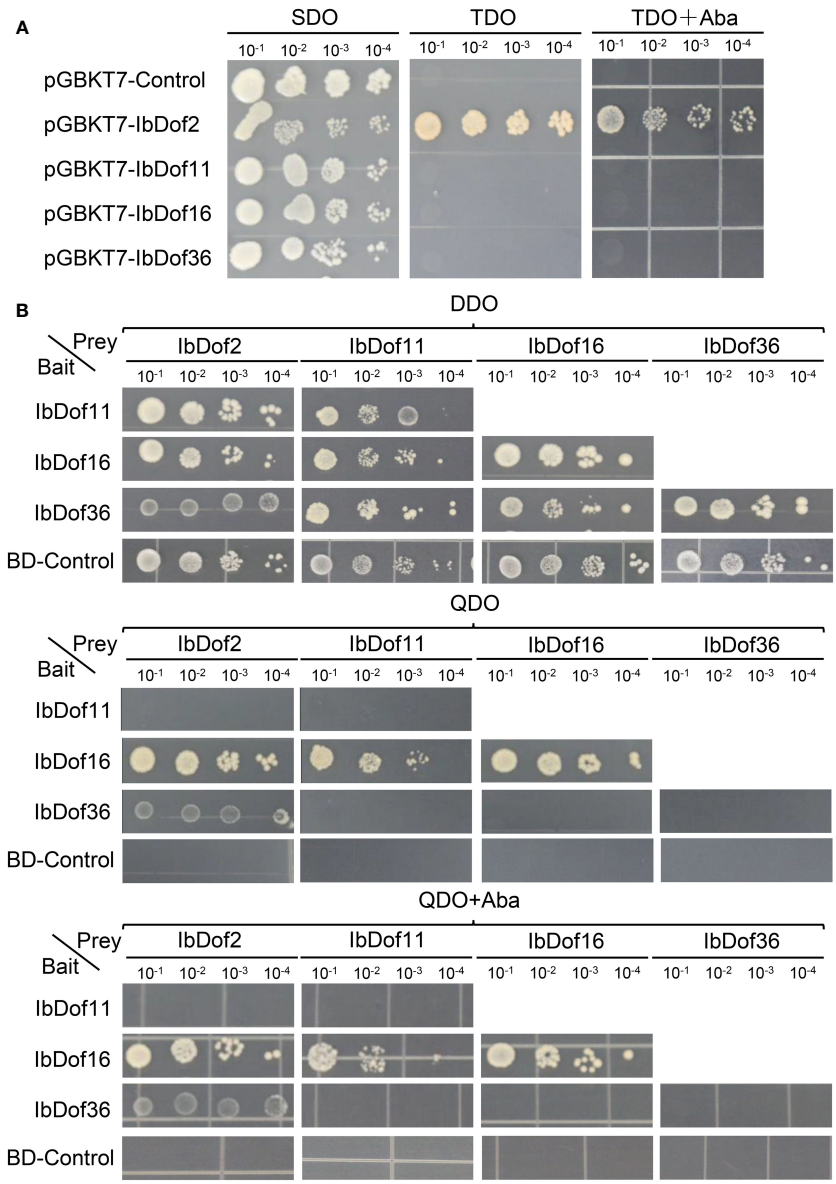


FIGURE 9
Analysis of transactivation activities and protein interactions of IbDof proteins. **(A)** Transactivation assays of IbDof-2/-11/-16/-36 in yeast strain Y2Hgold. Yeasts containing pGBKT7-IbDof-2/-11/-16/-36 or pGBKT7 empty vector (negative control) were streaked on the SDO (SD medium lacking Trp); TDO (SD medium lacking Trp, His, Ade) and TDO medium with 200 ng/mL AbA, respectively. **(B)** Yeasts containing both the indicated recombined pGBKT7 and pGADT7 plasmids were streaked on DDO (SD/-Trp-Leu) medium, QDO (SD/-Trp-Leu-His-Ade) medium with or without 200 ng/mL AbA. All the plates were recorded three days after 30° of incubation.

Gossypium hirsutum (114) (Li et al., 2018), *Brassica napus* (117) (Lohani et al., 2021), sugarcane (119 members) (Gupta et al., 2014), and wheat (108 members) (Fang et al., 2020), indicating that the amount of Dof TFs varies greatly in monocotyledons and dicotyledons.

Phylogenetic analysis showed that 43 IbDofs were classified into nine subgroups based on sequence homologies from known Dof family members from Arabidopsis and rice (Lijavetzky et al., 2003), consistent with the results of the Dof TFs in multiple plants such as rose (Nan et al., 2021), watermelon (Zhou et al., 2020), Tartary buckwheat (Li et al., 2022), *Gossypium hirsutum* (114) (Li et al., 2018), and *Brassica napus* (117) (Lohani et al., 2021). The gene structures can be applied as supporting evidence to confirm the evolutionary relationships among genes or organisms (Koralewski and Krutovsky, 2011). Gene structure analysis showed that there were significant differences amongst different subgroups, while similar structures were detected within the same subgroup, as was the case in motif analysis. For example, multiple subgroups contain their own unique motifs. The findings suggest the potential functional differentiation of IbDofs among different subgroups. And the intron numbers of *IbDof* genes were relatively small, most genes had only one intron or even no intron, which was similar to the Dof genes in wheat (Fang et al., 2020), rose (Nan et al., 2021), and Tartary buckwheat (Li et al., 2022). As reported in the

previous result, the intron-less genes may be associated with the quick stress response (Zhang et al., 2022).

Many plant gene families have evolved and expanded as a result of gene replication events, which may also encourage the formation of new functional genes and species that can better withstand harsh environments as plants evolve (Chen and Cao, 2015). Previous collinear studies on Dof gene family in Tartary buckwheat (Li et al., 2022), rose (Nan et al., 2021), and *Gossypium hirsutum* (Li et al., 2018) showed that segmental duplication events played a dominant role in Dof gene expansion. Similarly, no tandem repeat events were observed among sweetpotato *IbDof* genes, and segmental duplications were found to have a primary contribution to their expansions, indicating some *IbDof* genes may have originated from gene duplications in sweetpotato. However, studies also displayed that both tandem and segmental duplications exist simultaneously in the Dofs from *Brassica napus* (Lohani et al., 2021), wheat (Fang et al., 2020), and poplar (Wang et al., 2017).

Collinearity analysis can provide valuable insights into the evolutionary history of species (Wang et al., 2012). The synteny study examining the connections between *IbDof* genes and the equivalents from eight representative plant species were investigated. The biggest number of orthologous genes among them were found between sweetpotato and *Ipomoea triloba* and *Ipomoea trifida*, which was supported by their close evolutionary relationships, followed by tomato, pepper and Arabidopsis. These orthologous pairings may share common ancestors with the related plants and sweetpotato. In addition, a more intricate link between a single *Ipomoea triloba*/*Ipomoea trifida* gene and multiple sweetpotato *IbDof* genes was detected, suggesting that these orthologous genes may play an important role in the evolution of sweetpotato *IbDofs*. No orthologous gene pair was observed between sweetpotato and rice, probably due to numerous chromosomal rearrangements or fusions in the genomes (Paterson et al., 2012).

Expression profiling and functional prediction of *IbDofs* in sweetpotato

An increasing number of studies on the crucial roles of Dof TFs in regulating plant responses to multiple harmful environmental stimuli suggest that Dofs are excellent candidates for improving agricultural stress resistance through molecular breeding (Waqas et al., 2020). For instance, overexpression of tomato cycling DOF factor SlCDF1/SlCDF3 or grain amaranth AhDOF TF in Arabidopsis all showed enhanced drought and salt tolerance (Corrales et al., 2014; Massange-Sanchez et al., 2016). Besides, *Juglans regia* JrDof3 contributed to enhance the high temperature stress response of JrGRAS2, which could significantly regulate the expression of *HSPs* to improve high temperature stress tolerance (Yang et al., 2018). However, our understanding of how Dof TFs regulate stress responses in most plants, including sweetpotato, is still quite limited. The cis-elements present in the promoter regions play a key role in gene expression regulation (Meng et al., 2023). Numerous hormone- and stress-related elements in the promoters of *IbDof* genes were detected, such as the MBS, TC-rich repeats, P-box, and ABRE. Our transcriptome data and qRT-PCR results also revealed that the expression of multiple *IbDof* genes was clearly differed in response to different abiotic stress and hormone treatments, suggesting that sweetpotato *IbDof* genes may also have important and varied roles in response to environmental stresses. The data are supported by previous similar results, as many *TaDof* genes in wheat and *RchDof* genes in rose were also obviously upregulated by salt and drought treatments (Fang et al., 2020). And the transcription levels of Arabidopsis *CDF3* were significantly induced by drought, extreme temperature and ABA treatments, the transgenic assays showed that overexpression of *CDF3* could significantly improve the drought, cold and osmotic stress tolerance (Corrales et al., 2017; Nan et al., 2021). Therefore, sweetpotato *IbDof2*, *IbDof16*, and *IbDof36* genes showed obviously increased expression after stress treatments, indicating the accumulations of these genes may lessen the harm caused by unfavorable stresses in our analysis, but additional experimental confirmation is needed.

The dimerization of Dof TFs with other proteins depends on the conserved Dof domain (Gupta et al., 2015). The first protein-protein interaction of Dof TFs was found in Arabidopsis in which AtOBP1 interacted with the bZIP protein to regulate stress response. Similarly, many Dof proteins have also been found to exert their regulatory functions by forming dimers such as OsDof3 by interacting with GAMYB in rice, Dof1 by interacting with Dof2 and HMG1 in maize, and Dof 3.2 by interacting with TCP14 in Arabidopsis (Yanagisawa, 1997; Waqas et al., 2020). In the present study, the STRING database predicted that sweetpotato IbDof TFs may participate in the stress tolerance through a complicated protein interaction network based on the homologous proteins of Arabidopsis. For example, TMO6 (IbDof-3/-6/-9/-24) might interact with numerous Dof proteins including OBP2 (IbDof24), OBP3 (IbDof16), OBP4 (IbDof35 and IbDof39), and HCA2 (IbDof23). And IbDof16 could interact with IbDof2, IbDof11 and itself, and IbDof2 could also interact with IbDof36 according to the subsequent Y2H studies, indicating a complicated interaction connection amongst sweetpotato IbDof proteins. Collectively, our findings imply that multiple stress-responsive IbDof TFs can form intricate complexes with Dof and other types of proteins through direct protein-protein interactions, which may exert crucial roles in abiotic stress signaling cascades.

Collectively, in this study, we found that 43 *IbDof* genes were unevenly distributed on 14 of the 15 chromosomes of cultivated sweetpotato. The majority of *IbDof* genes lacked introns, and these IbDofs could be divided into nine subgroups according to the phylogenetic analysis. The *IbDof* genes within the same subgroup generally shared similar gene structures and motif compositions, while they were distinguishable among different subgroups. Segmental duplication events were shown to be the prominent driving force for the expansion of sweetpotato *IbDof* genes, and a collinearity analysis of orthologous genes from eight typical plants gave important hints about the evolutionary traits of the *IbDof* genes. RNA-seq data and qRT-PCR detection showed that multiple *IbDof* genes, particularly the *IbDof2*, *IbDof16* and *IbDof36*, were significantly upregulated in response to abiotic stressors and hormones, indicating that they might play a pivotal role in stress resistance. Additionally, IbDof2 protein has obvious transactivation activity, and a complicated interaction relationship between IbDof TFs was found, suggesting the complex connection and regulatory mechanism for IbDof TFs in regulating the response of sweetpotato

to environmental stresses. These findings provide valuable information for further comprehending the intricacy and importance of the Dof gene family, and multiple IbDof members with promising prospects in regulating sweetpotato response to environmental stresses are expected.

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

XM and MZ designed the experiments, CZ, MZ, and XM analyzed the data and wrote the manuscript. CZ, TD, JY, HH, SL, FG, and HM performed the experiments and analyzed the data, TD and JZ improved the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by National Natural Science Foundation of China (32201703, 32171936), the earmarked fund for CARS-10-Sweetpotato, Natural science fund for colleges and universities in Jiangsu Province (21KJB180014), and Xuzhou Science and Technology Planning Project (KC21117).

Acknowledgments

We appreciate the reviewers for their valuable comments on 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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1140727/full#supplementary-material

References

Cai, X., Zhang, Y., Zhang, C., Zhang, T., Hu, T., Ye, J., et al. (2013). Genome-wide analysis of plant-specific dof transcription factor family in tomato. *J. Integr. Plant Biol.* 55, 552–566. doi: 10.1111/jipb.12043

Chen, Y., and Cao, J. (2015). Comparative analysis of dof transcription factor family in maize. *Plant Mol. Biol. Rep.* 33, 1245–1258. doi: 10.1007/s11105-014-0835-9

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

Corrales, A. R., Carrillo, L., Lasierra, P., Nebauer, S. G., Dominguez-Figueroa, J., Renau-Morata, B., et al. (2017). Multifaceted role of cycling DOF factor 3 (CDF3) in the regulation of flowering time and abiotic stress responses in arabidopsis. *Plant Cell Environ.* 40, 748–764. doi: 10.1111/pce.12894

Corrales, A.-R., Nebauer, S. G., Carrillo, L., Fernández-Nohales, P., Marqués, J., Renau-Morata, B., et al. (2014). Characterization of tomato cycling dof factors reveals conserved and new functions in the control of flowering time and abiotic stress responses. *J. Exp. Bot.* 65, 995–1012. doi: 10.1093/jxb/ert451

Dong, C., Hu, H., and Xie, J. (2016). Genome-wide analysis of the DNA-binding with one zinc finger (Dof) transcription factor family in bananas. *Genome* 59, 1085–1100. doi: 10.1139/gen-2016-0081

Erpen, L., Devi, H. S., Grosser, J. W., and Dutt, M. (2018). Potential use of the DREB/ERF, MYB, NAC and WRKY transcription factors to improve abiotic and biotic stress in transgenic plants. *Plant Cell Tissue Organ Cult.* 132, 1–25. doi: 10.1007/s11240-017-1320-6

Fang, Z., Jiang, W., He, Y., Ma, D., Liu, Y., Wang, S., et al. (2020). Genome-wide identification, structure characterization, and expression profiling of dof transcription factor gene family in wheat (*Triticum aestivum* l.). *Agronomy-Basel* 10, 294. doi: 10.3390/agronomy10020294

Guo, F., Liu, S., Zhang, C., Dong, T., Meng, X., and Zhu, M. (2022). Genome-wide systematic survey and analysis of NAC transcription factor family and their response to abiotic stress in sweetpotato. *Scientia Hortic.* 299, 111048. doi: 10.1016/j.scienta.2022.111048

Gupta, S., Kushwaha, H., Singh, V. K., Bisht, N. C., Sarangi, B. K., and Yadav, D. (2014). Genome wide in silico characterization of dof transcription factor gene family of sugarcane and its comparative phylogenetic analysis with arabidopsis, rice and sorghum. *Sugar Tech.* 16, 372–384. doi: 10.1007/s12355-013-0288-8

Gupta, S., Malviya, N., Kushwaha, H., Nasim, J., Bisht, N. C., Singh, V., et al. (2015). Insights into structural and functional diversity of dof (DNA binding with one finger) transcription factor. *Planta* 241, 549–562. doi: 10.1007/s00425-014-2239-3

Ichimaru, K., Yamaguchi, K., Harada, K., Nishio, Y., Hori, M., Ishikawa, K., et al. (2022). Cooperative regulation of PBI1 and MAPKs controls WRKY45 transcription factor in rice immunity. *Nat. Commun.* 13, 2397. doi: 10.1038/s41467-022-30131-y

Jin, Y., Ding, F., Wang, J., Yi, Z., Gao, Y., Yang, L., et al. (2022). One-step conversion of sweet potato waste to butanol *via* fermentation by clostridium acetobutylicum. *Biomass Convers. Biorefin.* 1-12. doi: 10.1007/s13399-022-03314-2

Kang, W.-H., Kim, S., Lee, H.-A., Choi, D., and Yeom, S.-I. (2016). Genome-wide analysis of dof transcription factors reveals functional characteristics during development and response to biotic stresses in pepper. *Sci. Rep.* 6, 33332. doi: 10.1038/srep33332

Khan, I., Khan, S., Zhang, Y., and Zhou, J. (2021). Genome-wide analysis and functional characterization of the dof transcription factor family in rice (*Oryza sativa* l.). *Planta* 253, 101. doi: 10.1007/s00425-021-03627-y

Koralewski, T. E., and Krutovsky, K. V. (2011). Evolution of exon-intron structure and alternative splicing. *PloS One* 6, e18055. doi: 10.1371/journal.pone.0018055

Krohn, N. M., Yanagisawa, S., and Grasser, K. D. (2002). Specificity of the stimulatory interaction between chromosomal HMGB proteins and the transcription factor Dof2 and its negative regulation by protein kinase CK2-mediated phosphorylation. *J. Biol. Chem.* 277, 32438–32444. doi: 10.1074/jbc.M203814200

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., et al. (2009). Circos: An information aesthetic for comparative genomics. *Genome Res.* 19, 1639–1645. doi: 10.1101/gr.092759.109

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. And Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096

Kushwaha, H., Gupta, S., Singh, V. K., Rastogi, S., and Yadav, D. (2011). Genome wide identification of dof transcription factor gene family in sorghum and its comparative phylogenetic analysis with rice and arabidopsis. *Mol. Biol. Rep.* 38, 5037–5053. doi: 10.1007/s11033-010-0650-9

Li, H., Dou, L., Li, W., Wang, P., Zhao, Q., Xi, R., et al. (2018). Genome-wide identification and expression analysis of the dof transcription factor gene family in *Gossypium hirsutum* l. *Agronomy* 8, 186. doi: 10.3390/agronomy8090186

Li, J., Zhang, Y., Xu, L., Wang, C., Luo, Y., Feng, S., et al. (2022). Genome-wide identification of DNA binding with one finger (Dof) gene family in tartary buckwheat (*Fagopyrum tataricum*) and analysis of its expression pattern after exogenous hormone stimulation. *Biology* 11, 173. doi: 10.3390/biology11020173

Lijavetzky, D., Carbonero, P., and Vicente-Carbajosa, J. (2003). Genome-wide comparative phylogenetic analysis of the rice and arabidopsis dof gene families. *BMC evolutionary Biol.* 3, 17. doi: 10.1186/1471-2148-3-17

Liu, Q. C. (2017). Improvement for agronomically important traits by gene engineering in sweetpotato. *Breed. Sci.* 67, 15–26. doi: 10.1270/jsbbs.16126

Liu, S., Zhang, C., Zhu, Q., Guo, F., Chai, R., Wang, M., et al. (2022). Genome-and transcriptome-wide systematic characterization of bZIP transcription factor family identifies promising members involved in abiotic stress response in sweetpotato. *Scientia Hortic.* 303, 111185. doi: 10.1016/j.scienta.2022.111185

Lohani, N., Babaei, S., Singh, M. B., and Bhalla, P. L. (2021). Genome-wide in silico identification and comparative analysis of dof gene family in *Brassica napus*. *Plants* 10, 709. doi: 10.3390/plants10040709

Ma, J., Li, M.-Y., Wang, F., Tang, J., and Xiong, A.-S. (2015). Genome-wide analysis of dof family transcription factors and their responses to abiotic stresses in Chinese cabbage. *BMC Genomics* 16, 33. doi: 10.1186/s12864-015-1242-9

Massange-Sanchez, J. A., Palmeros-Suarez, P. A., Espitia-Rangel, E., Rodriguez-Arevalo, I., Sanchez-Segura, L., Martinez-Gallardo, N. A., et al. (2016). Overexpression of grain amaranth (*Amaranthus hypochondriacus*) AhERF or AhDOF transcription factors in *Arabidopsis thaliana* increases water deficit-and salt-stress tolerance, respectively, *via* contrasting stress-amelioration mechanisms. *PloS One* 11, e0164280. doi: 10.1371/journal.pone.0164280

Meng, X., Cai, J., Deng, L., Li, G., Sun, J., Han, Y., et al. (2020a). *SlSTE1* promotes abscisic acid-dependent salt stress-responsive pathways *via* improving ion homeostasis and reactive oxygen species scavenging in tomato *J. Integr. Plant Biol.* 62, 1942–1966. doi: 10.1111/jipb.12987

Meng, X., Li, G., Yu, J., Cai, J., Dong, T., Sun, J., et al. (2018). Isolation, expression analysis, and function evaluation of 12 novel stress-responsive genes of NAC transcription factors in sweetpotato. *Crop Sci.* 58, 1328–1341. doi: 10.2135/cropsci2017.12.0738

Meng, X., Liu, S., Dong, T., Xu, T., Ma, D., Pan, S., et al. (2020b). Comparative transcriptome and proteome analysis of salt-tolerant and salt-sensitive sweet potato and overexpression of *IbNAC7* confers salt tolerance in arabidopsis. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.572540

Meng, X., Liu, S., Zhang, C., He, J., Ma, D., Wang, X., et al. (2023). The unique sweet potato NAC transcription factor IbNAC3 modulates combined salt and drought stresses. *Plant Physiol.* 191, 747–771. doi: 10.1093/plphys/kiac508

Nan, H., Ludlow, R. A., Lu, M., and An, H. (2021). Genome-wide analysis of dof genes and their response to abiotic stress in rose (*Rosa chinensis*). *Front. Genet.* 12. doi: 10.3389/fgene.2021.538733

Noguero, M., Atif, R. M., Ochatt, S., and Thompson, R. D. (2013). The role of the DNA-binding one zinc finger (DOF) transcription factor family in plants. *Plant Sci.* 209, 32–45. doi: 10.1016/j.plantsci.2013.03.016

Park, S. C., Kim, Y. H., Chang, Y. J., Park, S., Jeong, J. C., Lee, H. S., et al. (2012). Stable internal reference genes for the normalization of real-time PCR in different sweetpotato cultivars subjected to abiotic stress conditions. *PloS One* 7, e51502. doi: 10.1371/journal.pone.0051502

Paterson, A., Wang, X., Tang, H., and Lee, T. (2012). doi: 10.1007/978-3-7091-1130-7_13

Qin, H., Wang, J., Chen, X., Wang, F., Peng, P., Zhou, Y., et al. (2019). Rice OsDOF15 contributes to ethylene-inhibited primary root elongation under salt stress. *New Phytol.* 223, 798–813. doi: 10.1111/nph.15824

Ramachandran, V., Tobimatsu, Y., Masaomi, Y., Sano, R., Umezawa, T., Demura, T., et al. (2020). Plant-specific dof transcription factors VASCULAR-RELATED DOF1 and VASCULAR-RELATED DOF2 regulate vascular cell differentiation and lignin biosynthesis in arabidopsis. *Plant Mol. Biol.* 104, 263–281. doi: 10.1007/s11103-020-01040-9

Ruta, V., Longo, C., Lepri, A., De Angelis, V., Occhigrossi, S., Costantino, P., et al. (2020). The DOF transcription factors in seed and seedling development. *Plants* 9, 218. doi: 10.3390/plants9020218

Ssamula, A., Okiror, A., Avrahami-Moyal, L., Tam, Y., Gaba, V., Gibson, R. W., et al. (2020). Factors influencing reversion from virus infection in sweetpotato. *Ann. Appl. Biol.* 176, 109–121. doi: 10.1111/aab.12551

Venkatesh, J., and Park, S. W. (2015). Genome-wide analysis and expression profiling of DNA-binding with one zinc finger (Dof) transcription factor family in potato. *Plant Physiol. Biochem.* 94, 73–85. doi: 10.1016/j.plaphy.2015.05.010

Wang, Y., Tang, H., Debarry, J. D., Tan, X., Li, J., Wang, X., et al. (2012). *MCScanX*: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 40, e49–e49. doi: 10.1093/nar/gkr1293

Wang, Z., Wang, Y., Tong, Q., Xu, G., Xu, M., Li, H., et al. (2021). Transcriptomic analysis of grapevine dof transcription factor gene family in response to cold stress and functional analyses of the *VaDof17d* gene. *Planta* 253, 55. doi: 10.1007/s00425-021-03574-8

Wang, H., Zhao, S., Gao, Y., and Yang, J. (2017). Characterization of dof transcription factors and their responses to osmotic stress in poplar (*Populus trichocarpa*). *PloS One* 12, e0170210. doi: 10.1371/journal.pone.0170210

Waqas, M., Shahid, L., Shoukat, K., Aslam, U., Azeem, F., and Atif, R. M. (2020). "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*. *Elsevier*, 1–14. doi: 10.1016/B978-0-12-819334-1.00001-0

Wei, Q., Wang, W., Hu, T., Hu, H., Mao, W., Zhu, Q., et al. (2018). Genome-wide identification and characterization of dof transcription factors in eggplant (*Solanum melongena* l.). *Peerj* 6, e4481. doi: 10.7717/peerj.4481

Wu, Z., Cheng, J., Cui, J., Xu, X., Liang, G., Luo, X., et al. (2016). Genome-wide identification and expression profile of dof transcription factor gene family in pepper (*Capsicum annuum* l.). *Front. Plant Sci.* 7. doi: 10.3389/fpls.2016.00574

Xiao, Q., Liu, T., Ling, M., Ma, Q., Cao, W., Xing, F., et al. (2022). Genome-wide identification of DOF gene family and the mechanism dissection of *SbDof21* regulating starch biosynthesis in sorghum. *Int. J. Mol. Sci.* 23, 12152. doi: 10.3390/ijms232012152

Yanagisawa, S. (1997). Dof DNA-binding domains of plant transcription factors contribute to multiple protein-protein interactions. *Eur. J. Biochem.* 250, 403–410. doi: 10.1111/j.1432-1033.1997.0403a.x

Yanagisawa, S., and Izui, K. (1993). Molecular cloning of two DNA-binding proteins of maize that are structurally different but interact with the same sequence motif. *J. Biol. Chem.* 268, 16028–16036. doi: 10.1016/S0021-9258(18)82353-5

Yang, G., Gao, X., Ma, K., Li, D., Jia, C., Zhai, M., et al. (2018). The walnut transcription factor JrGRAS2 contributes to high temperature stress tolerance involving in dof transcriptional regulation and HSP protein expression. *BMC Plant Biol.* 18, 367. doi: 10.1186/s12870-018-1568-y

Yang, J., Moeinzadeh, M. H., Kuhl, H., Helmuth, J., Xiao, P., Haas, S., et al. (2017b). Haplotype-resolved sweet potato genome traces back its hexaploidization history. *Nat. Plants* 3, 696–703. doi: 10.1038/s41477-017-0002-z

Yang, G., Yu, L., Wang, Y., Wang, C., and Gao, C. (2017a). The translation initiation factor 1A (TheIF1A) from tamarix hispida is regulated by a dof transcription factor and increased abiotic stress tolerance. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.00513

Zhang, C., Liu, S., Liu, D., Guo, F., Yang, Y., Dong, T., et al. (2022). Genome-wide survey and expression analysis of GRAS transcription factor family in sweetpotato provides insights into their potential roles in stress response. *BMC Plant Biol.* 22, 232. doi: 10.1186/s12870-022-03618-5

Zhou, Y., Cheng, Y., Wan, C., Li, J., Yang, Y., and Chen, J. (2020). Genome-wide characterization and expression analysis of the dof gene family related to abiotic stress in watermelon. *PeerJ* 8, e8358. doi: 10.7717/peerj.8358

Zhu, M., Chen, G., Dong, T., Wang, L., Zhang, J., Zhao, Z., et al. (2015). *SlDEAD31*, a putative DEAD-box RNA helicase gene, regulates salt and drought tolerance and stress-related genes in tomato. *PloS One* 10, e0133849. doi: 10.1371/journal.pone.0133849

Ziska, L. H., Runion, G. B., Tomecek, M., Prior, S. A., Torbet, H. A., and Sicher, R. (2009). An evaluation of cassava, sweet potato and field corn as potential carbohydrate sources for bioethanl production in Alabama and Maryland. *Biomass Bioenergy* 33, 1503–1508. doi: 10.1016/j.biombioe.2009.07.014+

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 06 March 2023
DOI 10.3389/fpls.2023.1147711

OPEN ACCESS

EDITED BY
Kaixuan Zhang,
Institute of Crop Sciences (CAAS), China

REVIEWED BY
Baoyuan Zhou,
Institute of Crop Sciences (CAAS), China
Jun Ma,
China Agricultural University, China

*CORRESPONDENCE
Mingming Yang
myang@nwsuaf.edu.cn
Wenchao Zhen
wenchao@hebau.edu.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Crop and Product Physiology, a section of the journal Frontiers in Plant Science

RECEIVED 19 January 2023
ACCEPTED 10 February 2023
PUBLISHED 06 March 2023

CITATION
Liu P, Yin B, Liu X, Gu L, Guo J, Yang M and Zhen W (2023) Optimizing plant spatial competition can change phytohormone content and promote tillering, thereby improving wheat yield.
Front. Plant Sci. 14:1147711.
doi: 10.3389/fpls.2023.1147711

Optimizing plant spatial competition can change phytohormone content and promote tillering, thereby improving wheat yield

Pan Liu[1,2,3,4†], Baozhong Yin[5†], Xuejing Liu[1,2,3,4,6†], Limin Gu[1,2,3,4], Jinkao Guo[1,2,3,4,7], Mingming Yang[8*] and Wenchao Zhen[1,2,3,4*]

[1]College of Agronomy, Hebei Agricultural University, Baoding, China, [2]State Key Laboratory of North China Crop Improvement and Regulation, Baoding, China, [3]Key Laboratory of North China Water-saving Agriculture, Ministry of Agriculture and Rural Affairs, Baoding, China, [4]Key Laboratory of Crop Growth Regulation of Hebei Province, Baoding, China, [5]College of Plant Protection, Hebei Agricultural University, Baoding, China, [6]College of Clinical Medicine, North China University of Technology, Tangshan, China, [7]Wheat Research Center, Shijiazhuang Academy of Agriculture and Forestry Sciences, Shijiazhuan, China, [8]College of Agronomy, Northwest A&F University, Xianyang, China

As an important type of interplant competition, line-spacing shrinkage and row-spacing expansion (LSRE) can increase the number of tillers and improve resource utilization efficiency in wheat. Wheat tillering is closely related to various phytohormones. However, it is unclear whether LSRE regulates phytohormones and their relationship to tillering and wheat yield. This study evaluated tillering characteristics, phytohormone content in tiller nodes at the pre-winter stage, and grain yield factors for the winter wheat variety Malan1. We used a two-factor randomized block trial design with two sowing spacings of 15 cm (15RS, conventional treatment) and 7.5 cm (7.5RS, LSRE treatment) at the same density and three sowing-date groups (SD1, SD2, and SD3). LSRE significantly promoted wheat tillering and biomass at the pre-winter stage (average increases of 14.5% and 20.9% in the three sowing-date groups, respectively) and shortened the accumulated temperature required for a single tiller. Changes in the levels of phytohormones, including decreased gibberellin and indole acetic acid and increased zeatin riboside and strigolactones, were determined by high-performance liquid chromatography and were shown to be responsible for the tillering process under LSRE treatment in winter wheat. LSRE treatment can improve crop yield by increasing the number of spikes per unit area and grain weight. Our results clarified the changes in tillering and phytohormones content of winter wheat under LSRE treatment and their correlation with grain yield. This study also provides insights into the physiological mechanisms of alleviating inter-plant competition to improve crop yield.

KEYWORDS

winter wheat, line-spacing shrinkage and row-spacing expansion, phytohormone, tiller, grain yield

1 Introduction

Wheat (*Triticum aestivum* L.) is one of the three major food crops worldwide (Yang et al., 2018) and the primary food source for more than one-third of the world's population. In particular, as the global population increases (expected to reach 10 billion by 2050), wheat demand will increase further (Cai et al., 2018). The North China Plain (NCP) is the largest wheat-producing area in China, accounting for approximately 56% of the total output, and plays a critical role in ensuring China's food security (Feng et al., 2023). The main planting system adopted by the NCP is double cropping of winter wheat and summer maize each year. This can improve the land multiple cropping index and effectively increase grain yield. However, this method has also postponed the winter wheat sowing date in this area, which is affected by the previous maize harvest date, weather, soil moisture, and other factors, resulting in an insufficient crop population and reduced resistance during the pre-winter stage (Xia et al., 2023). This situation is even more prominent in the heat resource shortage area north of the NCP. The ability to use limited heat resources to generate sufficient populations and healthy individuals is important for understanding the theory and technology of high-yield wheat cultivation in this region.

Tillering is a crucial agronomic characteristic for controlling the population as well as the individual structure and yield of wheat and other cereal crops (Fioreze and Rodrigues, 2014; Lecarpentier et al., 2019). Tillers have essential functions in organic matter storage and should be present in an appropriate quantity and quality to guarantee the number of effective ears in wheat fields. Tillers also promote the development of secondary roots, improve the utilization efficiency of water and nutrients on farmland, and ultimately affect yield (Yu et al., 2020). Therefore, regulating the growth and development of tillers, particularly the quantity and quality of tillers in the pre-winter stage, is essential to ensure high-yield and high-efficiency wheat cultivation. Wheat tillers develop from axillary meristems, and genetic factors mainly control their establishment and axillary bud formation. Axillary meristem growth is a typical quantitative genetic trait regulated by a complex regulatory network involving genetics, hormones, and the environment (Wang et al., 2016; Barbier et al., 2019).

Tillering of gramineous crops is initiated from the base of the plant, and the growth and development of tillers are significantly affected by the distance and density of adjacent plants (Lecarpentier et al., 2019). Therefore, spatial competition resulting from the individual distribution of plants is a critical environmental condition affecting the growth of wheat tillers (Yu et al., 2020). High-density planting inhibits the formation of wheat tiller buds and reduces the number of tillers per plant (Xu et al., 2016; Barbier et al., 2019). In addition, even at the same planting density and row spacing, wheat tillers exhibit differences. For example, at an equal planting density, moderately reducing row spacing will promote tillering, although row spacing that is too narrow will prevent tillering (McSteen and Leyser, 2005; Ozturk et al., 2006). Because wheat plant spacing is much smaller than row spacing, competition between plants in the same row is more significant than that between rows for wheat tillering. Therefore, wheat tillering can be promoted appropriately by reducing row spacing, expanding plant spacing, and alleviating plant competition. These control measures have become more practical in fields with fewer tillers because of late sowing dates and low accumulated temperatures.

Various phytohormones participate in the regulation of branch and tiller development (Müller and Leyser, 2011; Wang et al., 2018). For example, indole-3-acetic acid (IAA) and cytokinin (CTK) respectively promote and inhibit rice lateral bud formation. Other studies have also shown that the growth of wheat tiller buds is regulated by the concentrations of IAA, ABA, and zeatin (ZT) in tiller nodes, and the ratio of the contents of these phytohormones also plays a significant role (Shang et al., 2021). Strigolactones (SLs) are carotenoid-derived phytohormones that were first demonstrated in 2008 to inhibit the outgrowth of axillary buds (Liu et al., 2011; Liu et al., 2021b). Although discovered later, the effects of SL hormones play a crucial role in regulating the tillering of gramineous crops. Thus, their functions in different varieties of wheat and whether they have antagonistic or promotional relationships with other phytohormones should be examined.

The degree of competition between phytohormones and individual plant spaces significantly affects the tillering of gramineous crops and ultimately affects crop yield (Lafond, 1994; McSteen and Leyser, 2005; Evers et al., 2006; Hussain et al., 2012). However, there is little research on how phytohormones respond to plant competition at different tillering stages and how they regulate tillering and yield formation (Müller and Leyser, 2011). To understand these pivotal factors, we conducted a two-year field experiment in the northern part of the NCP, where heat resources are relatively scarce. We aimed to systematically examine the effects of line-spacing shrinkage and row-spacing expansion (LSRE) treatment on tillers and phytohormones of winter wheat at different sowing dates under the same sowing density as the different tillering effects on grain yield factors. This study provides some new ideas for expanding tillering regulation and efficiently utilizing light and heat resources, particularly for winter wheat production in areas with limited heat resources.

2 Materials and methods

2.1 Overview of test site

A two-year field experiment was performed during the winter wheat growing season (from early October to mid-June of the following year from 2019 to 2021) at the Malan R&D base of the National Grain High Yield Science and Technology Project in the NCP (Xinji City, Hebei Province, China, 115.22°E, 37.92°N, 43 m above sea level). The location and foundation soil fertility of the test station are shown in Figures 1A, B. From 1991 to 2018 at the test station, the average annual precipitation and temperature were 458.9 mm and 12.9°C, respectively; these values during the winter wheat season were 122.0 mm and 12.5°C, respectively. The

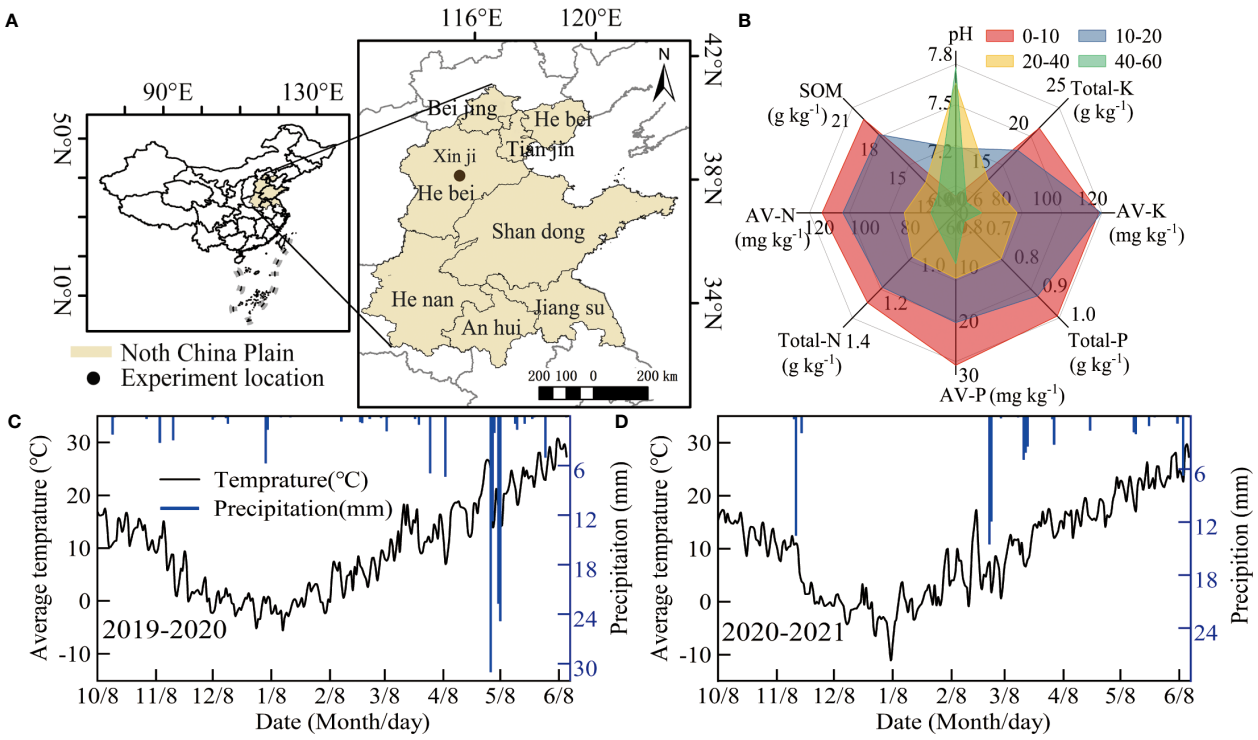


FIGURE 1
Basic information of the test site: **(A)** distribution of test sites and **(B)** foundation soil capacity of the test field. **(C, D)** are the daily precipitation and average temperature during the test.

precipitation and daily average temperature from 2019 to 2021 are shown in Figures 1C, D. The cumulative precipitation in the two winter wheat seasons was 139.8 mm and 72.4 mm, respectively.

2.2 Experimental design and treatments

A two-factor randomized block design was adopted: factor I involved a row spacing of 15 cm (15RS, conventional row spacing for wheat seeding in the NCP) and 7.5 cm row spacing (7.5RS, LSRE treatment). The field plant distributions for the two types of row spacing are shown in Figure 2. Factor II included a suitable sowing date (8 October, SD1), delayed sowing date 1 (13 October, SD2), and delayed sowing date 2 (18 October, SD3). The area of each plot was 65 m^2 (10 m × 6.5 m). The sowing densities of the two-row spacing treatments were the same on each sowing date. We selected Malan 1 (Variety approval No.: Jishenmai 20218011), a semi-compact winter wheat variety that is widely planted in the northern part of the NCP, as the test variety. The basic seedling design of three sowing-date groups (SD1, SD2, and SD3) was 3.3 × 10^6 kg hm^{-2}, 3.75 × 10^6 hm^{-2}, and 4.05 × 10^6 hm^{-2}. Their corresponding seeding amounts were 180.0 kg hm^{-2}, 210.0 kg hm^{-2}, and 225.0 kg hm^{-2}, respectively. After harvesting the previous summer's maize, the straw was crushed and returned to the field, and a 1GQN-230B rotary cultivator was used twice at a depth of 15 cm. Before sowing, 120 kg hm^{-2} of pure nitrogen, 112.5 kg hm^{-2} of P_2O_5, and 112.5 kg hm^{-2} of K_2O were added. Pure nitrogen was applied per hectare during the first irrigation in the spring. A Trimer Pico 64 portable soil moisture meter (TDR, IMIKO, Bochum, Germany) was used to measure soil moisture content (V/V). The target water content of irrigation at different stages of the wheat field was 80% of the field capacity (V/V); the calculated soil depth of irrigation before sowing was 0–40 cm, and the 0–60 cm soil layer for the jointing and anthesis stages. The irrigation amounts at different growth stages are shown in Table 1.

2.3 Number of basic seedlings and tillers

For each repeated plot, three rows of wheat exhibiting uniform growth were randomly selected as the survey sample section at the trilobal stage (BBCH 13, Cornelius et al., 2011) for each treatment and the basic number of wheat seedlings per unit area (BSN). We investigated the tiller number for every 60–80°C d increase in the active accumulated temperature (daily average temperature ≥3.0°C, Equation 1) from the trilobal stage to the overwintering stage (WS), and the survey data were recorded as D1, D2, D3,… Dn, WS, respectively (the specific measurement date is shown in Figure 3). The method used to calculate the active accumulated temperature (AT) in the prewinter stage is shown in Equation 2. From the standing stage of the wheat, the investigation was conducted once every 2–3 days until the number of tillers reached its maximum

10.3389/fpls.2023.1147711

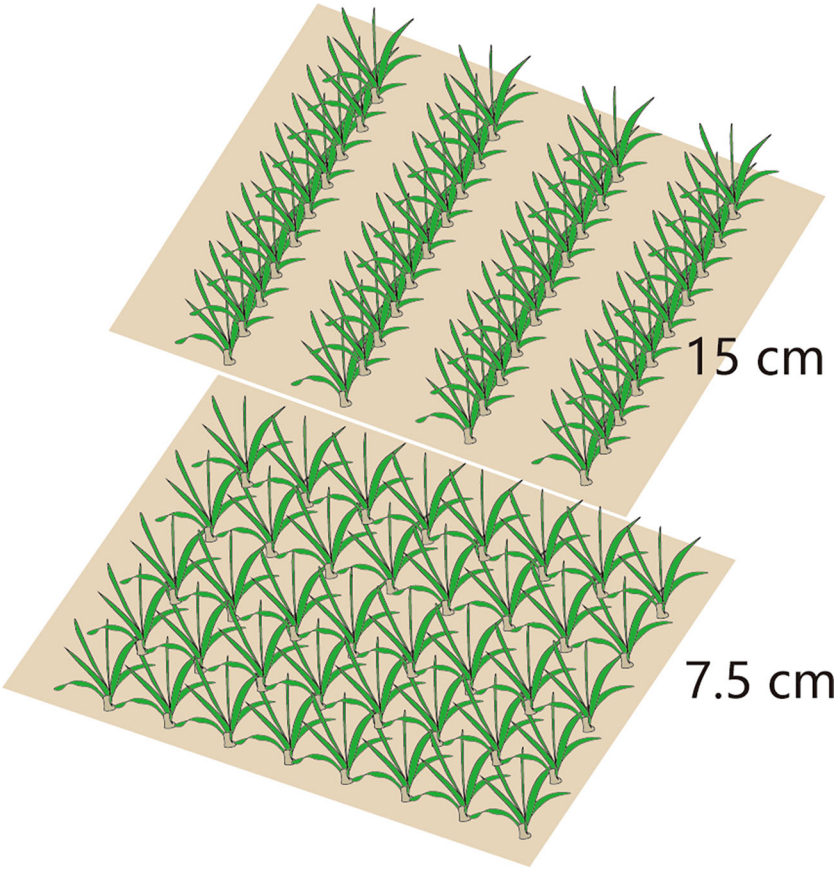


FIGURE 2
Spatial distribution of plants in different row spacing.

(BBCH 29) in spring. To avoid the impact of the survey on wheat growth, we only investigated one sample section of one repeated plot each time and alternated the measurement of different sections: tillers formed in winter (TP) and spring (TS) at the survey sample points were marked with red and yellow nylon threads (Tilley et al., 2019), respectively. When the tiller number reached its maximum, the sum was determined as the total tillers (PT) for the entire growth period.

A tiller was considered to have formed when the tiller bud protruded 2 cm from the base leaf axil (Figure 4A), and the tillering date of the entire field was considered when 50% of the plants reached this standard. The AT required for tillering formation is given by Equation 2. According to the sequence of primary tillering of the main stem, all tillers were named T_I, $T_{II,}$…T_n and the secondary tillers were named $T_{I\text{-}1}$, $T_{I\text{-}2,}$ and so on (Figure 4B). The green stalk number per unit area (N_{GS}) is the sum of the main

TABLE 1 Irrigation amount of wheat fields under different treatments (m^3 hm^{-2}).

Row space (cm)	Sowing date	2019–2020		2020–2021		
		Before sowing	Jointing stage	Before sowing	Jointing stage	Anthesis stage
15	SD1	538.5	873.0	384.0	631.5	826.5
	SD2	552.0	852.0	412.5	606.0	831.0
	SD3	580.5	798.0	423.0	568.5	790.5
7.5	SD1	568.5	822.0	336.0	591.0	756.0
	SD2	594.0	766.5	388.5	564.0	733.5
	SD3	550.5	723.0	393.0	525.0	678.0

No irrigation in the anthesis stage of wheat in 2020.

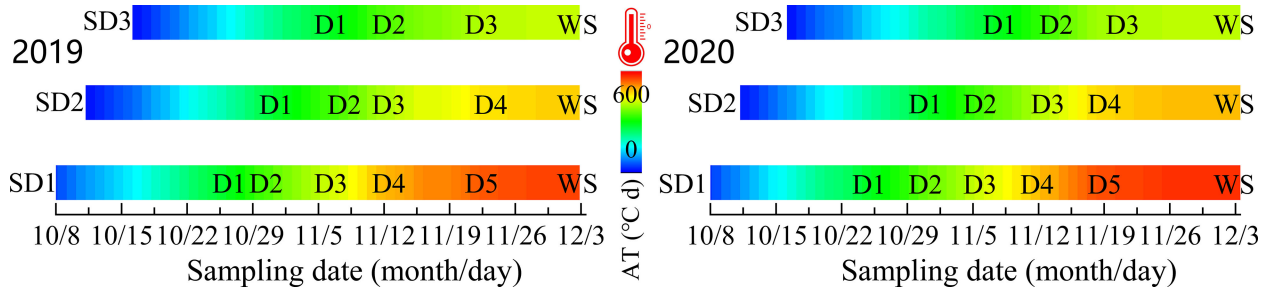


FIGURE 3
Sampling and determination date at pre-winter stage under different sowing dates.

stems and tillers (Equation 3).

$$\mathbf{AT}(°\mathrm{C}\ \mathbf{d}) = \sum_{i=1}^{n} t_i (\text{when } t_i < 3,\ t_i = 0) \qquad \text{Equation 1}$$

$$\mathbf{AT}_n(°\mathrm{C}\ \mathbf{d}) = \mathbf{AT}_{s \to n} - \mathbf{AT}_{s \to (n-1)} (\mathrm{n} \geq 1) \qquad \text{Equation 2}$$

$$\mathbf{N_{GS}}(\mathbf{stalks\ m^{-2}}) = \mathbf{N_{MS}}(\mathbf{main\ stems\ m^{-2}}) + \mathbf{N_T}(\mathbf{tillers\ m^{-2}}) \qquad \text{Equation 3}$$

where AT represents the active accumulated temperature from sowing to the *i*th day, n is the number of days after sowing (d), and t_i is the daily average temperature (Â °C) on the *i*th day after sowing.

A_n represents the AT required to form the n_{th} main stem tiller, $AT_{s \to n}$ and $AT_{s \to (n-1)}$ represent the AT required from sowing to the *n*th and $(n-1)$th main stem tiller formation, respectively. When $n = 1$, A_1 is the AT required from sowing to the trilobal period. N_{GS} is the number of green stalks per m^{-2}, N_{MS} is the number of green main stems per m^{-2}, and N_T is the number of green tillers per m^{-2}.

2.4 Biomass and phytohormone contents in tiller nodes at pre-winter stage

To determine the biomass according to the tillering survey time described above, another row with uniform growth was selected, and 10 plants were collected consecutively in each repeated plot. The roots were removed, and the tillers were separated from the main stems. The treated samples were dried at 120°C for 20 min and then dried to a constant weight. The biomass (g m^{-2}) of minimum stems and tillers at different tillering positions was measured. Five plants with representative growth potentials were collected from

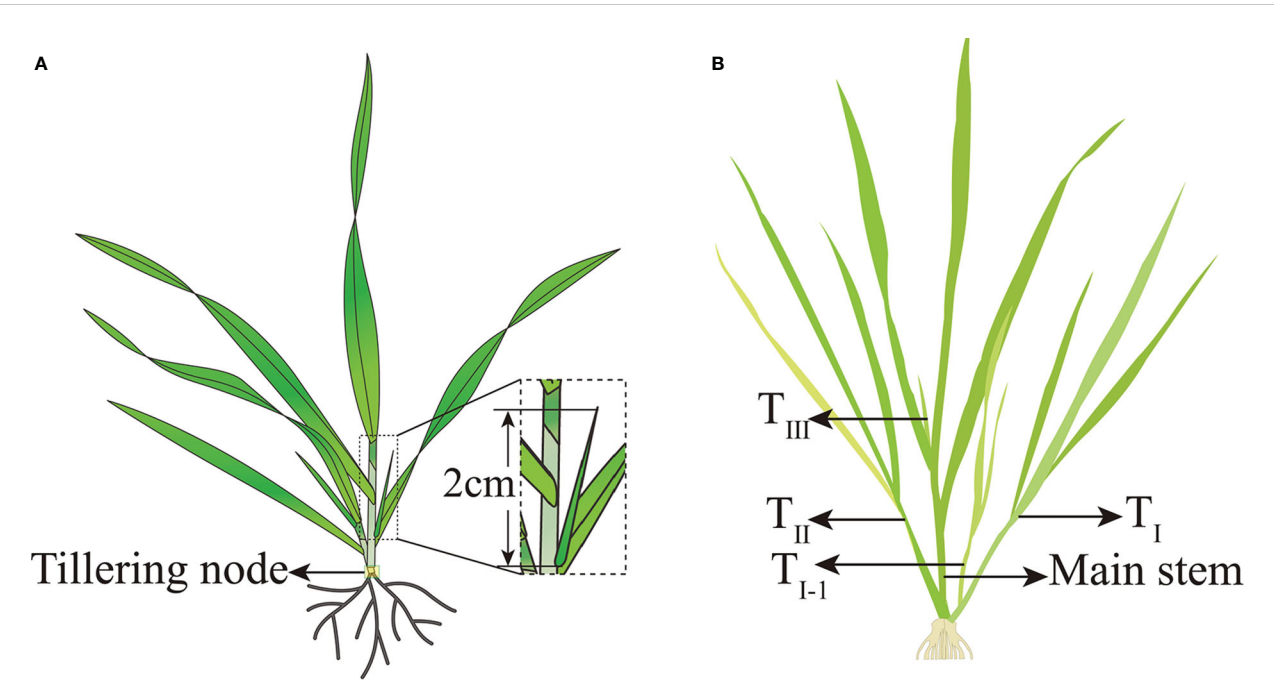


FIGURE 4
Sampling position, tiller criteria, and nomenclature of tillers for wheat. **(A)** Sampling position and tiller criteria for wheat; **(B)** nomenclature of tillers for wheat.

each plot during the sampling period, and the tiller nodes were cut and mixed (Figure 4A). Each repeated plot was used as a mixed sample and stored at −80°C.

The SL content in the tillering node was determined using high-performance liquid chromatography-mass spectrometry (Liu et al., 2011) on a TSQ Quantum triple quadrupole mass spectrometer and UltiMate 3000 RS chromatograph. Fresh wheat tiller buds (0.5 g) were mixed with 1 ml of pure methanol and zirconia grinding beads and ground for 5 min in liquid nitrogen. The supernatant was obtained by centrifugation at 8,000×*g* at 4°C for 10 min. The supernatant was filtered through a 0.22 μm membrane and used for analysis. SL chromatogram collection and integration were performed using Xcalibur software (version 3.0; Thermo Fisher Scientific, Waltham, MA, USA).

To determine the contents of IAA, zeatin riboside (ZR), and gibberellin (GA), 1.0 g of fresh wheat tiller bud was placed in a pre-cooled mortar, mixed with 6 ml of 80% pre-cooled methanol, ground in an ice bath, and extracted at 4.0 °C for 20 h. After centrifugation at 8,000×*g* at 4 °C for 10 min, the supernatant was collected to determine the IAA, GA, and ZR contents in the tillering node using an Agilent 1100 high-performance liquid chromatograph (Agilent, Santa Clara, CA, USA) (Liu et al., 2021a; Liu et al., 2022; Mao et al., 2022).

2.5 Percentage of effective tillers and grain yield

To determine the percentage of effective tillers, we counted all effective spikes (ES, kernels per spike ≥5) in the ripening stage (BBCH 89) in the sample segments marked as "BSN and tillers" and calculated the number of effective tillers in the overwintering stage (ET) and spring (EP) based on the color of nylon rope added to the tillers formed during different seasons. The percentage of ET at different growth stages was calculated using Equation 4.

$$\mathrm{ETgs}(\%) = \frac{\mathrm{ESgs}\ (\times 10^4\ \mathrm{hm}^{-2})}{\mathrm{Tgs}\ (\times 10^4\ \mathrm{hm}^{-2})} \times 100 \qquad \text{Equation 4}$$

ETgs is the effective tiller percentage at a specific growth stage, and ESgs and Tgs are the effective tillers and maximum tiller numbers at a specific growth stage.

Representative growth sample points (1.0 m length × 1.5 m width) were selected from each plot for GY measurements during the ripening stage (BBCH 89), calculated spikes per hectare, and then selected 20 randomly from all of them to count the grain number per spike. GY were calculated using Equation 5 after drying the samples to a standard moisture content of 12.5%.

$$\mathrm{GY}\ (\mathrm{kg\,hm}^{-2}) = \frac{\mathrm{ESW} \times \mathrm{GN} \times \mathrm{GW} \times 0.85}{100} \qquad \text{Equation 5}$$

where GY is the grain yield, ESW is the effective spike number during the entire growth period ($\times 10^4$ spikes hm^{-2}), GN is the kernel number per spike (kernels spike^{-1}), GW is the 1,000-grain weight (g), and 0.85 is an empirical coefficient.

2.6 Path analysis of tillering characteristics and yield factors

Wheat GY was considered the dependent variable, and 14 factors affecting its formation were considered independent variables (X_i): TP, TS, total tillers (TT), effective tillers in the pre-winter stage, effective tillers in spring, total effective tillers, effective spikes on the main stem, percentage of effective tillers in the pre-winter stage, percentage of effective tillers in spring, percentage of total effective tillers throughout the growth period, percentage of main stem spikes to total effective spikes, ESW, kernel number per spike (KN), and GW, abbreviated as X_1–X_{14}, respectively. After linear stepwise regression, X_i, with a partial regression coefficient of<0.05, was included in the regression model. The effect of X_i on GY is expressed as the direct path coefficient (DDC). The indirect effect of X_i on *GY* through X_i was expressed as the indirect path coefficient (IDC), calculated using Equation 6.

$$\mathbf{IDC}_{ijGY} = r_{ij} \times \mathbf{DDC}_{jGY} \qquad \text{Equation 6}$$

where IDC_{ijGY} is the IDC of X_i to *GY* through X_j ($X_{i\to jZ}$); r_{ij} is the Pearson correlation between X_i and X_j; and DDC_{jGY} is the direct path coefficient of X_j to GY.

2.7 Data processing and statistical analysis

Analysis of variance (ANOVA) was performed using SPSS 26.0 (IBM, NY, USA). Data are presented as the mean ± standard error (SE). A one-way ANOVA was conducted, and the Student's t-test was used to compare treatment means at the 5% level. Pearson correlation analysis and linear regression were used for correlation and path analysis using SPSS software (version 26.0).

3 Results

3.1 Optimizing plant spatial competition increases tillers number and biomass accumulation of winter wheat

3.1.1 Tillers number

Compared with those from the 15RS group, tillers from the 7.5RS group increased in the pre-winter stage for the SD1, SD2, and SD3 groups by 12.3%, 15.5%, and 15.6%, respectively, until the overwintering stage (annual average value over two years). There was no significant difference in the number of tillers between SD1 and SD2 sowing dates from D1 (trilobal stage) to D3. At the beginning of the D4 stage, the promotion effect of LSRE on tillering became obvious, and the number of 7.5RS tillers began to be significantly higher than that of 15RS, and the difference reached a maximum at the WS. LSRE began to promote tillering at the beginning of the D4 stage significantly, and the tiller number of 7.5RS started to become significantly higher than that of 15RS, and the difference reached a maximum until the overwintering stage. Under SD3 treatment, the tiller number differed between the

two-row spacing treatments at the D3 stage; 7.5RS was significantly higher than 15RS, and the difference reached a maximum until WS (Figures 5A–F).

Analysis of the stem composition under different row spacings before the overwintering stage showed that following LSRE treatment, the percentage of the third, fourth, and other tiller numbers of the main wheat stem at the SD1 sowing date and the percentage of the third level of the main stem and other tiller numbers at the SD2 sowing date increased. The average increase for the three sowing treatments was 9.8 percentage points (2-year average) (Figures 5G, H).

3.1.2 Tiller biomass accumulation

From the trilobal stage (D1) to the overwintering stage (WS) during SD1 and SD2, the tiller biomass of all treatments increased slowly and rapidly from D1 to D2 and D2 to D4, respectively (Figures 6A–D). The difference between 7.5RS and 15RS was slight from D1 to D2 and then gradually increased; 7.5RS showed a faster biomass accumulation rate. At WS, the biomass accumulation difference between the two-row spacing groups reached a maximum (Figures 6A–F). The biomass of 7.5RS tillers in the three sowing-date groups increased by 20.9% compared to that of 15RS tillers (2-year average). In addition, the difference in biomass accumulation between the two-row spacings increased with delays in the sowing date: compared with biomass accumulation in the 15RS wheat field, that of the 7.5RS field under SD1 treatment increased by 17.9%, SD2 treatment increased by 18.7%, and SD3 treatment increased by 23.1% (2-year annual average). In the winter stage (Figures 6G, H), the proportion of 7.5RS main stem biomass in plants for each sowing date was significantly lower than that of 15RS, with SD1, SD2, and SD3 reduced by 6.0%, 7.3%, and 8.5%, respectively (2-year average). The proportion of biomass of the third tiller and other tillers (fourth tiller and second tiller on the main stem) was significantly higher than that of the 15RS treatment, with an average decrease of 5.7% (2-year average).

3.1.3 Accumulated temperature of tillers

As shown in Figure 7A, for the two-row spacing treatments, the AT required for the formation of the first tiller (T_I) on the main stem of wheat for the three sowing-date groups did not differ significantly between the two-row spacing treatments, whereas that for forming the second tiller (T_{II}) differed. The AT required for SD2 and SD3 for 7.5RS was significantly lower than that for 15RS, with average decreases of 6.3% and 13.5%, respectively (2-year average). Compared with the AT required for forming higher tillers, LSRE promoted tillering. The AT required for forming the third and fourth tillers of the main stem at SD1 decreased by 7.2% and 11.2% for 7.5RS compared with those at 15RS, respectively, in the two test years.

Overall, LSRE reduced the AT required for wheat tillering in the pre-winter stage by 5.7, 7.4, and 8.0% for SD1, SD2, and SD3, respectively (2-year average). Furthermore, according to the linear fitting of the AT required for tillering in the different row spacing treatments, with the postponement of the tillering sequence, the AT required for 15RS tillering significantly increased by more than 7.5RS. The slopes of the fitting equations were 3.78 and 0.55 for 15RS and 7.5RS, respectively (Figure 7B). Under the same row spacing, the AT required for tillering at the same tiller position was significantly increased when the sowing date was delayed.

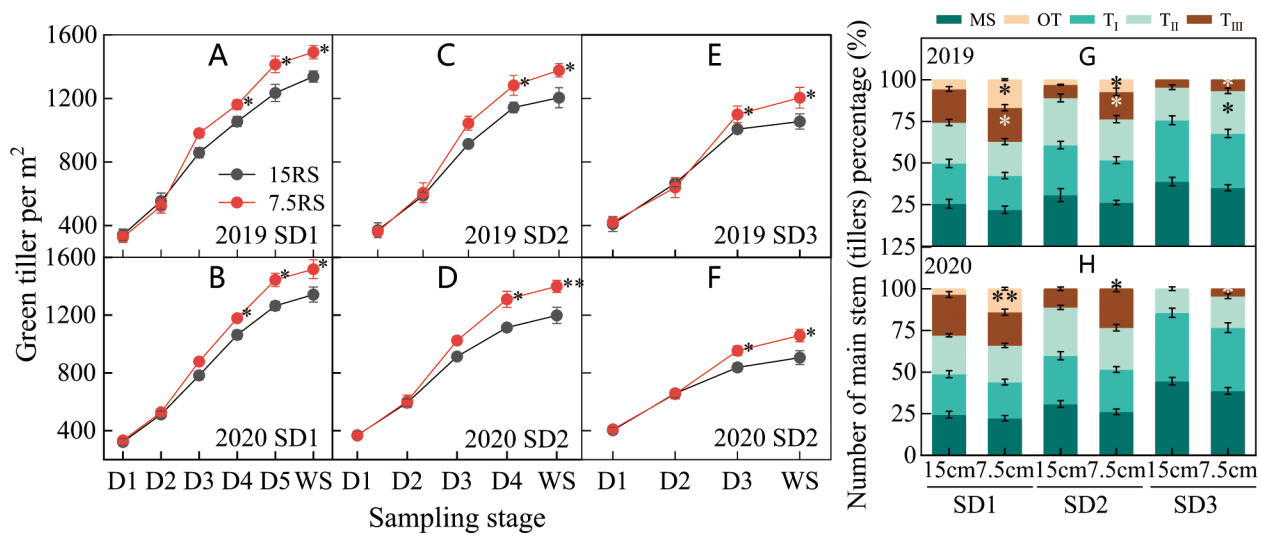


FIGURE 5
Number of wheat tillers at pre-winter stage under different sowing dates and row spacings. **(A, C, E)** show the green tillers number for wheat in pre-winter stage for SD1, SD2 and SD3 in 2019, respectively. **(B, D, F)** show the green tillers number for wheat in pre-winter stage of SD1, SD2 and SD3 in 2020, respectively. **(G, H)** show the quantity proportion of main stems and tillers for wheat in 2019 and 2020, respectively. "*, **" indicate significant difference at 0.05 level. $^{**}P<0.01$; $^{*}P<0.05$. Significant markers represent the differences between 7.5 cm and 15 cm row spacing treatments at the same sowing date and observation period. Number of wheat tillers at pre-winter stage under different sowing dates and row spacings.

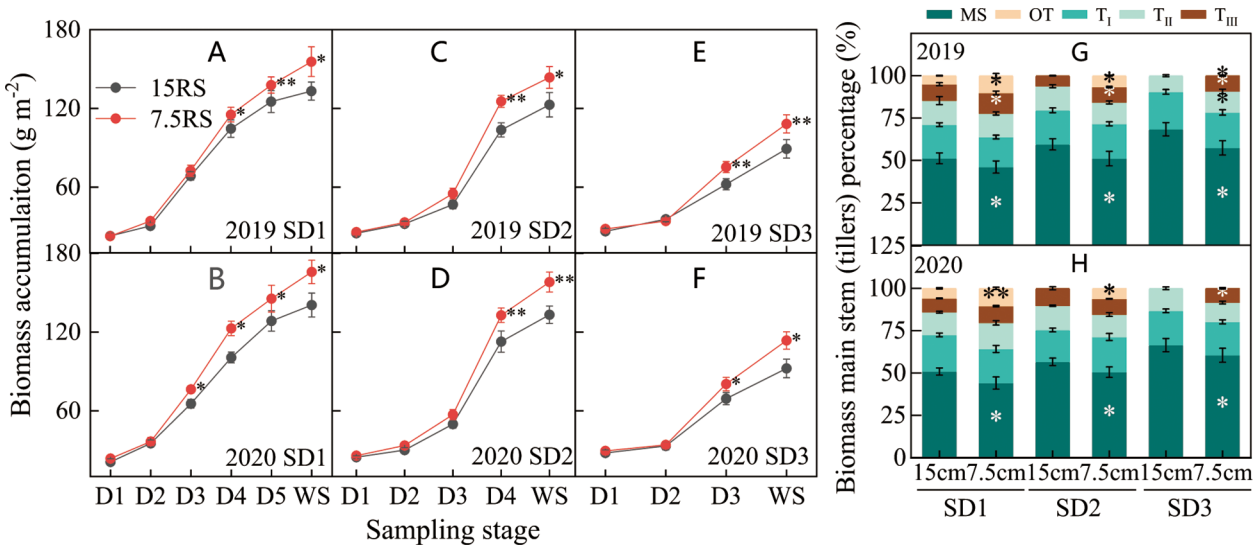


FIGURE 6
Biomass of wheat tillers at pre-winter stage under different sowing dates and row spacings. **(A, C, E)** show the dry matter accumulation for wheat in pre-winter stage for SD1, SD2 and SD3 in 2019, respectively. **(B, D, F)** show the dry matter accumulation for wheat in pre-winter stage of SD1, SD2 and SD3 in 2020, respectively. **(G, H)** show the biomass accumulation proportion of main stems and tillers for wheat in 2019 and 2020, respectively. "*, ** "indicate significant difference at 0.05 level. **P<0.01; *P<0.05. Significant markers represent the differences between 7.5 cm and 15 cm row spacing treatments at the same sowing date and observation period. MS stands for main stem, T_I, T_{II}, and T_{III} represent the first, second, and third tillers of the main stem respectively, and OT represents the fourth and other tillers of the main stem.

3.2 Optimizing plant spatial competition can change phytohormone content in tillering node of winter wheat

3.2.1 Contents of IAA and ZR

The IAA content of the wheat tillering nodes first decreased and then increased in each treatment from the trilobal stage (D1) to WS. The minimum IAA content during SD1 and SD2 was at the D3 stage, but the lowest IAA content during SD3 was at the D2 stage. At SD1 and SD2, the IAA content in the 15RS treatment decreased slightly (27.4%) at the D1–D3 stages, whereas that of the 7.5RS treatment decreased significantly (48.3%). After D3, the IAA content of the two-row spacing treatments increased rapidly, although 7.5RS remained lower than 15RS (20.0% lower than the 2-year average). For SD3, in the D1–D2 stages, the IAA content of the 7.5RS treatment decreased more significantly by 28.2%, and its decline was higher than that of 15RS (2-year average). During D2–WS, the IAA content of the two-row spacing groups began to increase. However, the growth rate of 7.5RS was much lower, at 52.7% lower than 15RS on average (Figures 8A–F).

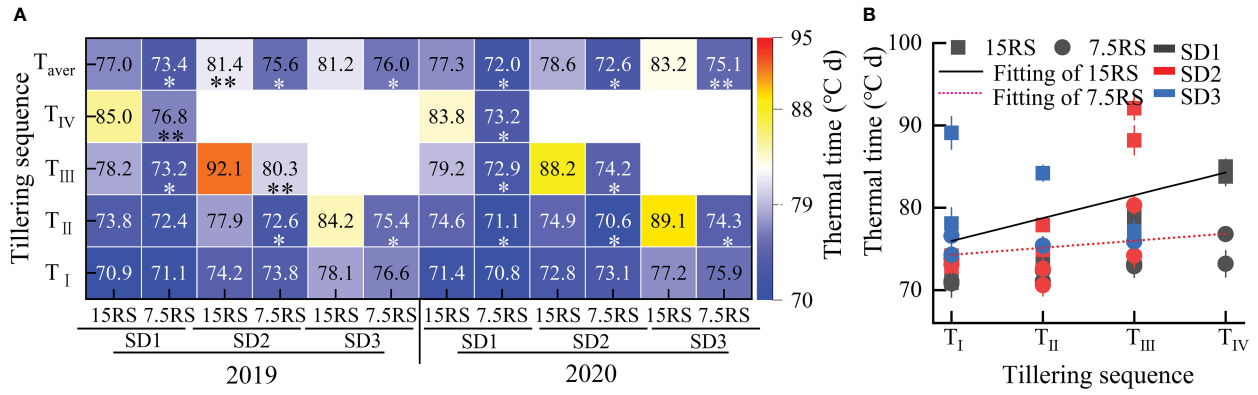


FIGURE 7
Accumulated temperature required for tillering at different tillers of wheat at pre-winter stage under different sowing dates and row spacings. **(A)** shows the AT required for different treatments of tillers at each level for wheat. **(B)** shows the linear fitting of AT required for wheat tillers with 7.5RS and 15RS. "*, ** "indicate significant difference at 0.05 level. ** P<0.01; *P<0.05. Significant markers represent the differences between 7.5 cm and 15 cm row spacing treatments at the same sowing date and observation period. T_I, T_{II}, T_{III}, and T_{IV} represent the first, second, and third tillers of the main stem, respectively. T_{aver} represents the average of T_I, T_{II}, T_{III}, and T_{IV}. 15RS and 7.5RS fitting equations are Y_{15RS} = +70.66+3.78x and $Y_{7.5RS}$ = 73.3 + 0.55x, respectively.

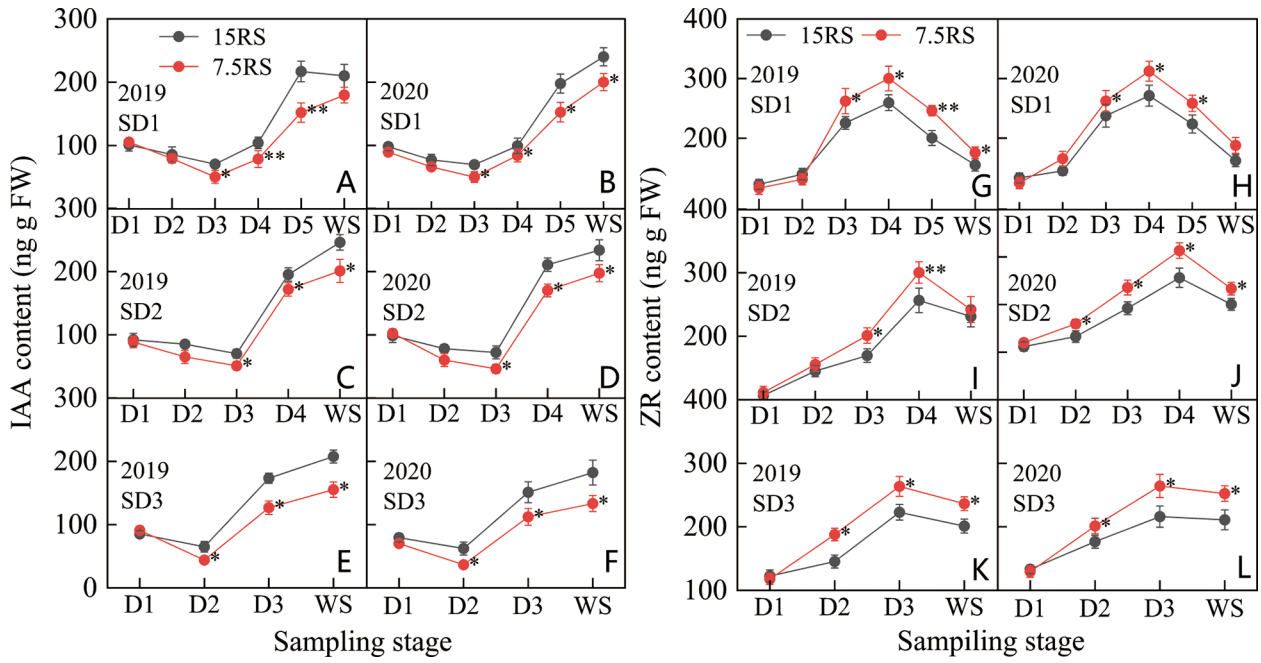


FIGURE 8
IAA and ZR content of wheat tillers at pre-winter stage under different sowing dates and row spacings. **(A, C, E)** show the IAA content for wheat in pre-winter stage for SD1, SD2 and SD3 in 2019, respectively. **(B, D, F)** show the IAA content for wheat in pre-winter stage of SD1, SD2 and SD3 in 2020, respectively. **(G, I, K)** show the ZR content for wheat in pre-winter stage for SD1, SD2 and SD3 in 2019, respectively. **(H, J, L)** show the ZR content for wheat in pre-winter stage of SD1, SD2 and SD3 in 2020, respectively. "*, **" indicate significant difference at 0.05 level. $^{**}P<0.01$; $^{*}P<0.05$. Significant markers represent the differences between 7.5 cm and 15 cm row spacing treatments at the same sowing date and observation period.

As shown in Figures 8G–L, the change in the ZR content in D1–WS for the three sowing-date groups was opposite to that of the IAA content; the ZR content first increased and then decreased. The highest ZR content in SD1 and SD2 was observed at the D4 stage, whereas that in SD3 was at the D3 stage. The peak value of ZR content for the three sowing-date groups was 19.9% higher for 7.5RS than for 15RS over two years. After the peak value, the ZR content of all treatments decreased rapidly; however, in the overwintering period, the ZR content of 7.5RS remained significantly higher than that of 15RS by 15.4%, 15.3%, and 19.4% for SD1, SD2, and SD3, respectively. After the peak ZR content was reached in each treatment group, 7.5RS showed a more significant decrease; the average decrease for the three sowing-date groups was 4.0% higher than that for 15RS (2-year average). Overall, in the pre-winter stage of winter wheat, lower IAA and higher ZR content were more conducive to the generation of wheat tillers. The IAA content of 7.5RS in the three sowing-date groups was 19.4% lower than that of 15RS, whereas the ZR content was 13.1% higher (2-year average).

3.2.2 Contents of SLs and GA

As shown in Figures 9A–F, the SL content in the tillering node of wheat under 7.5RS treatment was significantly higher than that under 15RS treatment (25.9% higher in the two years) at the D4 to WS stage of SD1; the maximum difference was observed at the D4 stage (37.9% higher in the two years). At the SD2 sowing date, the difference in SL content between the 7.5RS and 15RS treatments was like that at the SD1 sowing date, demonstrating that the difference was small before the D2 stage and large after the D2 stage. Additionally, after the D4 stage, the SL content in the 7.5RS treatment was significantly higher than that in the 15RS treatment (31.5% higher than that in the two years), with a peak difference of 21.5%. During the D3–WS stage of SD3 treatment, the SL content in the 7.5RS treatment was 20.0% higher than that in the 15RS treatment, with the largest difference observed in the D3 stage; 7.5RS was 22.5% higher than that in the 15RS. Overall, from the D1 to WS stage, the SL content in the wheat tiller nodes of the two-row spacing treatments for the three sowing-date groups first increased and then decreased; the LSRE treatment increased the SL content in the wheat tiller nodes by 17.5% in the two years. In addition, the SL content increased when the sowing date was delayed.

At the beginning of the D1 stage, the GA content in the tiller nodes of all treatments first decreased and then increased during the growth period. At SD1, the minimum GA content was observed at the D4 stage, whereas the minimum GA content of SD2 and SD3 appeared at the D3 stage (Figures 9G–L). Compared with the two-row spacing treatments, the GA content decreased from the trilobal stage to a minimum value, with SD1, SD2, and SD3 showing 40.5%, 1.0%, and 25.8% lower values in 7.5RS than in 15RS, respectively. The GA content of 7.5RS was significantly lower than that of 15RS for SD1 from D2 to D4, SD2 from D2 to D3, and SD3 from D3. For the different sowing dates, the lowest GA content tended to increase with a delay in the sowing date; the increase in 7.5RS was more obvious than that in 15RS. From SD1 to SD3, the lowest GA content of 7.5RS increased by 16.3%, whereas that of 15RS increased by only

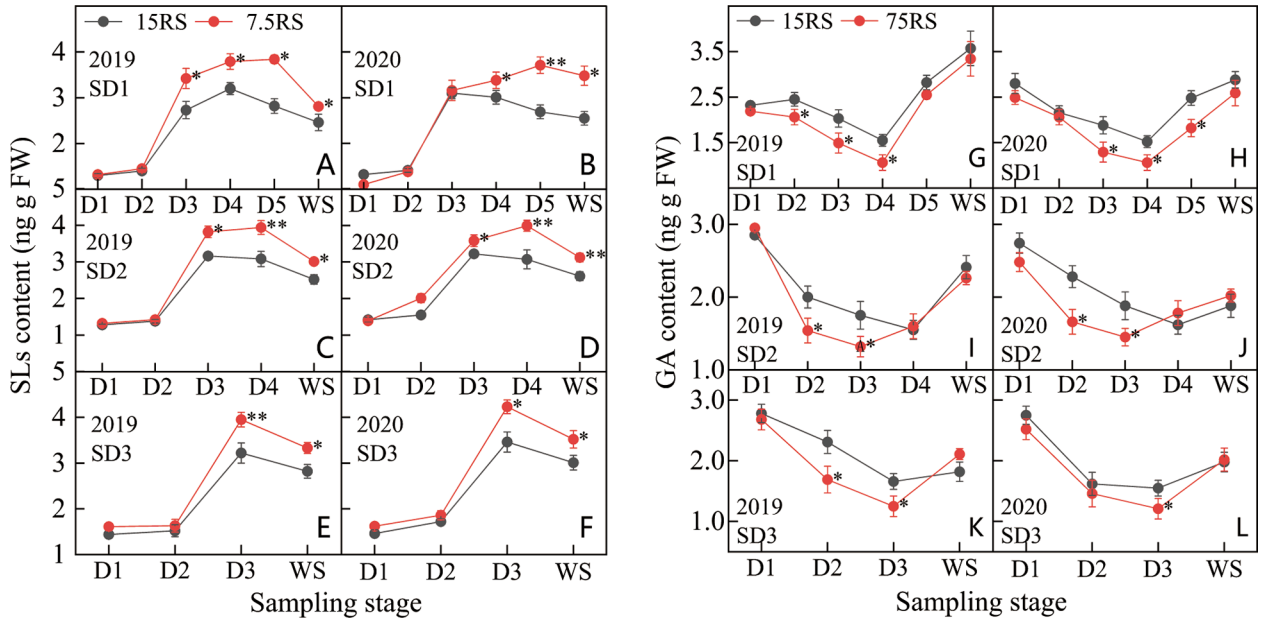


FIGURE 9
SL and GA content of wheat tillers at pre-winter stage under different sowing dates and row spacings. **(A, C, E)** show the SLs content for wheat in pre-winter stage for SD1, SD2 and SD3 in 2019, respectively. **(B, D, F)** show the SLs content for wheat in pre-winter stage of SD1, SD2 and SD3 in 2020, respectively. **(G, I, K)** show the GA content for wheat in pre-winter stage for SD1, SD2 and SD3 in 2019, respectively. **(H, J, L)**show the GA content for wheat in pre-winter stage of SD1, SD2 and SD3 in 2020, respectively. "*, **" indicate significant difference at 0.05 level. $^{**}P<0.01$; $^{*}P<0.05$. Significant markers represent the differences between 7.5 cm and 15 cm row spacing treatments at the same sowing date and observation period.

4.6%. There was no significant difference in GA content between the two rows at each sowing date until the overwintering stage. LSRE reduced GA content, with decreases of 15.7%, 9.1%, and 9.3% for SD1, SD2, and SD3, respectively.

3.2.3 IAA/ZR and SLs/GA

IAA/ZR (I/Z) decreased and then increased from the trilobal stage (D1) to WS. For SD1 and SD2, the minimum I/Z value appeared at the D3 stage, but the minimum value during SD3 appeared earlier (D2) (Figures 10A–F). At the lowest I/Z point, the 7.5RS group was lower than the 15RS group by 36.7%, 43.0%, and 48.0% for SD1, SD2, and SD3, respectively (2-year average). After the lowest point, the I/Z of each treatment began to increase rapidly, but the increase in 7.5RS was lower than that in 15RS, and the I/Z increase of 7.5RS was 35.7% lower than that of 15RS until the WS. The I/Z of 7.5RS treatment was significantly lower than that of 15RS at the beginning of the SD1 D5 stage and SD2 and SD3 D3 stages in WS, with decreases of 31.4%, 29.5%, and 35%, respectively (2-year average). Overall, reducing row spacing could reduce the I/Z of wheat tiller nodes during the pre-winter stage. The 7.5RS group showed 22.9%, 24.4%, and 30.1% lower values than the 15RS group at SD1, SD2, and SD3, respectively, with an average decrease of 25.8%.

As shown in Figures 10G–L, the SLs/GA (S/G) of D1-WS and each treatment first increased and then decreased. From D1 to D2, the S/G of the two-row spacing treatments increased slightly and showed no significant differences in the three sowing-date groups. At the beginning of the D2 stage, the S/G values of all treatments increased significantly. Except for SD1 in 2020, the peak value of the other treatments was observed in the D3 stage, with 7.5RS showing a higher value than 15RS. After D3, SD1 treatment continued to increase in 2020 and reached a peak in D4. The S/G of 7.5RS remained significantly higher than that of 15RS. In 2020, in addition to SD1, the S/G of the other sowing date treatments decreased to varying degrees from D3 to WS, whereas 7.5RS further decreased. There was no significant difference in the S/G between the two-row spacing treatments until WS. Overall, reducing row spacing improved the S/G of the tillering node of wheat in the pre-winter stage, and SD1, SD2, and SD3 led to increases of 51.7%, 32.1%, and 33.9%, respectively. A lower I/Z ratio and a higher S/G ratio significantly promote wheat tillering. In particular, under late sowing conditions, the lower I/Z and higher S/G that appeared earlier after the trilobal stage can promote tillering under lower temperatures at the tillering stage, thus achieving a sufficient population.

3.3 Correlation between endogenous phytohormone content and tillering characteristics under optimal plant distribution of winter wheat

As shown in Figures 11A–C, the number of tillers (TP) in the pre-winter stage was significantly positively correlated with SLs, IAA, and ZR at the tiller nodes for the three sowing-date groups ($P<0.05$), and the average correlation coefficients (r) were 0.85, 0.68, and 0.79, respectively. Among them, the correlation between the

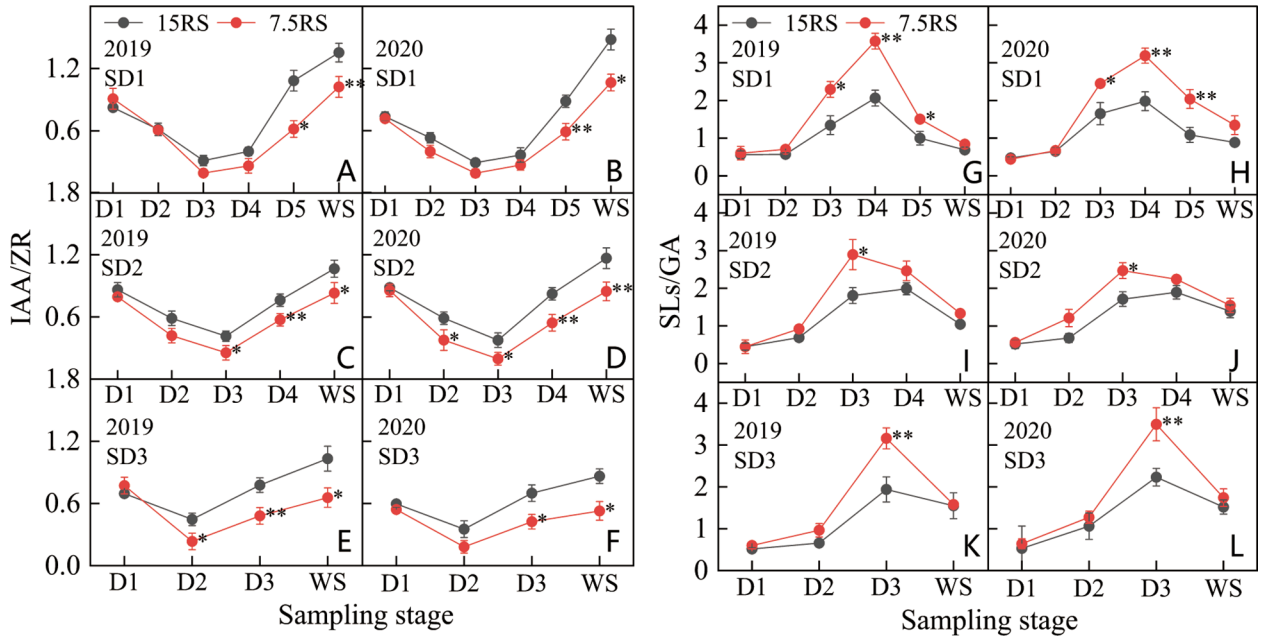


FIGURE 10
IAA/ZRand SLs/GA of wheat tillers at pre-winter stage under different sowing dates and row spacings. **(A, C, E)** show the IAA/ZR for wheat in pre-winter stage for SD1, SD2 and SD3 in 2019, respectively. **(B, D, F)** show the IAA/ZR for wheat in pre-winter stage of SD1, SD2 and SD3 in 2020, respectively. **(G, I, K)** show the SLs/GA for wheat in pre-winter stage for SD1, SD2 and SD3 in 2019, respectively. **(H, J, L)** show the SLs/GA for wheat in pre-winter stage of SD1, SD2 and SD3 in 2020, respectively. "*, **" indicate significant difference at 0.05 level. $^{**}P<0.01$; $^{*}P<0.05$. Significant markers represent the differences between 7.5 cm and 15 cm row spacing treatments at the same sowing date and observation period.

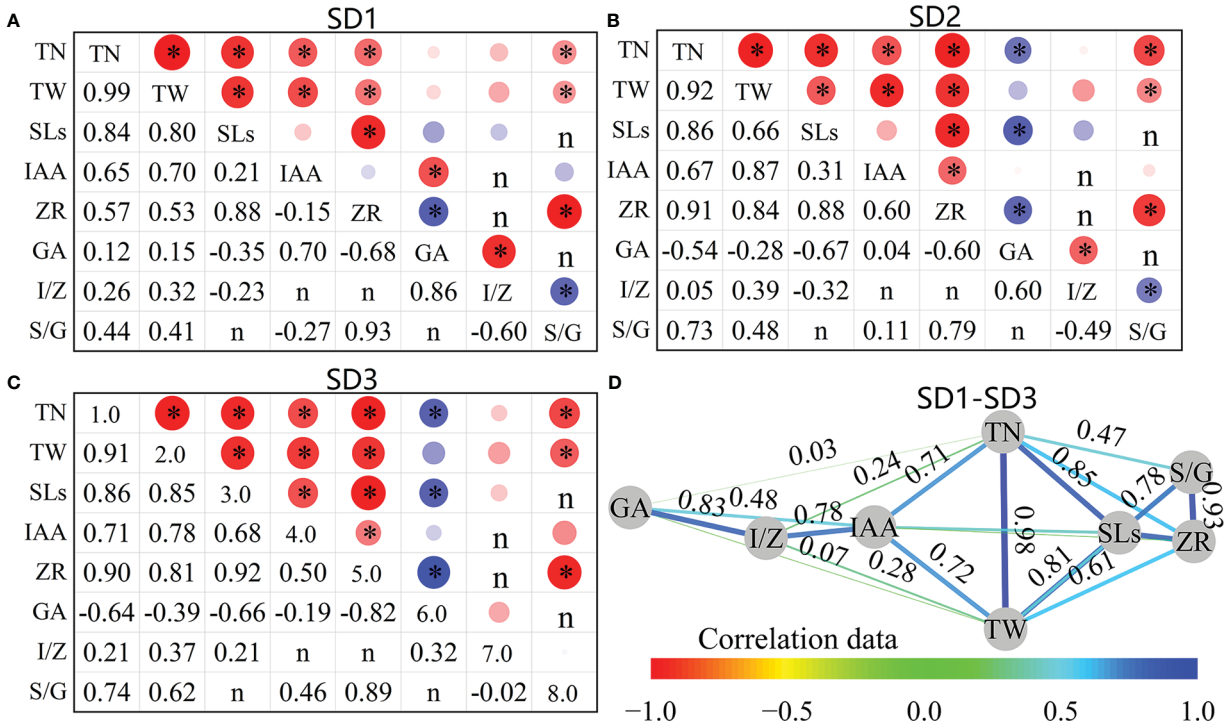


FIGURE 11
Correlation between tiller number, biomass and phytohormone contents of wheat tillers at pre-winter stage under different sowing dates and row spacings. **(A–C)** show the correlation data of SD1, SD2 and SD3, respectively. (D) shows the average correlation data of SD1-SD3. "*r*" is Person two tailed correlation coefficient, "*" indicate significant difference at 0.05 level. $^{*}P<0.05$. **(D)** Figure D is the network diagram of the correlation coefficient adjacency matrix between indicators, and the node layout adopts the Kamada–Kawai algorithm. The connection represents *r* between each two nodes. The thicker the connection, the greater the absolute value of *r*, and the data of **(D)** comes from the mean of SD1, SD2, and SD3.

TN and SLs of the SD1 treatment was the highest ($r = 0.84$), whereas that between the SD2 and SD3 treatments was the highest ($r = 0.91$ and 0.90, respectively). Tiller biomass (TW) was also significantly positively correlated with SLs, IAA, and ZR in the three sowing-date groups ($P<0.05$), with average correlation coefficients of 0.77, 0.78, and 0.73, respectively (Figure 11D). The TW of SD1 and SD3 showed the highest correlation coefficient with SLs ($r = 0.8$ and 0.85, respectively), whereas the TW of SD2 showed the highest correlation coefficient with IAA ($r = 0.87$). The TP and GA of SD2 and SD3 were significantly negatively correlated, whereas SD1 showed no significant correlation. There was also a strong correlation between tiller growth and hormone content; for example, there was a significant correlation between TP, TW, and S/G ($r = 0.64$ and 0.50).

The contents and ratios of different hormones were correlated at the tillering stage before winter. SLs in the three sowing-date groups were significantly positively correlated with ZR ($r = 0.89$) and significantly positively correlated with IAA only at SD3 ($r = 0.68$). SLs were negatively correlated with GA in all three sowing-date groups, and SD2 and SD3 were significantly correlated ($P<0.05$). IAA and ZR showed significant positive correlations with SD2 and SD3 ($r = 0.55$) but exhibited no significant correlation with SD1 ($P > 0.05$). The ZR and GA of the three sowing-date groups were negatively correlated (mean $r = 0.70$). Further analysis of the correlation between tiller growth characteristics and various hormones at different sowing dates showed that when the sowing date was postponed, the correlation between TP and TA, TW and IAA, and ZR and S/G showed an increasing trend. In contrast, the correlation between TP and GA changed from insignificant at suitable sowing dates to a significant negative correlation with delayed sowing dates. TN and TW were significantly and positively correlated with SLs, IAA, ZR, and SLs/GA, but not with IAA/ZR for the three sowing-date groups, and TP and GA were significantly negatively correlated when the sowing date was delayed.

3.4 Optimizing plant spatial distribution increased the number of effective tillers and yield of winter wheat

3.4.1 Number of effective tillers

As shown in Table 2, when the year (Y) was used as a single source of variation, the number of tillering panicles (ETP and ETS)

TABLE 2 Characteristics of tillering and spike formation of wheat under different sowing dates and row spacings.

Year	SD	RS (cm)	BSN ($\times10^4$/hm^2)	Number of tillers ($\times10^4$ hm^2)			Effective tillers number ($\times10^4$ hm^2)			Percentage of effective tillers (%)		
				TP	TS	TT	ETP	ETS	ETT	PETP	PETS	PETT
2019–2020	SD1	15	342.0	921.0	340.8	1261.8	372.0	75.0	447.0	40.4	22.1	35.5
		7.5	327.0	1156.5**	324.0	1480.5**	411.0*	60.0**	471.0	35.7*	18.5**	31.9*
	SD2	15	370.5	772.5	460.8	1233.3	313.5	79.5	393.0	40.4	17.2	31.8
		7.5	361.5	1015.5**	426.0	1441.5**	363.0	43.5**	406.5	35.7	10.2**	28.2*
	SD3	15	409.5	597.0	399.6	996.6	192.0	94.5	286.5	32.1	23.6	28.7
		7.5	423.0	781.5**	424.8	1206.3**	240.0*	60.0**	300.0	30.7	14.1**	24.9*
2020–2021	SD1	15	328.5	967.5	386.4	1353.9	369.0	88.5	457.5	38.2	22.9	33.8
		7.5	339.0	1227.0**	315.6*	1542.6**	405.0	51.0**	456.0	33.0*	16.1**	29.6*
	SD2	15	369.0	849.0	415.2	1264.2	318.0	64.5	382.5	37.5	15.5	30.3
		7.5	366.0	1045.5**	381.6	1427.1**	367.5**	36.0**	403.5	35.1	9.4**	28.3
	SD3	15	403.5	502.5	412.8	915.3	265.5	54.0	319.5	53.0	13.1	35.0
		7.5	409.5	709.5**	404.4	1113.9**	304.5	25.5**	330.0	43.3*	6.3**	29.8*
Year (Y)			ns	ns	ns	ns	**	**	ns	**	**	ns
Sowing date (SD)			**	**	**	**	**	**	**	*	**	**
Row space (RS)			ns	**	**	**	**	**	ns	**	**	**
Y×SD			ns	**	**	**	**	**	ns	**	**	**
Y×RS			ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
SD×RS			ns	ns	*	ns	ns	ns	ns	ns	**	ns
Y×SD×RS			ns	ns	ns	ns	ns	**	ns	ns	**	ns

RS and BSN represent row space and basic seedling number. TP, TS, and TT represent tillers number in pre-winter stage, spring, and the whole growth period, respectively. ETP, ETS and ETT represent effective tillers in pre-winter stage, spring, and the whole growth period, respectively. PETP, PETS and PETT represent the percentage of effective tillers in pre-winter stage, spring, and the whole growth period, respectively. "*, **" indicate significant difference at 0.05 level. **$P<0.01$; *$P<0.05$, ns, not statistically significant at $P<0.05$.

and the ratio of effective tillers (PETP and PETS) were significantly different (P<0.01), with no significant differences in the other indicators. Comparing different RS and SD, except for the lack of a significant difference between different row spacings for BSN and no significant difference for ETT between different SD, other indicators did not significantly differ when row spacing and SD were the single sources of variation. Further analysis of the interaction between different sources of variance revealed significant differences in the Y × SD interaction, except for those in BSN and ETT. The indicators showed no significant differences in the Y × RS interaction. Under the SD × RS interaction, TS and PETS showed significant and very significant differences, respectively, whereas the other indices showed no significant difference. However, under the interaction of factors, such as Y × SD × RS, as the source of variation for different tillering characteristics, only ETS and PETS showed highly significant differences, whereas the other indicators showed no significant differences.

LSRE treatment had no significant effect on BSN but significantly increased TP and TT by 29.9% and 17.3%, respectively, for the three sowing-date groups (2-year average). Although the TS of wheat treated with LSRE decreased, the difference was not statistically significant. LSRE increased the number of tillers and spikes by 3.7% for the three sowing-date groups, and ETP increased significantly by 15.2% for the three sowing-date groups (2-year average). LSRE also significantly inhibited ETS by 40.2% (2-year average) in the three sowing-date groups. The percentage of tillers and spikes at each stage of the LSRE treatment decreased by 5.0% (average of three sowing-date groups in two years).

Overall, the interaction of RS, SD, and Y × SD as sources of variation significantly affected the tillering and tiller panicle formation characteristics, whereas Y, as a single source of variation, significantly impacted ETP, ETS, PETP, and PETS.

3.4.2 Grain yield factors

Analysis of the interaction effects (Table 3) showed that the kernel number per spike (KN) and 1,000 grain weight (GW) significantly differed between years (P<0.01), whereas the ESW and GY showed no significant differences (P >0.05). Except for KN, the SD and row spacing significantly affected the other two yield factors (GW and ESW). The single effects of yield on Y, SD, and RS were significant (P<0.05) or highly significant (P<0.01). The reduction in row spacing tended to increase ESW and KN, and the average heights of plants from the three sowing-date groups were 1.9% and 0.4%, respectively, but the differences were not significant. In addition, after delaying the sowing date, the LSRE treatment had a more noticeable effect on the ESW increase. The LSRE treatment improved GW by 4.7% and 3.2% in 2020 and 2021, respectively. SD1 and SD2 in 2020 showed significant and highly significant differences, respectively, whereas no significant differences were observed for the other sowing dates. The LSRE treatment significantly increased yield by 6.8% and 5.7% in 2020 and 2021, respectively. However, the yield increase effect of the different sowing dates differed between the two trial years. In 2020, the yield increase effect of LSRE on a suitable sowing date was significant (7.6%), whereas in 2021, the yield increase effect of SD2 was significant (6.3%). LSRE can improve ESW and KN to some extent and significantly improve GW, although most

TABLE 3 Grain yield factors of wheat under different sowing dates and row spacings.

Sowing date	Row space (cm)	2019–2020				2020–2021			
		ESW km^{-2} ($\times10^{-4}$)	KN	GW (g)	GY (kg hm^{-2})	ESW km^{-2} ($\times10^{-4}$)	KN	GW (g)	GY (kg hm^{-2})
SD1	15	789.0	32.6	48.1	10,510.8	786.0	35.8	45.8	10,964.5
SD1	7.5	798.0	33.1	50.4*	11,311.1*	795.0	36.0	47.3	11,509.8*
SD2	15	763.5	33.3	49.0	10,592.0	751.5	35.8	46.6	10,651.7
SD2	7.5	768.0	32.6	52.8**	11,238.9*	769.5	36.1	48.0	11,322.6**
SD3	15	696.0	32.2	52.2	9,941.4	723.0	34.2	48.7	10,229.0
SD3	7.5	723.0	32.6	53.0	1,0615.0*	739.5*	34.3	50.3	10,836.4**
Year (Y)		ns	**	**	ns	ns	**	**	ns
Sowing date (SD)		**	**	**	**	**	**	**	**
Row space (RS)		**	ns	**	**	**	ns	**	**
Y × SD		ns	ns	ns	ns	ns	ns	ns	ns
Y × RS		ns	ns	ns	ns	ns	ns	ns	ns
SD × RS		ns	ns	ns	ns	ns	ns	ns	ns
Y × SD × RS		ns	ns	ns	ns	ns	ns	ns	ns

ESW, KN, GW, and GY represent total effective spikes in the whole growth period, kernel number per spike, 1,000-grain weight, and grain yield, respectively. "*, **" indicate significant difference at 0.05 level. **P<0.01; *P<0.05, ns, not statistically significant at P<0.05.

treatments lead to no significant difference. LSRE significantly increased yield by 6.3% for the three sowing-date groups.

3.5 Optimizing plant spatial distribution changes the contribution of tillers and yield factors to yield

After linear regression of 14 factors (X_i) with a substantial impact on the GY following each treatment, seven independent variables (action factors), including X_1 (TP), X_2 (TS), X_7 (effective spikes on the main stem), X_9 (percentage of effective tillers in spring), X_{12} (ESW), X_{13} (KN), and X_{14} (GW), were included in the regression model, and the DDC and IDC of X_i of the GY were determined (Figures 12A–F).

For SD1, four factors were included in the regression model of 15RS: X_1 and X_9 negatively affected GY (-1.011 in total), and X_7 and X_{13} had a positive effect (0.931 in total). X_7 had the most considerable indirect effect on GY through X_9 ($X_7 \rightarrow {}_9$) (IDC = 0.465), followed by $X_1 \rightarrow {}_{13}$ and $X_1 \rightarrow {}_9$ (IDC mean 0.380). In the 7.5 RS treatment, X_{13}, X_{14}, X_{12}, and X_1 were the main factors affecting *GY* (all positive). The effects of X_{13} and X_{14} were significant (mean DDC = 1.097), and X_{13} was significantly related to GY ($P<0.05$). X_{13} and X_{14} strongly negatively affected GY (mean IDC = -0.852). X_{13} and X_1 had significant positive and negative effects on GY through X_1 (IDC = 0.646 and −0.655, respectively).

For SD2, X_{13} and X_{14} were included in the two-row spacing treatment return models. Compared with 15RS, the DDC value of X_{13} versus GY under 7.5RS treatment increased by 21.6%, whereas X_{14} versus *GY* decreased significantly. For 15RS and 7.5RS, $X_{13} \rightarrow {}_2$ and $X_{13} \rightarrow {}_9$ had the most substantial effects on GY (IDC = 0.393 and 0.465, respectively).

For SD3, seven independent variables (X_{12}, X_{13}, X_{14}, X_1, X_7, X_9, and X_{14}) were included in 15RS and 7.5RS regression models. Among these, X_{12} and X_{14} in the 15RS and 7.5RS groups had the greatest effect on GY (DDC = 0.835 and −0.838, respectively). For the 15RS group, $X_{13} \rightarrow {}_9$ had the strongest effect on GY (IDC = 0.498), whereas $X_{14} \rightarrow {}_7$ of the 7.5RS group had the largest effect on GY (IDC = −0.269). Overall, the independent variable with the highest frequency of 15RS treatment in the regression model was X_{13}, whereas those in the 7.5RS treatment were X_1 and X_{14}. The direct effect of X_{13} on GY increased with a delay in the sowing date, whereas that of X_{14} decreased. Under 15RS treatment, X_1, X_{13}, and

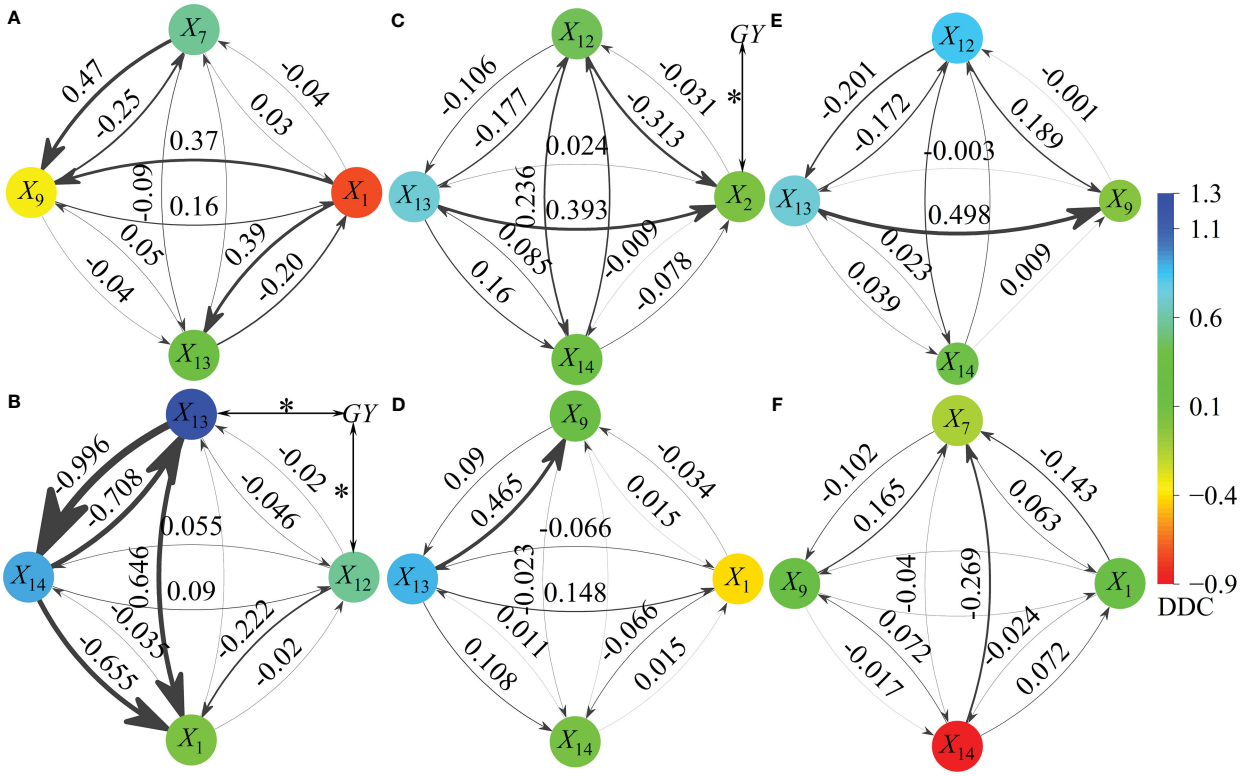


FIGURE 12
Path effect of different action factors (independent variables) on grain yield (dependent variables). X_1, X_2, X_7, X_9, X_{12}, X_{13}, X_{14}, and GY represent tillers in pre-winter stage (TP), tillers in spring stage (TS), effective spikes on main stem (ESM), percentage of effective tillers in spring (PESS), effective spikes in the whole growth period (ESW), kernel number per spike (KN), 1,000-grain weight (GW) and grain yield (GY) respectively. ANOVA: *P <0.05, it is the Pearson's two tailed correlation coefficient between action factor (X_i) and the GY. The arrow points to the effect of one factor on GY through another factor, and the thicker the line, the greater the indirect path coefficient (absolute value of IDC). Panels **(A, B)** show 15RS and 7.5RS treatment under SD1, respectively. Panels **(C, D)** show 15RS and 7.5RS treatment under SD2, respectively. Panels **(E, F)** show 15RS and 7.5RS treatment under SD3, respectively.

X_7 and X_1, X_{13}, and X_{13} in the SD1, SD2, and SD3 groups had the strongest direct effects on GY, whereas those for 7.5RS treatment were X_{14}, X_{13}, X_{14}, X_{13}, X_{13}, and X_{14}, respectively.

4 Discussion

4.1 Effect of optimized plant spatial competition on wheat tillering

Plants primarily compete for resources such as light, water, and nutrients (Sial et al., 2010). To fully use resources, plants must generate appropriate plant and leaf types to adapt to competition for living space (Evers et al., 2006; Dreccer et al., 2012; Fioreze and Rodrigues, 2014; Wang et al., 2018). However, as an essential aspect of the participation of gramineous crops in resource competition, the number of tillers and the start and end times of their formation are significantly affected by the distance between adjacent plants (Liu et al., 2011; Alzueta et al., 2012; Clerget et al., 2016; Lecarpentier et al., 2019). If the distribution of plants is not sufficiently uniform, it will lead to competition in the utilization of soil nutrients in areas with dense plant distribution, which will further lead to excessive growth. Further optimization of the allocation of plant row spacing to optimize crop spatial competition under conditions of equal density has attracted increasing research attention (Ballaré, 1999; Golubski et al., 2008; Liu et al., 2021c; Xiao et al., 2022).

Alzueta et al. (2012) reported that changing row spacing affects interactions between neighboring plants; such interactions involve competition for resources and active morphogenetic responses triggered by neighbors' perceptions. De Vita et al. (2017) and Fischer et al. (2019) reported that changing the density by increasing row spacing has a stronger impact than simply changing the plant density in rows because it leads to increased spatial heterogeneity of crops. In another study, plant spacing was reduced under wide row spacing, and the shading reaction of wheat plants cultivated at high density appeared at the early tillering stage. This study also suggested that wheat plants respond to changes in the distance between rows by altering their tillers, leaf size, and angle (de Wit et al., 2012).

Row spacing and density currently used in wheat cultivation do not lead to competition between plant rows. For example, at a conventional sowing date, density, and row spacing in the NCP, the canopy of winter wheat is not closed in the pre-winter stage, and thus there is almost no competition between rows in the pre-winter stage. However, competition among plants exists objectively. In the tillering stage, wheat leaves contact each other successively, leading to spatial competition in signal response plants and further phenotypic changes. For example, Hussain et al. (2012) reported that decreasing wheat row spacing from 30 to 15 cm increased the tiller number by 79.3%. Abichou et al. also reported that wheat row spacing was reduced from 35 cm to 17.5 cm without changing the density, and the plant spacing was correspondingly expanded, increasing the number of tillers of different varieties by 12%–19%. Other studies have shown similar trends for rice (Clerget et al., 2016) and *Aegilops tauschii* Coss (Yu et al., 2020).

We reduced the row spacing of winter wheat in the NCP area from 15 cm to 7.5 cm. The tillering and biomass of single plants before winter increased by 14.5% and 19.1%, respectively, demonstrating the feasibility of reducing row spacing and expanding plant spacing to regulate tillering in wheat production. Moreover, after postponing the sowing date, reducing row spacing promoted an increase in the number of tillers and wheat biomass. This result is applicable to tillering control in late-sown winter wheat by controlling various factors. The detailed mechanism of the above phenomena was revealed from the perspective of phytohormones (see below). Carbohydrates, plant temperature, and radiation energy utilization will be reported in subsequent papers.

4.2 Effects of optimized plant spatial competition on phytohormones content in wheat

A variety of phytohormones are single or synergistic and play important roles in the growth and development of crop branches and tillers (Müller and Leyser, 2011; Cai et al., 2018; Wang et al., 2018). Auxin, GA, and IAA inhibit plant tillering to different degrees, whereas CTK and endogenous ZT promote plant tillering. Additionally, the growth of wheat tiller buds is regulated by the concentrations of IAA and ZT in the tillering node and the ratios of IAA to ZT and abscisic acid to ZT (Shimizu-Sato et al., 2008; Cai et al., 2018). SLs regulate the tillering of gramineous crops, possibly due to the interaction between SLs and auxin (Liang et al., 2010). SLs and auxin may regulate tillering, as auxin regulates branching by inducing SL synthesis, or SLs may systematically regulate auxin transport (Tanaka, 2006; Prusinkiewicz et al., 2009).

The factors influencing phytohormones and environmental conditions during plant tillering are complex. The effects of phytohormone type and content, plant sowing density, plant spacing configuration, and other factors on tillering have been widely examined (Evers et al., 2006; Müller and Leyser, 2011; Hussain et al., 2012; Wang et al., 2018; Kaplan and Specht, 2022). However, there is a lack of direct evidence for the role of hormones and whether competition between plants directly affects hormones. For example, hormones are internal factors that link plant density to tillering development in *A. tauschii* (a gramineous crop), but the specific roles of these hormones remain unknown. We showed that the IAA and GA contents of 7.5RS for the three sowing-date groups decreased significantly, whereas the ZR and SL contents increased significantly compared to those of 15RS. Previous research on gramineous crops, such as *A. tauschii* Coss (Yu et al., 2020) and *Miscanthus* × *giganteus* (Schmidt et al., 2017), showed that when the density increased, the IAA and GA contents increased, whereas the content of CTK, a growth hormone, showed the opposite trend. Although the above studies and studies on crops differ from the present study, the results support those of our study.

Multiple hormones affect competition for space between plants, including plant spacing and density, and regulate tiller growth (Yu et al., 2020). For example, GA inhibits tillering in tall fescue through crosstalk with CTK (Zhuang et al., 2019). Another study showed

that plant hormone crosstalk, in which SL downregulated CTK levels by inducing the expression of CYTOKININ OXIDASE/DEHYDROGENASE 9, was involved in tillering regulation in rice (Duan et al., 2019). Additionally, natural auxin IAA stimulates SL biosynthesis, thus providing a potential mechanism for inhibiting bud outgrowth (Rameau et al., 2015).

We found that the number of tillers and biomass were significantly and positively correlated with SLs/GA for the three sowing-date groups ($P<0.05$), indicating that the environmental conditions changed when row spacing was reduced, which may inhibit the top development of wheat, promote the proportion of branching development hormones, and increase tillering. This study improves our understanding of the influence of plant spacing on endogenous hormones and their regulatory effects on tillering. However, we only evaluated the tillering node, and further in-depth analyses of the hormone characteristics of the root, other parts of the stem, and leaves are required. In addition, some hormones may not directly act on tillering but act as primary messengers, regulating tillering by changing the content of other hormones or the balance of several phytohormones (Cline and Oh, 2006), which requires further evaluation.

4.3 Effect of optimizing plant spatial competition on grain yield factors of wheat

Appropriate row spacing is important for improving crop productivity because plants growing in too wide rows may not efficiently utilize light, water, and nutrient resources, whereas growth in narrow rows may result in inter-row solid competition (Ali et al., 1999). Therefore, adjusting the plant and row spacing of crops and alleviating competition for plant growth space is important for cultivating crops with high yields, particularly closely planted crops with tillering characteristics (Abichou et al., 2019; Fischer et al., 2019). The yield of wheat with reduced row spacing (15 and 5 cm) was significantly higher than that at 25 cm (De Vita et al., 2017). Hussain et al. reported that wheat yield increased when row spacing was reduced, but there were large differences among varieties. Compared with the 30 cm and 25 cm row spacings, the yield of medium-plant height varieties was the highest at 20 cm, and that of short-plant height varieties was the highest at 15 cm, showing that the GW of wheat decreased after row spacing was reduced. This result corresponds to an increased number of spikes or grains per spike, thus increasing grain yield (Shimizu-Sato et al., 2008; Hussain et al., 2012).

Some studies have also pointed out that there is no evidence that increasing planting density can offset the broader row yield reduction (Fischer et al., 2019). Additionally, many studies have shown that reducing row spacing improves yield. However, when row spacing is reduced to a certain distance, the distribution of plant populations and individuals reaches a balance. If row spacing is further decreased, yield is not affected or even decreased (Liu et al., 2016; De Vita et al., 2017). However, the mechanism by which reducing row spacing can increase yield is complex. Decreasing row spacing can reduce the rectangularity of seed distribution (as the ratio of inter-row to intra-row distance), enabling greater uptake of soil nutrients and increasing ground cover by the crop (De Vita et al., 2017). Additionally, changing the density by increasing the inter-row distance has a stronger effect than changing the density of plants within a row, as it leads to increased crop spatial heterogeneity (Fischer et al., 2019). Furthermore, plants with a uniform spatial distribution show better tolerance to high-density planting, and the population and individuals receive more coordinated light, which is an important factor in reducing row spacing to improve yield (Liu et al., 2016).

We showed that reducing row spacing could significantly increase the number of pre-winters and total tillers. Row spacing, Y × SD interactions, and other sources of variation also had significant effects. The average yield of the three sowing-date groups increased by 6.3% when the row spacing was decreased from 15 cm to 7.5 cm (average of two years). The yield increase of the LSRE treatment was mainly reflected in the proportional increase of three yield factors, particularly the interaction between the grain number per spike and GW, which is different from the results of previous studies. According to path analysis (Figure 12), the number of grains per spike in 15RS had the most significant influence on yield, whereas 7.5RS had the greatest impact on GW. The contribution of GW to the 7.5RS yield increase increased when the sowing date was delayed. In addition, the pre-winter tillers of the three sowing-date groups in the 7.5RS treatment were included in the regression model, with the results supporting the importance of tillering in the LSRE treatment.

The tiller number was closely related to the changes in the contents of various phytohormone contents under LSRE treatments, specifically lower IAA and GA contents, higher ZR and SL contents, and higher SLs/GA (Figure 11). Although the phytohormone content in the pre-winter stage cannot directly control grain yield, the change in phytohormones significantly affects tillers, spikes the number of wheat, and further controls yield. Therefore, it is very necessary to study the change in hormone content as an important physiological mechanism of LSRE treatment to control wheat grain yield. In particular, the change in phytohormone content in the middle of the tillering stage of pre-winter has a more significant effect on the regulation of wheat tillering. It is necessary to continue studying.

5 Conclusion

Compared with the 15RS sowing date treatment, 7.5RS alleviated plant space competition, inhibited the accumulation of IAA and GA at the tillering nodes of wheat, increased the content of ZR and SLs, changed the IAA/ZR and SLs/GA, promoted tillering before winter, increased biomass accumulation, and delayed the sowing date more significantly. The increase in tillering and GW before winter was the main factor leading to increased yield in the LSER treatment. LSER can alter the phytohormone content, promote tiller growth, and increase the GW and grain yield of

wheat in the NCP, particularly in the heat resource shortage area in the northern part of the NCP.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

Author contributions

WZ and MY conceived the project and set scientific objectives. BY, JG, LG and PL contributed to the preparation of the field experiment and data acquisition. BY, XL, and PL wrote the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research was funded by the National Key Research and Development Project (Grant No. 2017YFD0300900), the Key R&D projects in Hebei Province (Grant No. 21327001D), and the Modern Agricultural Industrial Technology System in Hebei Province (Grant No. HBCT2018010205).

Acknowledgments

We would like to thank Editage (www.editage.cn) for the English language editing.

Conflict of interest

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

Publisher's note

References

Abichou, M., de Solan, B., and Andrieu, B. (2019). Architectural response of wheat cultivars to row spacing reveals altered perception of plant density. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.00999

Ali, Y., Haq, M. A., Tahir, G. R., and Ahmad, N. (1999). Effect of inter and intra row spacing on the yield and yield components of chickpea. *Pakistan J. @ Biol. Sci.* 2, 305–307. doi: 10.3923/pjbs.1999.305.307

Alzueta, I., Abeledo, L. G., Mignone, C. M., and Miralles, D. J. (2012). Differences between wheat and barley in leaf and tillering coordination under contrasting nitrogen and sulfur conditions. *Eur. J. Agron.* 41, 92–102. doi: 10.1016/j.eja.2012.04.002

Ballaré, C. L. (1999). Keeping up with the neighbours: Phytochrome sensing and other signalling mechanisms. *Trends Plant Sci.* 4, 97–102. doi: 10.1016/s1360-1385(99)01383-7

Barbier, F. F., Dun, E. A., Kerr, S. C., Chabikwa, T. G., and Beveridge, C. A. (2019). An update on the signals controlling shoot branching. *Trends Plant Sci.* 24, 220–236. doi: 10.1016/j.tplants.2018.12.001

Cai, T., Meng, X., Liu, X., Liu, T., Wang, H., Jia, Z., et al. (2018). Exogenous hormonal application regulates the occurrence of wheat tillers by changing endogenous hormones. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.01886

Clerget, B., Bueno, C., Domingo, A. J., Layaoen, H. L., and Vial, L. (2016). Leaf emergence, tillering, plant growth, and yield in response to plant density in a high-yielding aerobic rice crop. *Field Crops Res.* 199, 52–64. doi: 10.1016/j.fcr.2016.09.018

Cline, M. G., and Oh, C. (2006). A reappraisal of the role of abscisic acid and its interaction with auxin in apical dominance. *Ann. Bot.* 98, 891–897. doi: 10.1093/aob/mcl173

Cornelius, C., Petermeier, H., Estrella, N., and Menzel, A. (2011). A comparison of methods to estimate seasonal phenological development from BBCH scale recording. *Int. J. Biometeorol* 55, 867–877. doi: 10.1007/s00484-011-0421-x

De Vita, P., Colecchia, S. A., Pecorella, I., and Saia, S. (2017). Reduced inter-row distance improves yield and competition against weeds in a semi-dwarf durum wheat variety. *Eur. J. Agron.* 85, 69–77. doi: 10.1016/j.eja.2017.02.003

de Wit, M., Kegge, W., Evers, J. B., Vergeer-van Eijk, M. H., Gankema, P., Voesenek, L. A. C. J., et al. (2012). Plant neighbor detection through touching leaf tips precedes phytochrome signals. *Proc. Natl. Acad. Sci.* 109, 14705–14710. doi: 10.1073/pnas.1205437109

Dreccer, M. F., Chapman, S. C., Rattey, A. R., Neal, J., Song, Y., Christopher, J.(. T., et al. (2012). Developmental and growth controls of tillering and water-soluble carbohydrate accumulation in contrasting wheat (Triticum aestivumL.) genotypes: can we dissect them? *EXBOTJ* 64, 143–160. doi: 10.1093/jxb/ers317

Duan, J., Yu, H., Yuan, K., Liao, Z., Meng, X., Jing, Y., et al. (2019). Strigolactone promotes cytokinin degradation through transcriptional activation of CYTOKININ OXIDASE/DEHYDROGENASE 9in rice. *Proc. Natl. Acad. Sci. U.S.A.* 116, 14319–14324. doi: 10.1073/pnas.1810980116

Evers, J. B., Vos, J., Andrieu, B., and Struik, P. C. (2006). Cessation of tillering in spring wheat in relation to light interception and Red: Far-red ratio. *Ann. Bot.* 97, 649–658. doi: 10.1093/aob/mcl020

Feng, S., Ding, W., Shi, C., Zhu, X., Hu, T., and Ru, Z. (2023). Optimizing the spatial distribution of roots by supplemental irrigation to improve grain yield and water use efficiency of wheat in the north China plain. *Agric. Water Manage.* 275, 107989. doi: 10.1016/j.agwat.2022.107989

Fioreze, S. L., and Rodrigues, J. D. (2014). Tillering affected by sowing density and growth regulators in wheat. *Sem. Ci. Agr.* 35, 589. doi: 10.5433/1679-0359.2014v35n2p589

Fischer, R. A., Moreno Ramos, O. H., Ortiz Monasterio, I., and Sayre, K. D. (2019). Yield response to plant density, row spacing and raised beds in low latitude spring wheat with ample soil resources: An update. *Field Crops Res.* 232, 95–105. doi: 10.1016/j.fcr.2018.12.011

Golubski, A. J., Gross, K. L., and Mittelbach, G. G. (2008). Competition among plant species that interact with their environment at different spatial scales. *Proc. R. Soc B* 275, 1897–1906. doi: 10.1098/rspb.2008.0272

Hussain, M. Z., Mehmood, M. B., Khan, S. F., Lee a, D. J., and Farooq, M. (2012). Narrow row spacing ensures higher productivity of low tillering wheat cultivars. *Int. J. Agric. Biol.* 14, 413–418. doi: 10.1080/03650340.2012.725937

Kaplan, D. R., and Specht, C. D. (2022). *Shoot branching* (Boca Raton: CRC Press), 279–307. doi: 10.1201/9781315118642-12

Lafond, G. P. (1994). Effects of row spacing, seeding rate and nitrogen on yield of barley and wheat under zero-till management. *Can. J. Plant Sci.* 74, 703–711. doi: 10.4141/cjps94-127

Lecarpentier, C., Barillot, R., Blanc, E., Abichou, M., Goldringer, I., Barbillon, P., et al. (2019). WALTer: A three-dimensional wheat model to study competition for light through the prediction of tillering dynamics. *Ann. Bot.* 123, 961–975. doi: 10.1093/aob/mcy226

Liang, J., Zhao, L., Challis, R., and Leyser, O. (2010). Strigolactone regulation of shoot branching in chrysanthemum (Dendranthema grandiflorum). *J. Exp. Bot.* 61, 3069–3078. doi: 10.1093/jxb/erq133

Liu, G., Zhao, J., Liao, T., Wang, Y., Guo, L., Yao, Y., et al. (2021a). Histological dissection of cutting-inducible adventitious rooting in platycladus orientalis reveals developmental endogenous hormonal homeostasis. *Ind. Crops Products* 170, 113817. doi: 10.1016/j.indcrop.2021.113817

Liu, R., Hou, J., Li, H., Xu, P., Zhang, Z., and Zhang, X. (2021b). Association of TaD14-4D, a gene involved in strigolactone signaling, with yield contributing traits in wheat. *IJMS* 22, 3748. doi: 10.3390/ijms22073748

Liu, T., Wang, Z., and Cai, T. (2016). Canopy apparent photosynthetic characteristics and yield of two spike-type wheat cultivars in response to row spacing under high plant density. *PloS One* 11, e0148582. doi: 10.1371/journal.pone.0148582

Liu, X., Wang, X., Liu, P., Bao, X., Hou, X., Yang, M., et al. (2022). Rehydration compensation of winter wheat is mediated by hormone metabolism and de-peroxidative activities under field conditions. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.823846

Liu, Y., Ding, Y.-F., Wang, Q.-S., LI, G.-H., Xu, J.-X., Liu, Z.-H., et al. (2011). Effect of plant growth regulators on growth of rice tiller bud and changes of endogenous hormones. *Acta Agronomica Sin.* 37, 670–676. doi: 10.1016/s1875-2780(11)60019-9

Liu, Y., Liao, Y., and Liu, W. (2021c). High nitrogen application rate and planting density reduce wheat grain yield by reducing filling rate of inferior grain in middle spikelets. *Crop J.* 9, 412–426. doi: 10.1016/j.cj.2020.06.013

Mao, Y., Chai, X., Zhong, M., Zhang, L., Zhao, P., Kang, Y., et al. (2022). Effects of nitrogen and magnesium nutrient on the plant growth, quality, photosynthetic characteristics, antioxidant metabolism, and endogenous hormone of Chinese kale (Brassica albograbra bailey). *Scientia Hortic.* 303, 111243. doi: 10.1016/j.scienta.2022.111243

McSteen, P., and Leyser, O. (2005). "Shoot Branching. Annual Review of Plant Biology. 56, 121–128. doi: 10.1146/annurev.arplant.56.032604.144122

Müller, D., and Leyser, O. (2011). Auxin, cytokinin and the control of shoot branching. *Ann. Bot.* 107, 1203–1212. doi: 10.1093/aob/mcr069

Ozturk, A., Caglar, O., and Bulut, S. (2006). Growth and yield response of facultative wheat to winter sowing, freezing sowing and spring sowing at different seeding rates. *J. Agron. Crop Sci.* 192, 10–16. doi: 10.1111/j.1439-037X.2006.00187.x

Prusinkiewicz, P., Crawford, S., Smith, R. S., Ljung, K., Bennett, T., Ongaro, V., et al. (2009). Control of bud activation by an auxin transport switch. *Proc. Natl. Acad. Sci.* 106, 17431–17436. doi: 10.1073/pnas.0906696106

Rameau, C., Bertheloot, J., Leduc, N., Andrieu, B., Foucher, F., and Sakr, S. (2015). Multiple pathways regulate shoot branching. *Front. Plant Sci.* 5. doi: 10.3389/fpls.2014.00741

Schmidt, C. S., Mrnka, L., Frantík, T., Motyka, V., Dobrev, P. I., and Vosátka, M. (2017). Combined effects of fungal inoculants and the cytokinin-like growth regulator thidiazuron on growth, phytohormone contents and endophytic root fungi in miscanthus × giganteus. *Plant Physiol. Biochem.* 120, 120–131. doi: 10.1016/j.plaphy.2017.09.016

Shang, Q., Wang, Y., Tang, H., Sui, N., Zhang, X., and Wang, F. (2021). Genetic, hormonal, and environmental control of tillering in wheat. *Crop J.* 9, 986–991. doi: 10.1016/j.cj.2021.03.002

Shimizu-Sato, S., Tanaka, M., and Mori, H. (2008). Auxin–cytokinin interactions in the control of shoot branching. *Plant Mol. Biol.* 69, 429–435. doi: 10.1007/s11103-008-9416-3

Sial, M. A., Arain, M. A., Dahot, M. U., Markhand, G. S., Laghari, K. A., Mangrio, S., et al. (2010). Effect of sowing dates on yield and yield components on mutant-cum-hybrid lines of bread wheat. *Pak. J. Bot.* 42, 269–277. doi: 10.1127/1438-9134/2010/0136-0325

Tanaka, Y. (2006). Cytokinin and auxin inhibit abscisic acid-induced stomatal closure by enhancing ethylene production in arabidopsis. *J. Exp. Bot.* 57, 2259–2266. doi: 10.1093/jxb/erj193

Tilley, M. S., Heiniger, R. W., and Crozier, C. R. (2019). Tiller initiation and its effects on yield and yield components in winter wheat. *Agron. J.* 111, 1323–1332. doi: 10.2134/agronj2018.07.0469

Wang, Y., Miao, F., and Yan, L. (2016). Branching shoots and spikes from lateral meristems in bread wheat. *PloS One* 11, e0151656. doi: 10.1371/journal.pone.0151656

Wang, B., Smith, S. M., and Li, J. (2018). Genetic regulation of shoot architecture. *Annu. Rev. Plant Biol.* 69, 437–468. doi: 10.1146/annurev-arplant-042817-040422

Xia, H., Qiao, Y., Li, X., Xue, Y., Wang, N., Yan, W., et al. (2023). Moderation of nitrogen input and integration of legumes *via* intercropping enable sustainable intensification of wheat-maize double cropping in the north China plain: A four-year rotation study. *Agric. Syst.* 204, 103540. doi: 10.1016/j.agsy.2022.103540

Xiao, Z., Gu, H., Wu, H., Jing, W., Zhu, K., Zhang, W., et al. (2022). Effects of planting density, levels, and forms of nitrogen application on the yield and nitrogen utilization of wheat following rice in East China. *Agronomy* 12, 2607. doi: 10.3390/agronomy12112607

Xu, T., Bian, N., Wen, M., Xiao, J., Yuan, C., Cao, A., et al. (2016). Characterization of a common wheat (Triticum aestivum l.) high-tillering dwarf mutant. *Theor. Appl. Genet.* 130, 483–494. doi: 10.1007/s00122-016-2828-6

Yang, D. Q., Luo, Y. L., Dong, W. H., Yin, Y. P., Li, Y., and Wang, Z. L. (2018). Response of photosystem II performance and antioxidant enzyme activities in stay-green wheat to cytokinin. *Photosynt* 56, 567–577. doi: 10.1007/s11099-017-0708-1

Yu, H., Yang, J., Cui, H., Li, Z., Jia, F., Chen, J., et al. (2020). Effects of plant density on tillering in the weed grass aegilops tauschii coss. and its phytohormonal regulation. *Plant Physiol. Biochem.* 157, 70–78. doi: 10.1016/j.plaphy.2020.10.013

Zhuang, L., Ge, Y., Wang, J., Yu, J., Yang, Z., and Huang, B. (2019). Gibberellic acid inhibition of tillering in tall fescue involving crosstalks with cytokinins and transcriptional regulation of genes controlling axillary bud outgrowth. *Plant Sci.* 287, 110168. doi: 10.1016/j.plantsci.2019.110168

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 17 March 2023
DOI 10.3389/fpls.2023.1142814

OPEN ACCESS

EDITED BY
Meng Kou,
Sweet Potato Research Institute (CAAS), China

REVIEWED BY
Wenliang Xu,
Central China Normal University, China
Shuangxia Jin,
Huazhong Agricultural University, China

*CORRESPONDENCE
Chuangyun Wang
wrwcy@139.com
Dengxiang Du
ddx@whpu.edu.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Crop and Product Physiology, a section of the journal Frontiers in Plant Science

RECEIVED 12 January 2023
ACCEPTED 16 February 2023
PUBLISHED 17 March 2023

CITATION
Huo D, Hao Y, Zou J, Qin L, Wang C and Du D (2023) Integrated transcriptome and metabonomic analysis of key metabolic pathways in response to cadmium stress in novel buckwheat and cultivated species. *Front. Plant Sci.* 14:1142814. doi: 10.3389/fpls.2023.1142814

Integrated transcriptome and metabonomic analysis of key metabolic pathways in response to cadmium stress in novel buckwheat and cultivated species

Dongao Huo[1,2], Ying Hao[1], Juan Zou[3], Lixia Qin[4], Chuangyun Wang[4*†] and Dengxiang Du[3*†]

[1]Guizhou Normal University, Guiyang, China, [2]College of Biological Sciences and Technology, Taiyuan Normal University, Taiyuan, China, [3]School of Life Science and Technology, Wuhan Polytechnic University, Wuhan, China, [4]College of Agriculture, Shanxi Agricultural University, Taiyuan, China

Introduction: Buckwheat (*Fagopyrum tataricum*), an important food crop, also has medicinal uses. It is widely planted in Southwest China, overlapping with planting areas remarkably polluted by cadmium (Cd). Therefore, it is of great significance to study the response mechanism of buckwheat under Cd stress and further develop varieties with excellent Cd tolerance.

Methods: In this study, two critical periods of Cd stress treatment (days 7 and 14 after Cd treatment) of cultivated buckwheat (Pinku-1, named K33) and perennial species (*F. tatari-cymosum* Q.F. Chen) (duoku, named DK19) were analyzed using transcriptome and metabolomics.

Results: The results showed that Cd stress led to changes in reactive oxygen species (ROS) and the chlorophyll system. Moreover, Cd-response genes related to stress response, amino acid metabolism, and ROS scavenging were enriched or activated in DK19. Transcriptome and metabolomic analyses highlighted the important role of galactose, lipid (glycerophosphatide metabolism and glycerophosphatide metabolism), and glutathione metabolism in response to Cd stress in buckwheat, which are significantly enriched at the gene and metabolic levels in DK19.

Discussion: The results of the present study provide valuable information for a better understanding of the molecular mechanisms underlying Cd tolerance in buckwheat and useful clues for the genetic improvement of drought tolerance in buckwheat.

KEYWORDS

buckwheat, cadmium (Cd) stress, metabolome, transcriptome, reactive oxygen species (ROS)

1 Introduction

1.1 Cadmium pollution and influence

Heavy metal pollution in soil is a global problem, which has posed a serious threat to public health. It is reported that in China, the farmland polluted by heavy metals cadmium, lead, arsenic and mercury covers an area of 20 million hm^2, accounting for about 20% of the total farmland area, of which 14000 hm^2 of farmland soil is seriously polluted by cadmium (Jiang et al., 2019; Lin et al., 2021). Cd in soil will have adverse effects on crops, agricultural products and groundwater (Qin et al., 2021). Cadmium (Cd) is a non-essential trace heavy metal element for plant growth (Hoseini and Zargari, 2013; Bhattacharya, 2022), it will accumulate in all parts of plants after being absorbed. The toxic effect of Cd on plants is firstly shown as inhibition of root growth, which causes root browning, reduction of lateral root number and root tip death. Secondly, Cd will cause leaf curl, terminal leaf chlorosis and stem growth retardation (Clemens, 2006; Ran et al., 2015; Li et al., 2021a; Meng et al., 2022). At the same time, Cd will replace the essential elements of the active center of enzymes involved in plant biochemical reactions, replace the sulfhydryl groups of proteins and enzymes, change the conformation of macromolecules, cause protein denaturation, and damage the cell membrane, thus affecting the growth and development of plants (Romero-Puertas et al., 2002; Riaz et al., 2021). More dangerous is that Cd accumulated in plants will enter the human body through the food chain, which will cause serious damage to the body (Alengebawy et al., 2021; Rehman et al., 2021).

1.2 Detoxification mechanism of plants under cadmium stress

When plants grow in soil rich in Cd, their growth and development are affected comprehensively, and produces a series of tolerance and detoxification mechanisms to control the absorption, transportation and accumulation of Cd to alleviate the toxicity (Zhang et al., 2021b). Plant root epidermis, trichome and cuticle all produce protective tissue to isolate Cd (Feng et al., 2021), and secrete organic acids (citric acid, lactic acid, malic acid) to form a complex with cadmium to inhibit the transmembrane transport of cadmium, and change the pH and EH of the rhizosphere to reduce the migration of soil cadmium into the root (Chen et al., 2019a). Polysaccharides and proteins in the cell wall can bind with Cd to prevent Cd from entering the cell, Cd entering cells will be enriched in vacuoles to reduce the toxicity of cadmium to cell activity (Wei et al., 2021). Chelators such as glutathione, plant chelating peptide and metallothionein in plant cells can bind to Cd to avoid Cd binding to proteins with important physiological functions (Baghayeri et al., 2018).

1.3 Application of transcriptomics in the study of cadmium stress in plants

Transcriptome connects gene expression and biological function, and shows a strong advantage in identifying key genes that respond to abiotic stresses during plant growth and development (Tang et al., 2011). Many transcriptome analyses have been conducted to analyze the response of plants to abiotic stress (Dossa et al., 2019; Zhang et al., 2019; Kang et al., 2020), and studies on the response to Cd stress have been carried out in rape, Arabidopsis, maize, rice, and other crops (Rui et al., 2018; Shu et al., 2019). Using transcriptome technology to analyze ceruleuca glauca with different genotypes under Cd stress, found that *NcNramp1* played a key role in the process of plant absorption and accumulation of Cd^{2+} (Milner et al., 2014). The results of transcriptome analysis of winter wheat under cadmium stress deteced 429 up-regulated genes and 998 down-regulated genes, and found that these differential genes were involved in the defense and detoxification mechanism under Cd stress (Xiao et al., 2019). Similarly, sequencing analysis of maize under Cd stress showed that auxin affected the Cd content of maize seedlings, and may improve maize yield and Cd tolerance by regulating auxin signals (Yue et al., 2016).

1.4 Application of metabonomics in the study of cadmium stress in plants

Metabolomics is a technology to study the metabolic network of biological systems (Weckwerth, 2003), using metabolomics to explore the response mechanism and metabolic network regulation of plants under abiotic stress is of great significance (Sung et al., 2015; Feng et al., 2020; Nephali et al., 2020; Kumar et al., 2021). After Cd treatment of amaranth, 41 different metabolites were identified, of which 12 differential metabolites were significantly related to three pathways (the synthesis of valine, leucine and isoleucine) (Mei et al., 2014). The metabonomic analysis of Brassica napus by LC-MS proved that the production of reducing agents (monoterpenoids, carotenoids, etc.) plays an important role in promoting the detoxification mechanism and stabilizing the membrane (Mwamba et al., 2020; Wang et al., 2021). In addition, the metabonomic analysis of indica rice grains under Cd stress was carried out by mass spectrometry metabonomic, the results showed that the metabolism of carbohydrates, organic acids and amino acids was affected by Cd stress, and the cysteine content increased, which could improve the tolerance to Cd toxicity (Zeng et al., 2021). Yang et al. (2019) performed a metabolomic analysis of alfalfa under Cd stress to determine the response of proline and free amino acid content to improve its resistance to Cd toxicity.

1.5 Response of buckwheat to cadmium stress

Buckwheat, as an important minor cereal, has important nutritional value. However, as a major crop in many remote areas, it has a potential risk of Cd exceeding the standard (Zhang et al., 2017). Therefore, it is of great significance to study the molecular mechanism of buckwheat response to Cd stress. The research on buckwheat under Cd stress mainly focused on physiological characteristics such as seedling growth, subcellular

distribution and the transfer coefficient to plants (Mitrus and Horbowicz, 2020; Li et al., 2021b). There are few reports on the use of multiomics integration analyses to study the response law of buckwheat to Cd stress in a deep level, or even to find the core regulatory network sunder it. Therefore, this study uses transcriptome sequencing technology and non-targeted metabolomics technology to integrate and analyze the response mechanism of buckwheat to Cd stress, reveal the key metabolic pathways under it, and provide a new idea for a comprehensive understanding of buckwheat Cd stress response mechanism.

2 Materials and methods

2.1 Plant material

The samples used in this study were obtained from College of Biological Sciences and Technology, Taiyuan Normal University (Taiyuan, Shanxi, China). The cultivated species named Kuqiao 33 (K33) is an excellent new line developed from Qinghai landrace black Tartary Buckwheat by N ion beam implantation. The field performance is good, the maturity period is consistent, and the lodging resistance and seed falling resistance are good (Zhang et al., 2017).

The novel buckwheat named duoku19 (DK19) is a new perennial species (*F. tatari-cymosum* Q.F. Chen) by crossing annual auto-tetraploid tartary buckwheat and perennial tetraploid *F. cymosum*, using the main agronomic traits (plant height, yield, grain color, etc.) and key molecular markers as screening criteria (Sytar et al., 2014). In a traditional agricultural production model, the establishment of annual crops has both economic and agronomic implications such as high seed and nutrient inputs, ploughing, and may involve a number of sowings each year (Komori et al., 2007).

2.2 Growth conditions and cadmium treatment

The experiment was carried out in the greenhouse of Taiyuan Normal University with planting method. Chose buckwheat seeds of similar size sterilized them with 3% NaClO for 10 minutes, then washed them with distilled water for 3 times, and put the seeds on the wet filter paper in the germination box for germination. The culture conditions were: illumination at 25°C for 16 hours, darkness at 25°C for 8 hours, humidity 80%, and 400 $\mu molm^{-2}$ s^{-1} intense luminosity. The seedlings were transplanted into a polyethylene pot (contain Hoagland nutrient solution) at the cotyledons and radicle reach 3-5 cm, 12 holes in each pot and 1 plant in each hole for each cultivated variety in the greenhouse. The greenhouse cultivation conditions were 16 hours of light at 25-30°C and 8 hours of dark light at 15-18°C. The air pump supplemented the nutrient solution with oxygen to provide the oxygen content required for root respiration, and the nutrient solution was replaced every 3 days to ensure the consistency of nutrient content in all treatments and before and after each treatment (Zhang et al., 2021a).

Buckwheat with consistent growth was selected and its roots were placed in deionized water for 12 h to remove excess nutrient ions on the root surface. After that, the buckwheat root system was rinsed with deionized water for times, gently wiped dry and immersed in Cd solution (Cd solution, 1 mg/kg) for 2 h under magnetic stirring. Each material was treated with 200 plants. After the adsorption was completed, the buckwheat roots were taken out from the Cd solution, washed with deionized water, and the excess water on the root surface was gently wiped with filter paper. In this study, DK19 and K33 were divided into 6 groups for analysis. The untreated materials were used as the control group (DK19-0 and K33-0), DK19-1 and K33-1 were treated for 7 days, and the materials treated for 14 days were named K19-2 and K33-2.

2.3 Determination of chlorophyll content and antioxidant enzyme activity

Phenotype of leaves and chlorophyll content were measured on the 7th and 14th days after treatment, and fresh leaf samples were taken for transcriptome analysis and metabolome analysis. Ten samples were collected at different stages after Cd treatment, and the un-treated samples as the control, the samples were quickly stored in liquid nitrogen. The buckwheat leaf cells were resuspended with 80% acetone solution, incubated for 48 hours, and vibrated at 25 ° at 80 rpm speed in dark culture. Measure the absorbance of the supernatant at 652 nm to calculate the chlorophyll content (Pérez-Patricio et al., 2018).

Take 0.5 g of fresh leaves, put the samples (three replicates) stored at - 80°C into liquid nitrogen precooled mortar, and grind them to powder. Put into a 10 ml centrifuge tube. Add 8 ml 0.05 $molL^{-1}$ sodium phosphate buffer (pH = 7.8) precooled in advance, centrifuge at 10000 rpm for 20 min, and store the supernatant at 4°C for enzyme activity determination (He et al., 2013). The content of malondialdehyde (MDA) was determined by thiobarbituric acid TBA. Peroxidase (POD) activity was determined by guaiacol method (Peter-Katalinić, 2005). The activity of catalase (CAT) was determined by TBA-TCA method (Stewart and Bewley, 1980). Each physiological and biochemical index of each treatment was measured three times, and the results were expressed as the average value.

2.4 Untargeted metabolomics of buckwheat leaf tissues

Untargeted metabolomics, a useful approach for the simultaneous analysis of samples with significant phenotypic changes, has been used to detect metabolites after Cd treatment (Gevi et al., 2019). The samples with significant phenotypic changes were used to detect the metabolites after drought treatment. Three experimental replicates were carried out in each treatment stage, and ten materials with consistent growth were used for each replicate. Samples of each group were fully crushed in liquid nitrogen, extracted with 10g leaf samples extract with countryman filter paper. After lyophilization for 7 days, the metabolite samples

were ground to powder in a ball mill (MM400, Retsch, Germany) at 30 Hz for 45 s. Then, 0.05–0.1 g powdered sample was extracted as previously description (Ramalingam et al., 2021).

Non-targeted metabolic profiling analyses were profiled by Perkin Elmer 680 GC (Perkin Elmer Inc, Akron, OH, USA) and Q Exactive Focus Orbitrap LC-MS/MS (Thermo Scientific, Waltham, MA, USA). Scanning mass ranged from m/z 100–1000 with an accumulation time of 0.10s. The scanning mode was full MS/ddMS2. The recorded data were processed with compound discoverer (CD) 3.1 software to obtain the mass to charge ratio, retention time, MS/MS2 information of all detected substances. Then, the detected signals were automatically matched through the internally established reference libraries of chemical standard entries of software to predict and identify the metabolite information. The multiple reaction monitoring (MRM) mode with QTRAP 6500+ LC-MS/MS (Shimadzu, Kyoto, Japan) was used for targeted metabolome analyses. The detection window was set to 80 s, and the targeted scanning time was 1.5 s. The original data were processed by Multi Quant 3.0.3 software. The chromatographic column was C18 column (Shim-pack GLSS C18, 1.9UM, 2.1*100, Shimadzu). Mobile phase A and B was 0.04% acetic acid–water solution, and mobile phase B was 0.04% acetic acid–methanol solution. The qualitative and quantitative chromatographic conditions were consistent.

2.5 Bioinformatics analysis of metabolome data

Normalize the total peak area by dividing each metabolite in the sample by the total peak area of the sample, so as to normalize the data before analysis. Principal component analysis (PCA), correlation analysis between samples, Orthogonal Partial Least Squares-Discriminant Analysis (OPLS-DA) are carried out by using R software package ropls to classify and discriminate between samples (http://bioconductor.org/packages/release/bioc/html/ropls.htm) (Thevenot, 2016). We combined the multivariate statistical analysis of the VIP value of OPLS-DA and the univariate statistical analysis of the T-test P-value to screen differentially accumulated metabolites (DAMs) among different comparison groups (Wu et al., 2022). Screening criteria FC>1.50, P-value<0.05 and VIP>1 were considered to determine the significant differences between samples. The Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis was performed on the dam using KOBAS software (Mao et al., 2005).

2.6 High throughput transcriptome analysis

After treatment, ten samples were taken for transcriptome analysis with untreated materials as control. Total RNA from plant tissues was extracted using the total RNA extraction kit (Sangon, Shanghai, China, SK1321) and remove genomic DNA with RNase-free DNase-I treated. RNA integrity was evaluated using the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) (Du et al., 2019). The libraries were constructed using TruSeq Stranded mRNA LTSample Prep Kit (Illumina, San Diego, CA, USA) according to the manufacturer's instructions and sequenced using Illumina Hi-Seq platform (Bolger et al., 2014). The fluorescent image processing, base-calling, and calculation of quality value were performed by Illumina data processing pipeline 1.4 (Illumina®, San Diego, California, USA).

The raw data was filtered and trimmed for low-quality score reads, adaptor, and primer sequences were removed and rawdata (rawreads) were processed using Trimmomatic. The reads containing ploy-N and the low quality reads were removed to obtain the clean reads (Bolger et al., 2014). HISAT2 V2.1.0 was selected to hierarchical indexing for spliced alignment of transcripts to compare transcriptome sequencing reads to the reference genome (Kim et al., 2015). Transcripts Per Million mapped reads (TPM) value of each gene was calculated using cuff links, and there adcounts of each gene were obtained by htseq-count (Trapnell et al., 2010; Anders et al., 2015). DEGs were identified using the DESeq package function estimate Size Factors and nbinomTest. P value < 0.05 and fold Change > 2 or fold Change < 0.5 was set as the threshold for significantly differential expression (Anders and Huber, 2012). Gene Ontology (GO) enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of DEGs were respectively performed using R based on the hypergeometric distribution using data from http://www.geneontology.org/ and http://www.genome.jp/kegg/, respectively.

2.7 Statistical analysis

The statistical analysis was performed using Graph Pad Prism 9 (https://www.graphpad.com/). The experiments were performed with three biological replicates, and plant materials from ten seedlings were pooled for each biological replicate. The statistical significance was determined through a Two-way ANOVA and Tukey's test. The difference was considered to be statistically significant as **** $P<0.0001$, *** $P \le 0.001$, ** $P \le 0.01$, and * $P \le 0.05$.

3 Results

3.1 Physiological analysis of the response of buckwheat seedling to cadmium stress

Buckwheat varieties DK19 and K33 were treated with Cd at the two-leaves seedling stage, and the samples were taken at an interval of one week, the diversity of physiological and phenotypic responses under Cd stress was studied. Compared with normal planting condition (CK), the leaves of DK19 began to curl and wilt from the edge on the 7th day of Cd treatment (Figure 1B). On the 14th day, the leaves still maintained a relatively complete shape, and only chlorosis occurred at the edge (Figure 1C). In order to further study the effect of Cd stress on the photosynthetic capacity of tartary buckwheat, we detected the chlorophyll content of material DK19 at different Cd treatment times. Results in Figure 1D show the variation trend with the extension of Cd stress treatment time. After 7 days of treatment, the chlorophyll content in DK19 decreased, which was 4.112 mg/g. When treated for 14 days, the

chlorophyll content in leaves decreased significantly, which was 2.058 mg/g compared with the control.

In K33, there was no significant difference in the leaf phenotype, relative to DK19, in the normal planting condition (Figure 1E). All leaves produced significant curly wilting on the seventh day of treatment, which could not be flattened (Figure 1F). By the 14th day, the edges of leaves significantly turned yellow and withered, and extended to the center (Figure 1G). Significant differences in leaf phenotypic changes were observed between the two materials. It can be seen from the figures that there were no significant difference in chlorophyll content between DK19 (4.623 mg/g) and K33 (4.879 mg/g) leaves under the same conditions without Cd treatment ($P<0.05$), which was consistent with the leaf phenotype. After seven days of treatment, the chlorophyll content in K33 decreased, which was 3.343 mg/g. The chlorophyll content in K33 decreased significantly. When treated for 14 days, the chlorophyll content in K33 leaves decreased significantly, which was 1.836 mg/g respectively compared with the control.

3.2 Difference of malondialdehyde content and antioxidant enzyme system

MDA content of two buckwheat varieties increased with the prolongation of cadmium stress treatment time. The MDA content of the two buckwheat varieties increased with increasing Cd stress treatment time. On the 7th and 14th days of treatment, the MDA content in DK19 increased significantly ($P < 0.01$). The promotion rate on the 7th day was 23.752mmol/g, and the promotion rate on the 14th day was 23.729mmol/g (Figure 2A). There was no significant difference in MDA content between K33 and DK19 when untreated ($P>0.05$). The content was 33.125mmol/g (K33) and 52.168mmol/g (Figure 2B) after 7 and 14 days of treatment, respectively. After 14 days of treatment, the MDA content of the two varieties was significantly different from that of the control group ($P < 0.001$). In different Cd treatment periods, the MDA content in K33 leaves was always higher than that in DK19, indicating that the membrane damage in K33 was more serious than that in DK19. It can be seen from Figure 2C that the activity of POD kept up-regulated in the two treatment stages. With the passage of time (Figure 2B), the activity of POD in DK19 leaves was significantly up-regulated ($P < 0.01$) and extremely significantly up-regulated ($P < 0.001$), while the activity of POD in K33 leaves was significantly lower than DK19 in each period ($P < 0.001$) (Figure 2D). Figure 2E and Figure 2F show the change of CAT activity in leaves. CAT activity in DK19 leaves was lower than K33 in each period of treatment, and showed significant difference and extremely significant down-regulation in control and treatment for 7 days respectively. CAT activity in K33 leaves was slightly higher than DK19 in treatment for 14 days. It indicated that more drastic changes in CAT activity occurred in DK19 leaves with the passage of treatment time. The changes of MDA content and antioxidant enzyme activity in buckwheat leaf tissue under Cd stress showed that the treatment of Cd stress significantly changed the redox reaction in buckwheat, and there was a significant difference between the two materials.

3.3 Statistical description of metabolic analysis of untargeted metabolomics sequencing

Untargeted metabolomics was used to detect the accumulation of metabolites in buckwheat leaves tissues at the period with significant phenotypic variation (day 7 and day 14) after Cd treatment. The content of metabolites in buckwheat leaves without Cd treatment was used as a control in DK19 and K33, respectively. A total of 1798 metabolites were identified in this study, which were divided into 19 categories (Figure 3), and the metabolites were annotated according to their chemical structure. Among them, 381 metabolites are organic acids and derivatives. There were 362, 295, 277, 169 and 158 metabolites detected in the categories of organic heterocycles, benzene, lipids and lipid molecules, phenylpropane, polyketones and organic oxygen compounds. The third kind of super categories, like organic nitrogen compounds, nucleosides, nucleotides, and analogues and alkaloids and derivatives, containing 52, 39 and 22 metabolites were detected respectively. In addition, a large number of organo sulfur compounds (9), organic polymers (6), mixed metal/non-metal compounds (5), hydrocarbons (5), hydrocarbon derivatives (4), lignans, neolignans and related compounds (4), organohalogen compounds (4), organic 1, 3-dipolar compounds (1), organic salts (1), organophosphorus compounds (1) were deteced.

A total of 153 classes were identified through an extensive comparison of the literature. Among these classes, the groups with the highest number of metabolites were carboxylic acids and derivatives (324), benzene and substituted derivatives (194), organooxygen compounds (153), and fatty acids (116). In addition, classes containing more than 20 kinds of metabolites include flavonoids (81), steroids and steroid derivatives (67), prenol lipids (59), organonitrogen compounds (52), indoles and derivatives (45), pyridines and derivatives (28), imidazopyrimidines (27), phenol ethers (26), coumarins and derivatives (25), naphthalenes (25), peptidomimetics (25), phenols (22), benzopyrans (21), azoles (20) and glycerophospholipids (20). Rich metabolite types were first detected in buckwheat after cadmium stress treatment in this study. Among 153 types, 27 types detected extensive variation before and after treatment. Among them, amino acid metabolism, antioxidant metabolism and energy metabolism all contain a large number of variations, which will be analyzed later in this study.

3.4 Screen of differentially accumulated metabolites

Through the screening of metabolites, differentially accumulated metabolites (DAMs) with criteria FC>1.50, P-value<0.05 and VIP>1 were considered to determine the significant differences between samples (Punia et al., 2021). In DK19, compared with CK, 534 metabolites were up-regulated and 632 metabolites were down-regulated after 7 days of treatment (Figure 4A). In the second treatment period (14 days), 810 metabolites were up-regulated and 602 metabolites were down-regulated, taking the materials before treatment as the control (Figure 4B) and 734 up-regulated metabolites

10.3389/fpls.2023.1142814

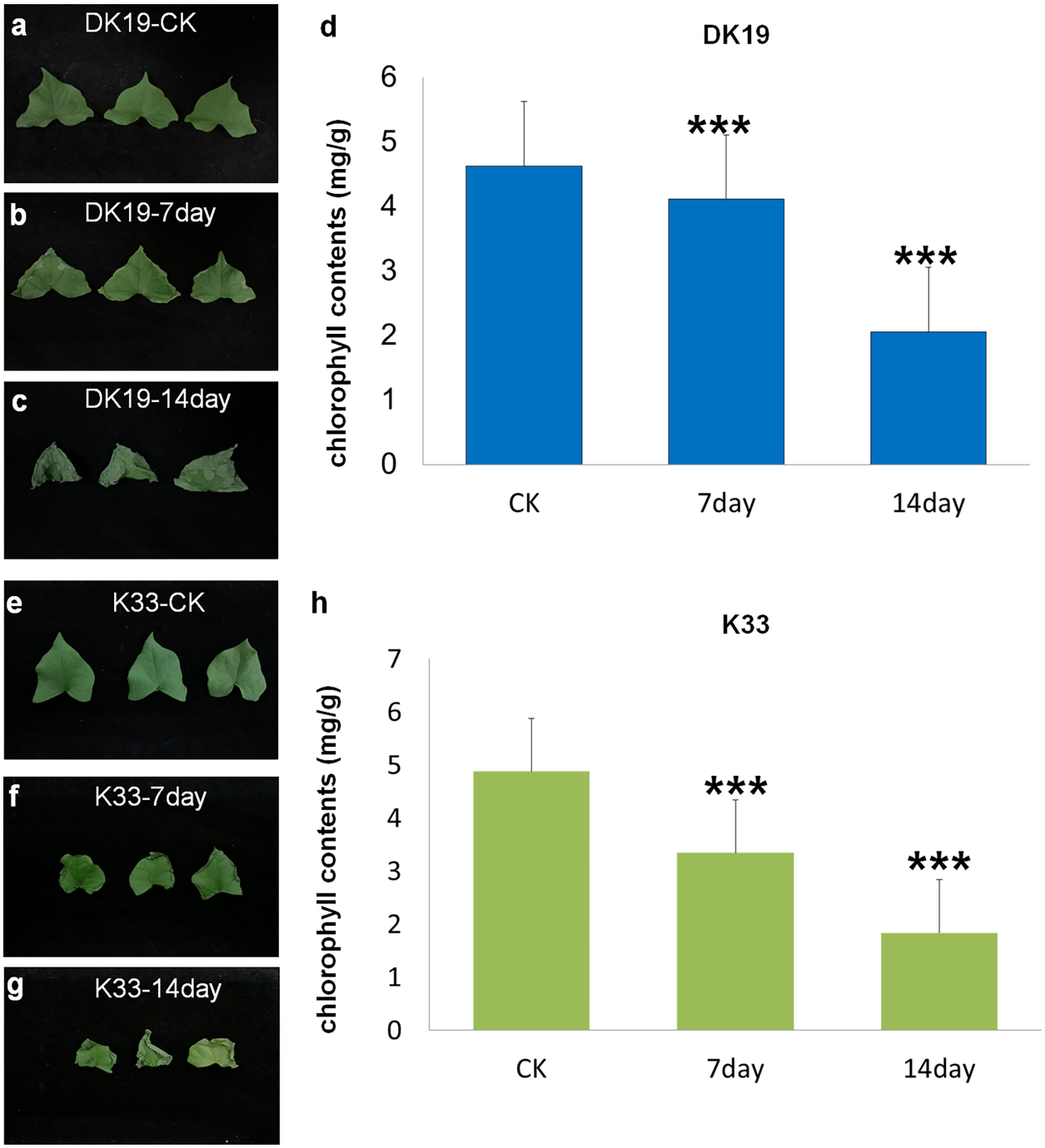


FIGURE 1
Phenotypic detected before and after Cd stress in DK19 and K33. **(A)** The phenotype of fresh buckwheat leaves was before treatment (CK) in buckwheat DK19. The leaves contain sufficient water, showing good spreading and extension, without wilting and drying. **(B)** The phenotype of fresh buckwheat leaves after seven days of Cd treatment in DK19. Leaves did not show significant variation relative to the control group. **(C)** The phenotype of fresh buckwheat leaves after 14 days of Cd treatment in DK19. The leaves wilted, showing a significant water loss state, but the leaves did not appear withered and yellow. **(D)**, The diversity in chlorophyll contents. The abscissa represents the samples with different treatments, and the ordinate is the chlorophyll content detected, in mg/g. **(E)**, The phenotype of fresh buckwheat leaves was before treatment (CK) in buckwheat K33. The leaves contain sufficient water, showing good spreading and extension, without wilting and drying. **(F)**, The phenotype of fresh buckwheat leaves after seven days of Cd treatment in K33. The leaves wilted, showing a significant water loss state, but the leaves did not appear withered and yellow. **(G)**, The phenotype of fresh buckwheat leaves after 14 days of Cd treatment in K33. The leaves showed significant wilting and serious water loss. A large number of withered and yellow states appeared at the edge of the leaves, and gradually extended to the center. **(H)**, The diversity in chlorophyll contents. The abscissa represents the samples with different treatments, and the ordinate is the chlorophyll content detected, in mg/g. *** $P \leq 0.001$.

10.3389/fpls.2023.1142814

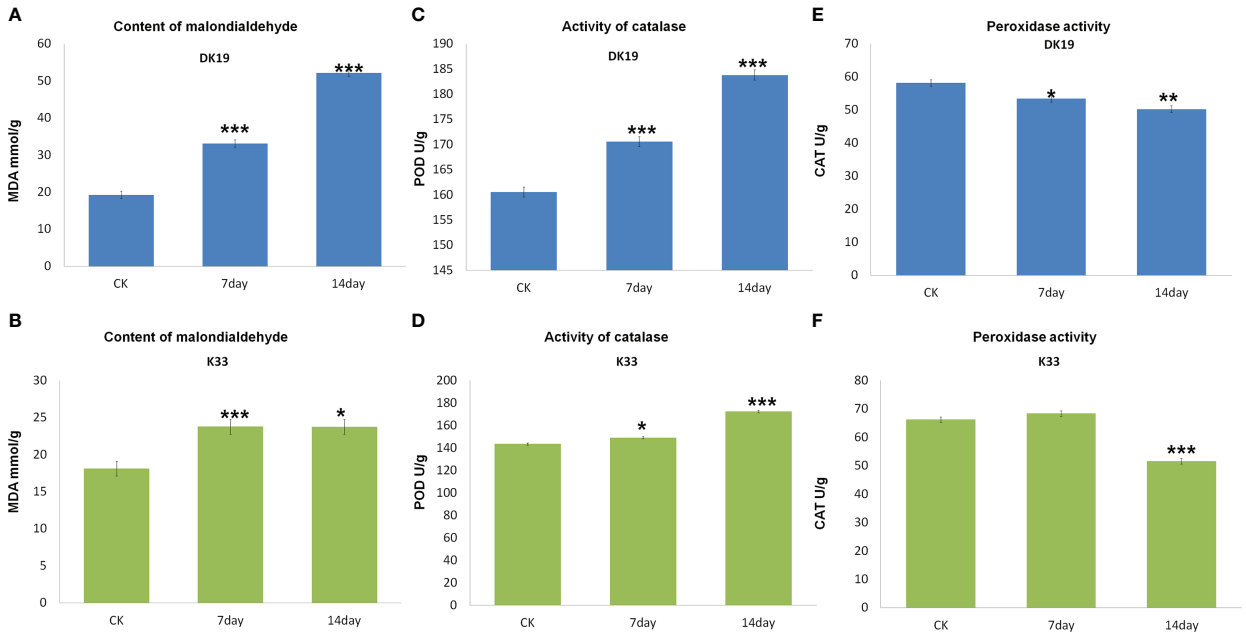


FIGURE 2
Difference of malondialdehyde content and antioxidant enzyme system between buckwheat DK19 and K33 under Cd stress. **(A)** Difference of MDA content in buckwheat leaves under Cd stress in DK19. **(B)** Difference of MDA content in buckwheat leaves under Cd stress in K33. **(C)** Difference of POD content in buckwheat leaves under Cd stress in DK19. **(D)** Difference of POD content in buckwheat leaves under Cd stress in K33. **(E)** Difference of CAT content in buckwheat leaves under Cd stress in DK19. **(F)** Difference of CAT content in buckwheat leaves under Cd stress in K33. * in green indicated that the difference of the same variety under different treatments is significant (P<0.05). * in black indicated significant difference (P<0.05) between different varieties under the same treatment. ** P ≤ 0.01, *** P ≤ 0.001.

and 377 down-regulated metabolites were identified between the two stages of the Cd treatment (Figure 4C). Correspondingly, in K33, 390 up-regulated metabolites and 416 down regulated metabolites were detected in the first stage of treatment (7 days) relative to that before treatment (Figure 4D); On the fourteenth day of treatment (the second stage), 448 up-regulated metabolites and 425 depressed metabolites were detected relative to the control group (Figure 4E); Between the two treatments, the up-regulated and down-regulated metabolites were 487 and 412, respectively (Figure 4F). Corresponding to the observed phenotypic variation, with the dramatic change of phenotype, the metabolism in buckwheat cells changed dramatically under Cd stress. There were significant differences in the trends of these differential metabolites between the two materials, shows that DK19 has undergone drastic metabolic changes to respond to Cd stress, thus improving the resistance.

The Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis was performed on the DAMs using KOBAS software. KEGG enrichment was carried out for DAMs between different treatments, and the 25 top pathways are shown in Figure 5. In DK19, in the first stage of treatment, the pathways which enrich the most metabolites are Riboflavin metabolism, Vitamin B6 metabolism, Phenylalanine, tyrosine and tryptophan, Nicotinate and nicotinamide metabolism (Figure 5A). In the second treatment period (14 days), the top five KEGG pathways, Phenylalanine metabolism, Phenylalanine, tyrosine and tryptophan, Riboflavin metabolism, Purine metabolism and Glutathione metabolism are enriched, taking the materials before treatment as the control (Figure 5B). Between the two stages after Cd treatment, DAMs were enriched in Riboflavin metabolism, Nicotinate and nicotinamide metabolism, Vitamin B6 metabolism, Phenylalanine metabolism and Phenylalanine, tyrosine and tryptophan (Figure 5C). For material K33, in the first stage of treatment, the pathways enriched to the most metabolites were Riboflavin metabolism, D-Glutamine and D-glutamate metabolism, Purine metabolism, Phenylalanine, tyrosine and tryptophan and Arginine biosynthesis (Figure 5D). In the second stage, there were differences pathways between the DK19 and K33, except for a enriched pathway Phenylalanine, tyrosine and tryptophan were enriched in two materials, four pathways, Nicotinate and nicotinamide metabolism, Glutathione metabolism, Pantothenate and CoA biosynthesis were special enriched in K33 (Figure 5E). Between the two stages of treatment, the metabolites were similarly enriched in three pathways (Riboflavin metabolism, Nicotinate and nicotinamide metabolism and Vitamin B6 metabolism), between the two materials (Figure 5F).

The metabolite changes of the two materials after Cd treatment showed consistency. For example, among the pathways with the most enriched genes, the three pathways Riboflavin metabolism, Nicotinate and nicotinamide metabolism and Phenylalanine, tyrosine and tryptophan were significant present almost every period. In addition, among other significantly enriched pathways, many pathways were significantly enriched in the two materials at different periods. This result showed that different buckwheat materials, whether resistant or sensitive, showed similar metabolic changes after Cd treatment, which may be the basis of buckwheat tolerance. On the other hand, the difference of metabolite

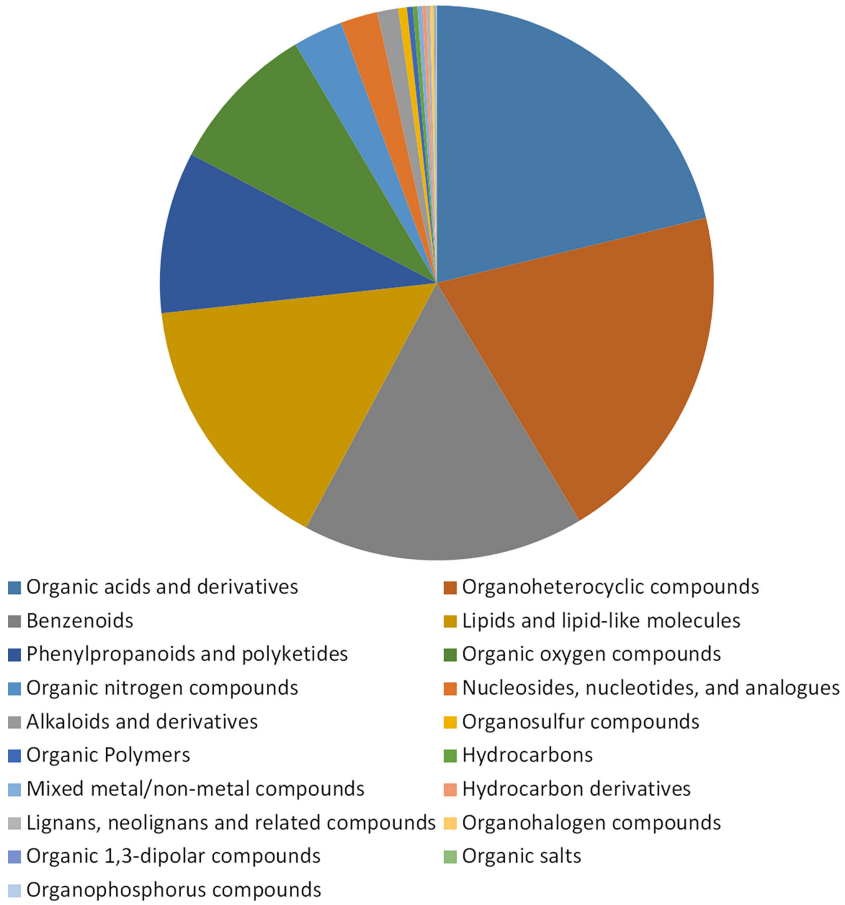


FIGURE 3
Statistical description of metabolic analysis of buckwheat on Cd response. Classification of the 1798 known identified metabolites in non-targeted metabolome analysis. The ratio of the detected metabolites to the total metabolites in the pie chart. Blocks with different colors indicate different supertypes of metabolite values, and different types are arranged in alphabetical order of names.

enrichment pathways between the two materials in the same period also indicates the basis of the resistance difference between the two materials. It is suggested that the differences of specific metabolic pathways may be the basis of resistance differences among materials on the basis of the consistent overall trend.

Subsequently, we analyzed the differential metabolites of two materials (DK19 and K33) at the same treatment stage (Figure 6). As shown in Figure 6A, there were 536 metabolites up-regulated and 540 metabolites down-regulated in DK19 relative to K33. These metabolites were mainly enriched in 16 pathways including riboflavin metabolism, arginine and proline metabolism, phosphonate and phosphinate metabolism, vitamin B6 metabolism and glutathione metabolism (Figure 6B). In the first period (7 days) of treatment, 539 metabolites were up-regulated and 518 metabolites were down-regulated in DK19, compared with K33 (Figure 6C). These metabolites were mainly enriched in 18 pathways, of which riboflavin metabolism, phenylalanine, tyrosine and tryptophan, ubiquinone and other terpenoid quinone, vitamin B6 metabolism and phenytalanine metabolism were significantly enriched (Figure 6D). After 14 days of treatment, 403 up-regulated metabolites and 717 down-regulated metabolites were detected in DK19, and K33 was used as control (Figure 6E). KEGG enrichment showed that these differentially expressed metabolites were significantly enriched in riboflavin metabolism and 20 pathways such as nicotinate and nicotinamide metabolism, vitamin B6 metabolism, phenytalanine metabolism and arginine and proline metabolism (Figure 6F). These results showed that there were some metabolic differences between the two materials before Cd treatment, which may be the basis of the difference in resistance between the two materials. In a period defined by treatment, the two materials had significant differences in metabolite accumulating in arginine and proline metabolism, phosphonate and phosphinate metabolism, phenytalanine, tyrosine and tryptophan, ubiquinone and other terpenoid-quinone, glutathione metabolism, phenytalanine metabolism and nicotinate and nicotinamide metabolism pathways on the basis of original differences. In the second period, the differences to be negotiated are nicotinate and nicotinamide metabolism, phenytalanine metabolism and glutathione metabolism, which continue the change trend of the first period, and purine metabolism, butanoate metabolism, citrate cycle (TCA cycle) and alaine, aspartate and glutamate metabolism, which change significantly in this period. The comparison of the two materials at the same treatment stage showed that at the metabolic level, the

10.3389/fpls.2023.1142814

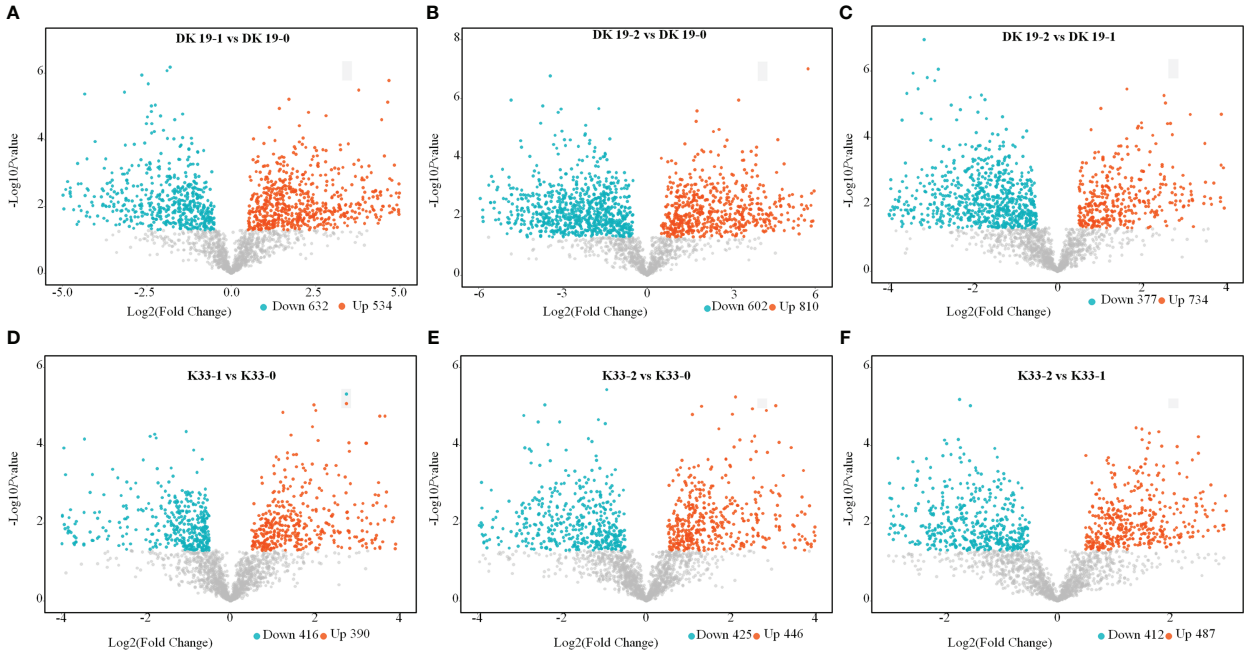


FIGURE 4
Screen of differentially accumulated metabolites (DAMs), generated using untargeted metabolomics sequencing. The volcano plot presents the expression of the DAMs in different treatments, the horizontal axis is Log2 (Fold Change) value, and the vertical axis is - Log10*p*value, the red dots represent up-regulated metabolites, and the blue dots represent down-regulated metabolites. **(A–C)** DAMs in DK19 between different stages of treatment. **(D–F)** DAMs in K33 between different stages of treatment.

resistance to Cd stress had different changes, and the phenotypic differences were obtained.

3.5 Identification and functional analysis of differentially expressed genes

High throughput transcriptome analysis was performed using the same materials, a total of 122.58 GB of clean data was obtained, the transcript abundance of each gene was evaluated by Transcripts Per Million mapped reads (TPM) on the basis of the number of uniquely mapped reads that overlapped with exon regions. A total of 29,739 expressed genes were identified. The genes which were detected twice in three replicates and with an average expression of more than 0.5 in TPM were selected as expression gene, and the differentially expressed genes at different stages of treatment were detected according to the standard of P value < 0.05 and fold Change > 2 or fold Change < 0.5. A Venn diagram was created to identify the expressed genes of each material, 367 (DK19-1 *vs.* DK19-0), 211 (DK19-2 *vs.* DK19-0), 629 (DK19-2 *vs.* DK19-1), 263 (K33-1 *vs.* K33-0), 265 (K33-2 *vs.* K33-0) and 370 (K33-2 *vs.* K33-1) differentially expressed genes (DEGs) were identified, respectively (Figures 7A, B). Genes with continuous changes in the three periods were not detected in the two materials, and 91 genes were changed in the two periods of treatment, compared with the control, in DK19. The expression of 41 genes changed in the first stage of treatment compared with the control and between the two periods of treatment. 47 genes had expression changes in the second period of treatment, whether compared with the control or the first period. The number of differentially expressed genes distributed in k33 was 51 (K33-1 *vs.* K33-0 compared with K33-1 *vs.* K33-0), 52 (K33-1 *vs.* K33-0 compared with K33-2 *vs.* K33-1) and 94 (K33-2 *vs.* K33-0 compared with K33-2 *vs.* K33-1).

In order to further analyze the DEGs, Gene Ontology (GO) classification analysis of the identified unique DEGs was performed. All the DEGs were divided into the biological process, cellular component, and molecular function categories, the top 35 GOs in DK19 were shown in Figure 7C, and top 35 GOs in K33 were shown in Figure 7D. In DK19, five GOs, biological regulation, cellular process, metabolic process, response to stimulus, belong to biological process, with the most genes enriched. Three GOs, extracellular region, membrane, organelle, belong to cellular component, with the most genes enriched. Binding and catalytic activity were two GOs with the much genes enriched in molecular function. In most GOs, the combination of DK19-2 *vs.* DK 19-0 down enriched the most differentially expressed genes, followed by DK 19-1 *vs.* DK 19-0 up and DK 19-2 *vs.* DK 19-0 down. Similar GO enrichment trends were also detected in K33, and the GOs with a large number of genes include cellular process, metabolic process, reproductive process, macromolecular complex, membrane-enclosed lumen, binding, catalytic activity. The combinations containing the most genes were K33-2 *vs.* K33-1 Down, K33-1 *vs.* K33-0 up and K33-2 *vs.* K33-0 down. The results of GO analysis showed that the two materials had significant changes in cell composition and expression regulation during cadmium treatment. Especially in the first stage of change, many differentially expressed genes were detected in both materials, which was consistent with the stage of stress response. On

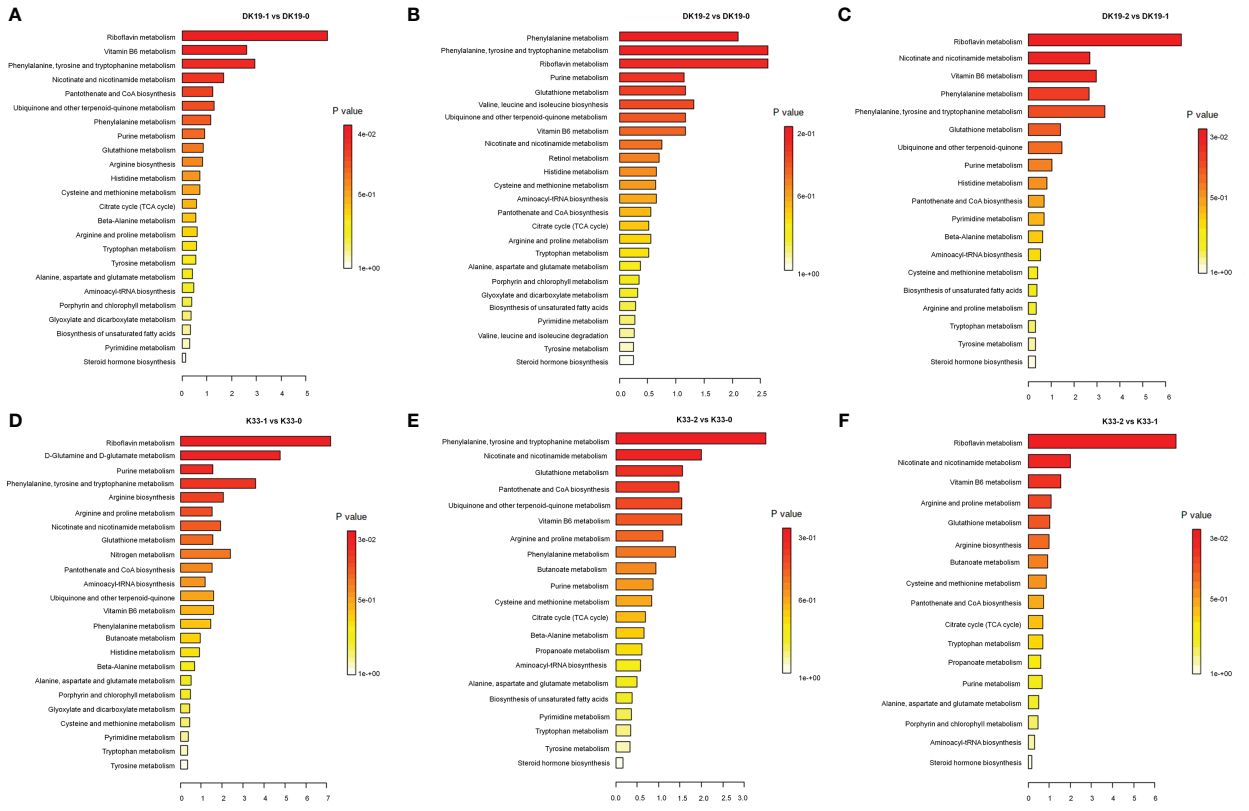


FIGURE 5
Screen of differentially accumulated metabolites (DAMs), generated using untargeted metabolomics sequencing. The Top KEGG pathways of all DAMs between different stages of treatment. The abscissa is the percent of differentially accumulated metabolites in different KEGG, and the ordinate is the type of KEGG enrichment. The darker the color, the higher the *p* value. **(A–C)** DAMs in DK19 between different stages of treatment. **(D–F)** DAMs in K33 between different stages of treatment.

the basis of some similar trends, there are also significant differences, even showing opposite expression changes in the two materials, which are related to the changes of phenotypes and physiological indicators.

When subjected to a Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis, using Q-value less than 0.05 as the significant enrichment threshold to all the DEGs for KEGG pathway, the significantly enriched 15 pathways are shown in Figure 8. All the pathways were divided into cellular processes, environmental information processing, genetic information processing, metabolism, and organismic systems. In DK19, at the 7^{th} day, the differentially expressed genes were mainly enriched in pathways such as amino acid metabolism (13), carbohydrate metabolism (22), lipid metabolism (16). After two weeks treatment, the DEGs were significantly enriched in carbohydrate metabolism (12) and lipid metabolism (7). Between the two stages of treatment, DEGs were enriched in signal transduction (25), translation (15), amino acid metabolism (14), carbohydrate metabolism (26), lipid metabolism (16), the most DEGs were enriched in carbohydrate metabolism and signal transduction, in particularly. In the lines K33, the most enriched KEGGs were amino acid metabolism (12) and biosynthesis (11) of DEGs between K33-1 *vs.* K33-0, the most enriched KEGG was carbohydrate metabolism (11) between K33-2 *vs.* K33-0, carbohydrate metabolism (16), lipid metabolism (15) and translation (11) were the most three KEGGs between K33-2 *vs.* K33-1. By comprehensively comparing the enrichment of differential expression between the two lines in the three periods, it is obvious that through KEGG analysis, the differentially expressed genes are significantly enriched in carbohydrate metabolism, lipid metabolism, and translation. The correlation analysis between the KEGG enriched by differentially expressed genes and the KEGG enriched by differentially expressed metabolites found that the relevant pathways were mainly in the arginine and proline metabolism, the citrate cycle (TCA cycle) and the metabolism of alanine, aspartic acid and glutamic acid, as well as the metabolism of amino acid, carbohydrate metabolism, lipid metabolism and translation regulation, and were further analyzed in the follow-up study.

3.6 Key metabolic pathways identified by joint analysis

According to the previous research results, the obtained differential genes were uploaded to GO database and KEGG database for functional annotation and pathway enrichment. The results showed that the differential genes were significantly enriched in nitrogen metabolism, galactose metabolism, phenylpropanoid biosynthesis, glutathione metabolism, carbohydrate metabolism, linolenic acid metabolism and other metabolic pathways. The metabolomic analysis of buckwheat samples by non-targeted

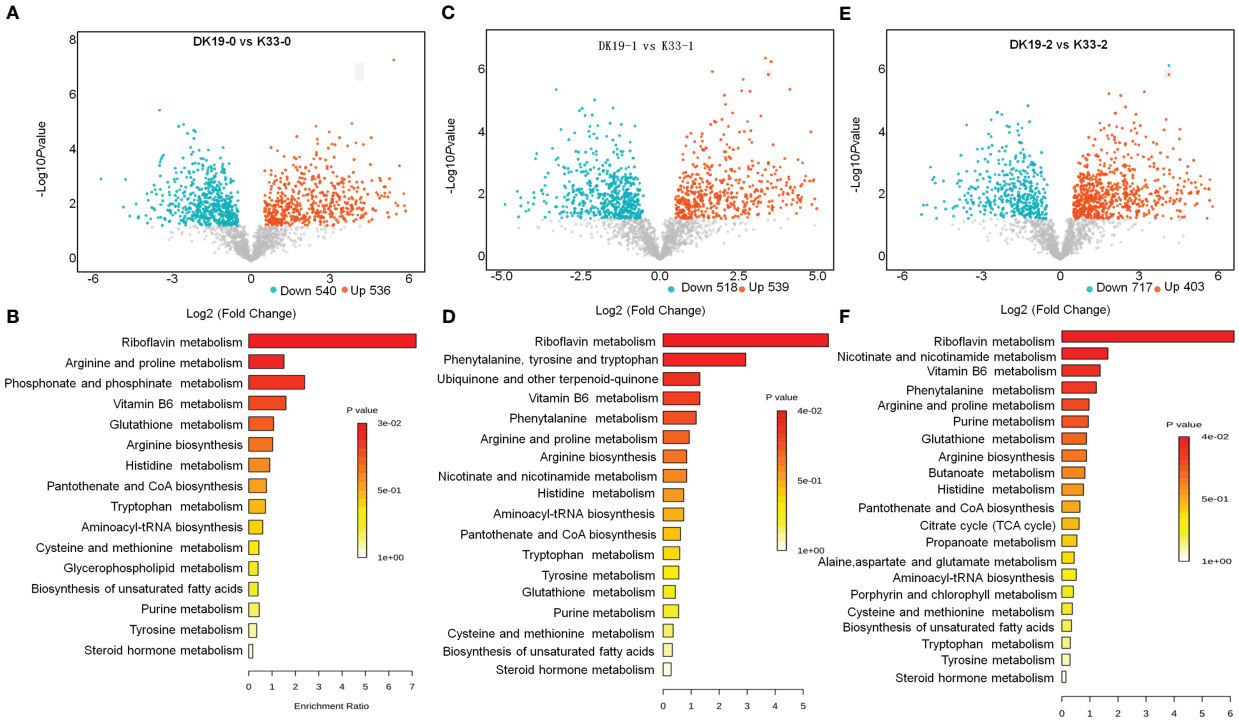


FIGURE 6
Screen of differentially accumulated metabolites (DAMs) at the same stages between DK19 and K33, generated using untargeted metabolomics sequencing. **(A, C, E)** Identification of differentially accumulated metabolites (DAMs) between DK19 and K33 at the control, 7days after the treatment and 17days after the treatment, generated using untargeted metabolomics sequencing technology. The volcano plot presents the expression of the DAMs in different treatments, the red dots represent up-regulated metabolites, and the blue dots represent down-regulated metabolites. **(B, D, F)** KEGG pathways of all DAMs between the DK19 and K33 at the same stages. The abscissa is the percent of differentially accumulated metabolites in different KEGG, and the ordinate is the type of KEGG enrichment. The darker the color, the higher the *p* value.

metabolomics showed that the main metabolic pathways for the enrichment of differential metabolites were the biosynthesis of secondary metabolites, amino acid biosynthesis, galactose metabolism, glycerophosphatide metabolism, and so on. Combined with transcriptome and metabolome data, galactose metabolism, lipid metabolism (glycerophosphatide metabolism and glycerophosphatide metabolism) and glutathione metabolism, which are significantly enriched at gene level and metabolic level, are further analyzed.

Transcriptome of this study showed that 25 differential genes were enriched in galactose metabolism (galactose metabolism, ko00052), 7 differential genes were enriched in fructose and mannose metabolism (fructose and mannose metabolism, ko00051), and 43 differential genes were enriched in starch and sucrose metabolism (starch and sucrose metabolism, ko00500) after 7 days of Cd stress. After 14 days of Cd stress, 20 differential genes were enriched in galactose metabolism, 8 differential genes were enriched in fructose and mannose metabolism, and 30 differential genes were enriched in starch and sucrose metabolism. The metabolome results showed that 6 differential metabolites were enriched in galactose metabolism, 3 differential metabolites were enriched in fructose and mannose metabolism, and 1 differential metabolite was enriched in starch and sucrose metabolism. We uploaded the differential genes and metabolites to the KEGG website for enrichment analysis. In this study, the selected 10 candidate genes were tested to verify the galactose metabolic pathway, and the results are shown in Figure 9.

In this study, the transcriptional results showed that after 7 days of Cd stress, 15 differential genes were enriched in glycolipid metabolism (ko00561), 12 differential genes were enriched in glycerophospholipid metabolism (ko00564), and 20 genes were enriched in α- Linolenic acid metabolism (alpha linolenic acid metabolism, ko00592). A total of 9 differential genes were enriched in glycerol metabolism, 13 differential genes were enriched in glycerol phospholipid metabolism, and 20 genes were enriched in α-Linolenic acid metabolism. Metabolomic results showed that 2 differential metabolites were enriched in glycerophospholipids metabolism, 3 differential metabolites were enriched in glycerophospholipids metabolism, and 1 metabolite was enriched in α-Linolenic acid metabolism. Metabolites involved in lipid metabolism include glycerol triphosphate (1.69 times), lecithin (1.53 times), palmitic acid (0.45 times), betaine (1.28 times), etc. The up regulation of genes was the main reason for the increase of lysophosphatide and glycerol-3-phosphate, and the down regulation of genes were the main reason for the decrease of palmitic acid content (Figure 10).

10.3389/fpls.2023.1142814

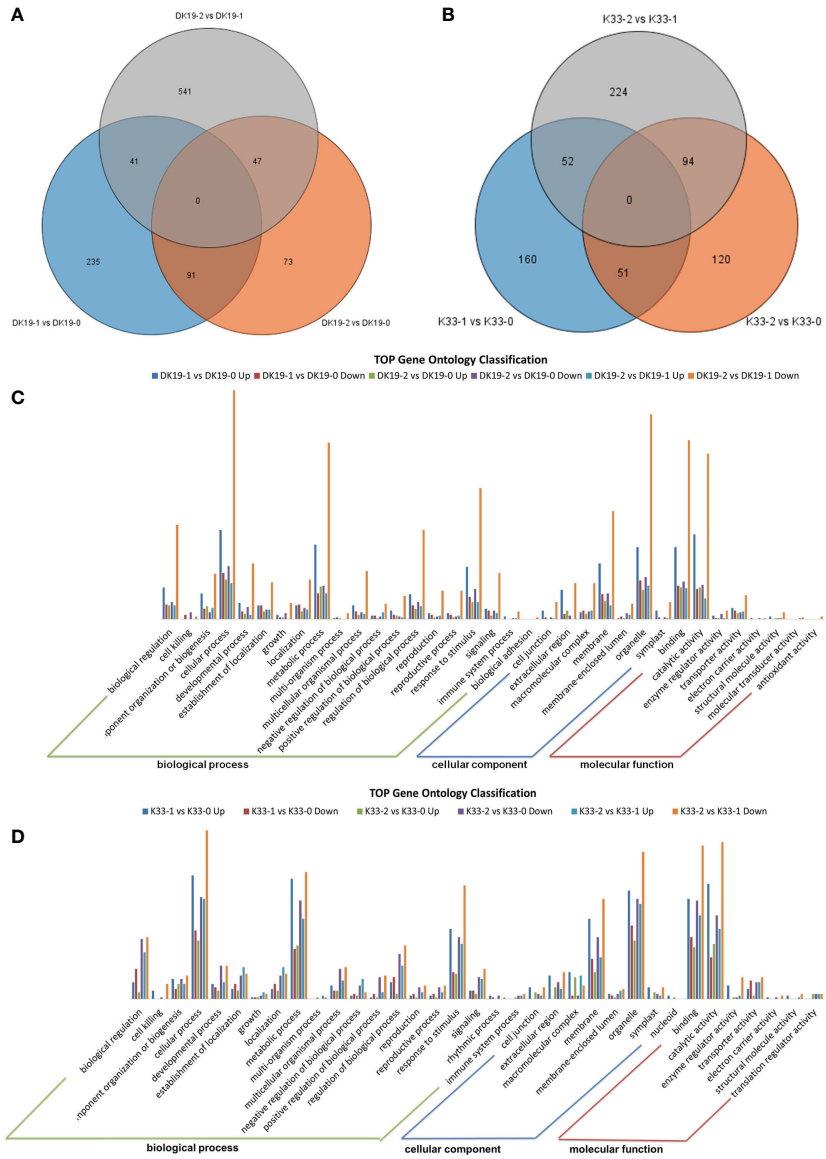


FIGURE 7
The quality and quantity of the RNA sequencing, generated using high-throughput deep sequencing technology; **(A)** Venn diagrams of differentially expressed genes (DEGs) among the three stages. DK19 and K33 represent two materials, DK9-0 and K33-0 are the control without Cd treatment, DK19-1 and K33-1 are the test results after 7 days of treatment, and DK19-2 and K33-2 are the test results after 14 days of treatment. The selected gene is average TPM> 0.5. **(B)** Relationship between samples in three biological replicates and summary of expressed genes in different treatments. **(C, D)** Venn diagrams of differentially expressed genes (DEGs) among the four stages. 0D represents the control without IAA treatment, 1D represents the sample after 1 day of treatment, 2D represents the sample after 2 days of treatment, and 3D represents the sample after 3 days of treatment. The selected gene is average TPM> 0.5.

4 Discussion

4.1 Effects of cadmium stress on plant

Excess Cd in soil has become one of the major environmental problems of global concern, because once Cd is discharged into the environment, it is difficult to collect and treat, and it has super durability (Li et al., 2019; Naeem et al., 2019; Rahman and Singh, 2019). Cd pollution in farmland soil seriously threatens agricultural productivity, food safety and human health (Sarwar et al., 2010; Wang et al., 2019). Therefore, exploring the key genes of low Cd accumulation in buckwheat is helpful to screen buckwheat with low Cd accumulation, so as to solve the important technical guarantee and industrial foundation of buckwheat safety production (Tiryakioglu et al., 2006; Borges et al., 2019). In addition, high-throughput sequencing can analyze the expression regulation under specific abiotic stress, so as to study the response of the body to changes in internal and external environment, and study the changes of gene expression regulation and metabolic regulation network (Chen et al., 2019b; Guo et al., 2020). Combining the advantages of transcriptome and metabolome can solve the deep-seated problems (Szymański et al., 2020).

10.3389/fpls.2023.1142814

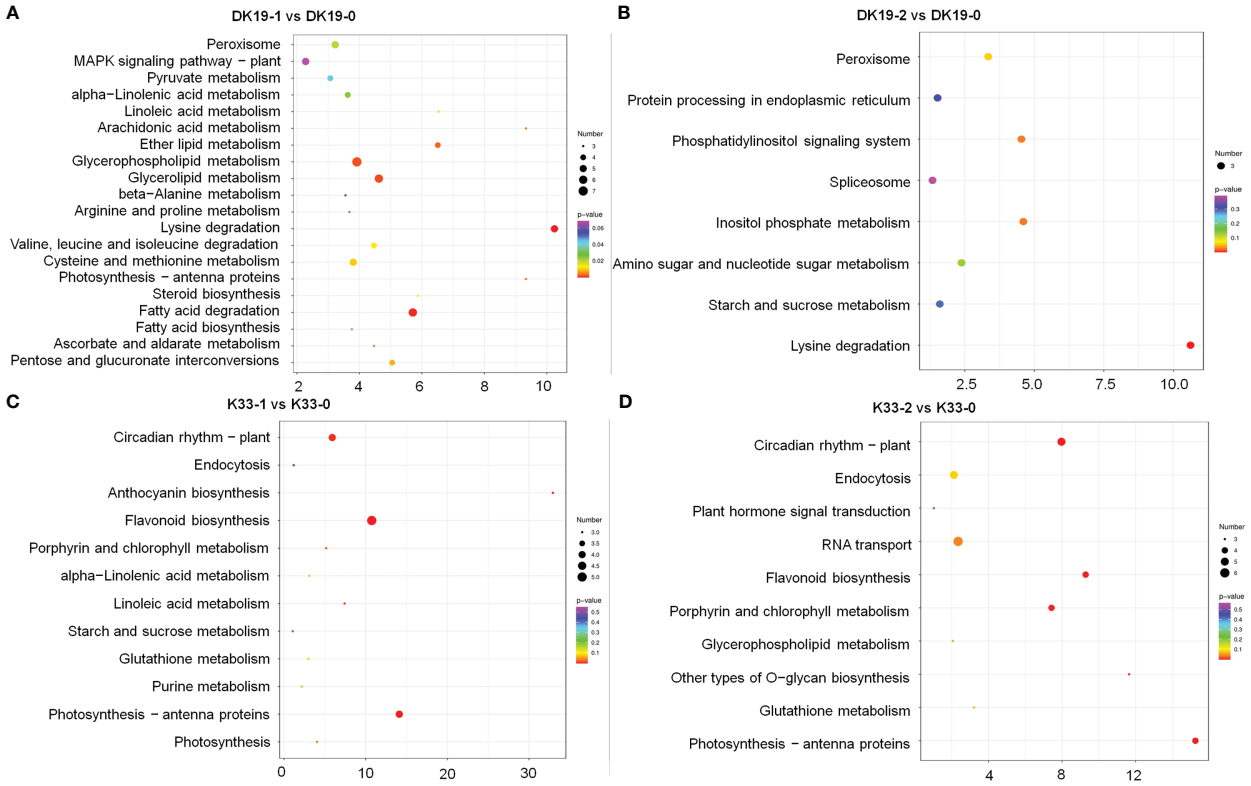


FIGURE 8
KEGG analysis of differentially expressed genes (DEGs) between DK19 and K33 at the same stages, generated using high-throughput deep sequencing technology. **(A)** The KEGG pathways of all DEGs between the DK19-1 and DK19-0. The abscissa is the percent of differentially accumulated metabolites in different KEGG, and the ordinate is the type of KEGG enrichment. The darker the color, the higher the p value. **(B)** The KEGG pathways of all DEGs between the DK19-2 and DK19-0. The abscissa is the percent of differentially accumulated metabolites in different KEGG, and the ordinate is the type of KEGG enrichment. The darker the color, the higher the p value. **(C)** The KEGG pathways of all DEGs between the K33-1 and K33-0. The abscissa is the percent of differentially accumulated metabolites in different KEGG, and the ordinate is the type of KEGG enrichment. The darker the color, the higher the p value. **(D)** The KEGG pathways of all DEGs between the K33-2 and K33-0. The abscissa is the percent of differentially accumulated metabolites in different KEGG, and the ordinate is the type of KEGG enrichment. The darker the color, the higher the p value.

4.2 Advances in molecular mechanisms related to cadmium accumulation and tolerance in plants

The results of this study are also consistent with those of other crops. Under Cd stress, the growth of leaves of seedlings slowed down, and even withered, withered and necrotic after treatment. Under Cd stress for 7 and 14 days, although the growth of DK19 was inhibited to some extent, K33 was obviously more severely poisoned (Figure 1). It can be seen that DK19 is a very good material for studying Cd resistance. Based on metabonomics and transcriptomics methods, this anti Cd mechanism of DK19 can be further analyzed. 632 (7 days of

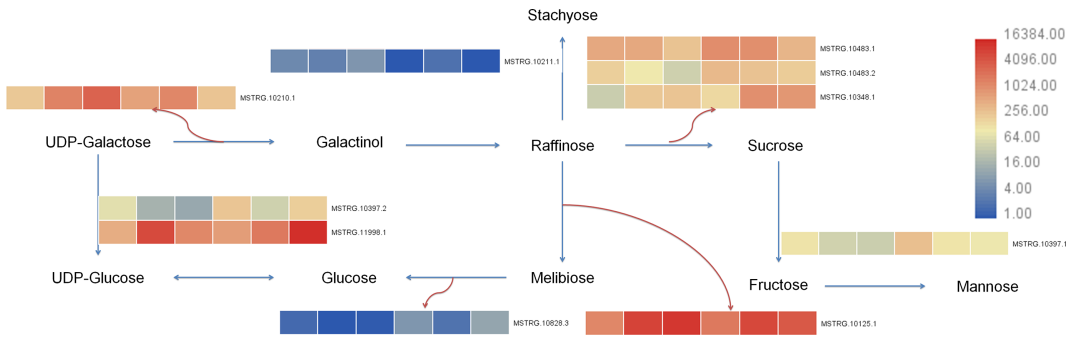


FIGURE 9
Galactose metabolism pathway in buckwheat leaves under Cd stress. The sample number in the order from left to right in the heat map: DK9-0, DK19-1, DK19-2, K33-0, K33-1 and K33-2.

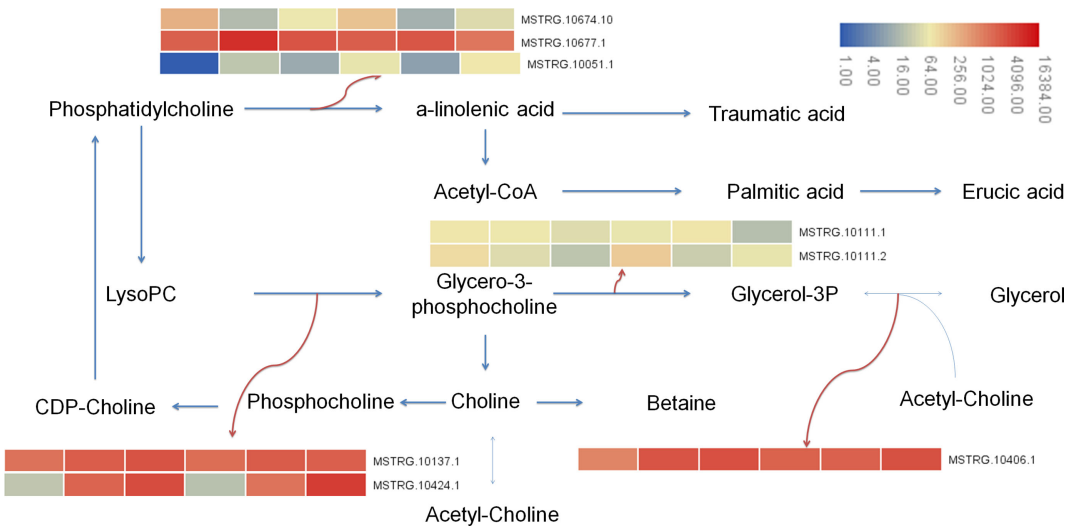


FIGURE 10
Lipid metabolism pathway in buckwheat leaves under Cd stress. The sample number in the order from left to right in the heat map: DK9-0, DK19-1, DK19-2, K33-0, K33-1 and K33-2.

treatment) and 602 (14 days of treatment) up-regulated and enriched metabolites and 534 (7 days of treatment) and 810 (14 days of treatment) down-regulated and enriched metabolites were detected in resistant material DK19 during the two periods of significant surface changes, corresponding to relatively few differentially enriched metabolites detected in sensitive material K33 (Figure 4). Through GO annotation and KEGG enrichment of differential genes and metabolites, it was found that cadmium stress affected carbon metabolism, nitrogen metabolism, galactose metabolism, plant signal transduction, glutathione metabolism, fatty acid metabolism and phenylpropanoid biosynthesis and other metabolic processes, which may be a stress response of buckwheat under cadmium stress (Figure 5). These induce oxidative stress and osmotic stress, and the change trend of carbohydrate content is consistent with the research in rice, wheat and other crops (Liu et al., 2018; Li et al., 2020).These induce oxidative stress and osmotic stress, and the change trend of carbohydrate content is consistent with the research in rice, wheat and other crops (Liu et al., 2018; Li et al., 2020).

Hossain et al. believed that the resistance of plants to a certain type of heavy metals means that plants can normally survive in a specific heavy metal environment with high content, without growth stagnation or even death. SOD, POD, GSH and other factors in plants can repair cell functions and maintain homeostasis by activating the antioxidant system to remove stress induced free radicals (Hossain and Komatsu, 2013). In recent years, many studies have used transcriptome analysis to study Cd responsive genes of different species, and identified a large number of differentially expressed genes (DEGs) responding to Cd stress. These differentially expressed genes are involved in many processes such as cell wall biosynthesis, GSH metabolism, TCA cycle and antioxidant system, which may play a key role in cell wall binding, vacuole isolation and detoxification (Peng et al., 2015; Xu et al., 2015). Similarly, transcriptome analysis showed that many differentially expressed genes in response to Cd stress were involved in cell wall modification, heavy metal transport and phenylacetone biosynthesis, respectively (Feng et al., 2018). At present, there have been many studies on metabolite analysis in plant response to various stresses, including but not limited to nutrient deficiency (Hirai and Saito, 2004), mineral toxicity (Uraguchi et al., 2011), temperature and oxidative stress and osmotic stress (Gong et al., 2005). In addition, mass spectrometry metabonomics was used for the first time to conduct metabonomic analysis of indica rice grains under Cd stress. The results showed that carbohydrate, organic acid and amino acid metabolism were affected by Cd stress, and the content of cysteine increased, which could improve the tolerance to Cd toxicity (Zeng et al., 2021).

4.3 Effects of cadmium stress on plant antioxidant system

Cd also has negative effects on physiological and biochemical processes such as photosynthetic efficiency and enzyme activity of plants (Mostofa et al., 2015). The accumulation of Cd in plants leads to the impairment of amino acid biosynthesis, inhibition of enzyme activity, oxidative stress response, interference in mineral nutrition absorption, and metabolic imbalance. Cd can destroy cell membranes, biomolecules, and organelles in plants by increasing the production of reactive oxygen species (ROS). Plant resistance to Cd stress depends on their ability to scavenge ROS. In this study, we detected significant content changes in malondialdehyde (MDA), peroxidase (POD) activity and activity of catalase (CAT) on the 7th and 14th days of cadmium treatment of the two materials. Through the analysis of transcriptome and metabolomic data, enriched differential metabolites and differentially expressed genes detected pathways including galactose metabolism, lipid metabolism (glycerol phospholipid metabolism and glycerol phospholipid metabolism) and glutathione metabolism. Ten key genes with

different expression were identified in Galactose metabolism pathway and nine key genes with different expression were identified in Lipid metabolism pathway. The further identified differential metabolites and related core genes are also consistent with previous studies.

Previous studies have shown that oxidative stress under Cd stress enhances the chemical reactivity in plants, and higher levels of reactive oxygen species will lead to protein and nucleic acid degradation, lipid peroxidation, and cell membrane damage (Kaya et al., 2020; Hayat et al., 2021). Scavenging reactive oxygen species is an important defense mechanism for plants to resist external stress. Plants have evolved antioxidant mechanisms and ascorbic acid glutathione mechanisms to avoid the harmful effects of reactive oxygen species. Among them, the main antioxidants are superoxide dismutase (SOD), catalase (CAT) and peroxidase (POD) and ascorbic acid peroxidase (APX), glutathione reductase (GR) and dehydroascorbic acid reductase (DHAR) of ascorbic acid glutathione system (Dat et al., 2000; Sharma et al., 2012; Sachdev et al., 2021).

5 Conclusion

Using transcriptome and metabolome data, we identified a large number of candidate genes and metabolites involved in the key biological pathways of buckwheat Cd stress. Our results show that the regulatory mechanism of Buckwheat under Cd stress may not only be a single gene or metabolite, but a complex regulatory and signaling mechanism. How these candidate genes or metabolites participate in buckwheat cadmium stress deserves further study.

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: NGDC (https://ngdc.cncb.ac.cn/), accession number SRA716568.

Author contributions

DD conceived and designed the experiments. DH and CW wrote the manuscript. JZ and LQ performed the experiments. YH and JZ analyzed the data. DD and CW revised the manuscript. CW and DD contributed equally to the paper. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Key R&D Program of China (2019YFD1001300, 2019YFD1001303), the Critical Talent Workstation Project (TYSGJ202201), the National Natural Science Foundation of China (31960415), the Academician Workstation Project (TYYSZ201707), the Research Program Sponsored by State Key Laboratory of Integrative Sustainable Dryland Agriculture, Shanxi Agricultural University 202105D121008-3-5), the Fundamental Research Project of Shanxi Province (Grant No. 20210302123381) and the Teaching Content and Curriculum System Reform Project of Higher Education Institutions in Guizhou Province (2019202, 2020035).

Acknowledgments

Special thanks to Professor Chen Qingfu of Guizhou Normal University for his outstanding contribution in the early creation of perennial germplasm! Thanks to Ms. Li Bin, Institute of Botany, Chinese Academy of Sciences, for her help in analyzing the results and writing the article! All authors state that this study complies with the Chinese legislation and the Convention on the Trade in Endangered Species of Wild Fauna and Flora.

Conflict of interest

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

Publisher's note

References

Alengebawy, A., Abdelkhalek, S. T., Qureshi, S. R., and Wang, M. Q. (2021). Heavy metals and pesticides toxicity in agricultural soil and plants: Ecological risks and human health implications. *Toxics.* 9 (3), 42. doi: 10.3390/toxics9030042

Anders, S., and Huber, W. (2012). *Differential expression of RNA-seq data at the gene level the DESeq package* Vol. 10 (Heidelberg, Germany: European Molecular Biology Laboratory (EMBL), f1000 research.

Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq: a Python framework to work with high-throughput sequencing data. *Bioinformatics.* 31 (2), 166–169. doi: 10.1093/bioinformatics/btu638

Baghayeri, M., Amiri, A., Maleki, B., Alizadeh, Z., and Reiser, O. (2018). A simple approach for simultaneous detection of cadmium (II) and lead (II) based on glutathione coated magnetic nanoparticles as a highly selective electrochemical probe. *Sensors Actuators B: Chemical.* 273, 1442–1450. doi: 10.1016/j.snb.2018.07.063

Bhattacharya, S. (2022). Protective role of the essential trace elements in the obviation of cadmium toxicity: glimpses of mechanisms. *Biol. Trace Element Res.* 200 (5), 2239–2246. doi: 10.1007/s12011-021-02827-7

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics* 30 (15), 2114–2120. doi: 10.1093/bioinformatics/btu170

Borges, K. L. R., Hippler, F. W. R., Carvalho, M. E. A., Nalin, R. S., Matias, F. I., and Azevedo, R. A. (2019). Nutritional status and root morphology of tomato under Cd-induced stress: comparing contrasting genotypes for metal-tolerance. *Scientia horticulturae.* 246, 518–527. doi: 10.1016/j.scienta.2018.11.023

Chen, H., Li, Y., Ma, X., Guo, L., He, Y., Ren, Z., et al. (2019a). Analysis of potential strategies for cadmium stress tolerance revealed by transcriptome analysis of upland cotton. *Sci. Rep.* 9 (1), 1–13. doi: 10.1038/s41598-018-36228-z

Chen, L., Wu, Q., He, W., He, T., Wu, Q., and Miao, Y. (2019b). Combined *de novo* transcriptome and metabolome analysis of common bean response to *Fusarium oxysporum* f. sp. *phaseoli* infection. *Int. J. Mol. Sci.* 20(24), 6278. doi: 10.3390/ijms20246278

Clemens, S. (2006). Toxic metal accumulation, responses to exposure and mechanisms of tolerance in plants. *Biochimie.* 88 (11), 1707–1719. doi: 10.1016/j.biochi.2006.07.003

Dat, J., Vandenabeele, S., Vranova, E. V. M. M., Van Montagu, M., and Van Breusegem, F. (2000). Dual action of the active oxygen species during plant stress responses. *Cell. Mol. Life Sci.* 57 (5), 779–795. doi: 10.1007/s000180050041

Dossa, K., Mmadi, M. A., Zhou, R., Zhang, T., Su, R., Zhang, Y., et al. (2019). Depicting the core transcriptome modulating multiple abiotic stresses responses in sesame (*Sesamum indicum* L.). *Int. J. Mol. Sci.* 20(16), 3930. doi: 10.3390/ijms20163930

Du, D., Jin, R., Guo, J., and Zhang, F. (2019). Infection of embryonic callus with *Agrobacterium* enables high-speed transformation of maize. *Int. J. Mol. Sci.* 20 (2), 279. doi: 10.3390/ijms20020279

Feng, Z., Bartholomew, E. S., Liu, Z., Cui, Y., Dong, Y., Li, S., et al. (2021). Glandular trichomes: new focus on horticultural crops. *Horticulture Res.* 8. doi: 10.1038/s41438-021-00592-1

Feng, Z., Ding, C., Li, W., Wang, D., and Cui, D. (2020). Applications of metabolomics in the research of soybean plant under abiotic stress. *Food Chem.* 310, 125914. doi: 10.1016/j.foodchem.2019.125914

Feng, J., Jia, W., Lv, S., Bao, H., Miao, F., Zhang, X., et al. (2018). Comparative transcriptome combined with morpho-physiological analyses revealed key factors for differential cadmium accumulation in two contrasting sweet sorghum genotypes. *Plant Biotechnol. J.* 16 (2), 558–571. doi: 10.1111/pbi.12795

Gevi, F., Fanelli, G., Zolla, L., and Rinalducci, S. (2019). "Untargeted metabolomics of plant leaf tissues," in *High-throughput metabolomics* (New York, NY: Humana), 187–195). doi: 10.1007/978-1-4939-9236-2_12

Gong, Q., Li, P., Ma, S., Indu Rupassara, S., and Bohnert, H. J. (2005). Salinity stress adaptation competence in the extremophile thellungiella halophila in comparison with its relative *Arabidopsis thaliana*. *Plant J.* 44 (5), 826–839. doi: 10.1111/j.1365-313X.2005.02587.x

Guo, Y., Gao, C., Wang, M., Fu, F. F., El-Kassaby, Y. A., Wang, T., et al. (2020). Metabolome and transcriptome analyses reveal flavonoids biosynthesis differences in ginkgo biloba associated with environmental conditions. *Ind. Crops Products.* 158, 112963. doi: 10.1016/j.indcrop.2020.112963

Hayat, K., Khan, J., Khan, A., Ullah, S., Ali, S., and Fu, Y. (2021). Ameliorative effects of exogenous proline on photosynthetic attributes, nutrients uptake, and oxidative stresses under cadmium in pigeon pea (*Cajanus cajan* L.). *Plants*. 10 (4), 796. doi: 10.3390/plants10040796

He, J., Li, H., Luo, J., Ma, C., Li, S., Qu, L., et al. (2013). A transcriptomic network underlies microstructural and physiological responses to cadmium in *Populus* × *canescens*. *Plant Physiol.* 162 (1), 424–439. doi: 10.1104/pp.113.215681

Hirai, M. Y., and Saito, K. (2004). Post-genomics approaches for the elucidation of plant adaptive mechanisms to sulphur deficiency. *J. Exp. Botany.* 55 (404), 1871–1879. doi: 10.1093/jxb/erh184

Hoseini, S. M., and Zargari, F. (2013). Cadmium in plants: a review. *Int. J. Farming Allied Sci.* 2 (17), 579–581. doi: 10.1016/S0269-7491(97)00110-3

Hossain, Z., and Komatsu, S. (2013). Contribution of proteomic studies towards understanding plant heavy metal stress response. *Front. Plant Science.* 3. doi: 10.3389/fpls.2012.00310

Jiang, Y., Jiang, S., Li, Z., Yan, X., Qin, Z., and Huang, R. (2019). Field scale remediation of cd and Pb contaminated paddy soil using three mulberry (*Morus alba* L.) cultivars. *Ecol. Engineering.* 129, 38–44. doi: 10.1016/j.ecoleng.2019.01.009

Kang, W. H., Sim, Y. M., Koo, N., Nam, J. Y., Lee, J., Kim, N., et al. (2020). Transcriptome profiling of abiotic responses to heat, cold, salt, and osmotic stress of *Capsicum annuum* L. *Sci. Data.* 7 (1), 1–7. doi: 10.1038/s41597-020-0352-7

Kaya, C., Ashraf, M., Alyemeni, M. N., and Ahmad, P. (2020). Responses of nitric oxide and hydrogen sulfide in regulating oxidative defence system in wheat plants grown under cadmium stress. *Physiologia plantarum.* 168 (2), 345–360. doi: 10.1111/ppl.13012

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12 (4), 357–360. doi: 10.1038/nmeth.3317

Komori, A., Naoto, I., Kaori, F., Shinya, K., and Asakazu, H. (2007). Measurement of rutin and quercetin in tartary buckwheat flour by ultraviolet-induced fluorescence. *In Proc. 10th Int. symposium buckwheat*, 401–409).

Kumar, M., Kumar Patel, M., Kumar, N., Bajpai, A. B., and Siddique, K. H. (2021). Metabolomics and molecular approaches reveal drought stress tolerance in plants. *Int. J. Mol. Sci.* 22(17), 9108. doi: 10.3390/ijms22179108

Li, C., Liu, Y., Tian, J., Zhu, Y., and Fan, J. (2020). Changes in sucrose metabolism in maize varieties with different cadmium sensitivities under cadmium stress. *PloS One* 15 (12), e0243835. doi: 10.1371/journal.pone.0243835

Li, Y., Sun, M., He, W., Wang, H., Pan, H., Yang, Q., et al. (2021a). Effect of phosphorus supplementation on growth, nutrient uptake, physiological responses, and cadmium absorption by tall fescue (*Festuca arundinacea schreb.*) exposed to cadmium. *Ecotoxicology Environ. Safety.* 213, 112021. doi: 10.1016/j.ecoenv.2021.112021

Li, L., Yan, X., Li, J., and Tian, Y. (2021b). Physiological and *FtCHS* gene expression responses to PEG-simulated drought and cadmium stresses in tartary buckwheat seedlings. *J. Plant Growth Regulation.*, 1–12. doi: 10.1007/s00344-021-10530-z

Li, C., Zhou, K., Qin, W., Tian, C., Qi, M., Yan, X., et al. (2019). A review on heavy metals contamination in soil: effects, sources, and remediation techniques. *Soil Sediment Contamination: Int. J.* 28 (4), 380–394. doi: 10.1080/15320383.2019.1592108

Lin, X., Li, S., Wei, Z., Chen, Y., Hei, L., and Wu, Q. T. (2021). Indirect application of sludge for recycling in agriculture to minimize heavy metal contamination of soil. *Resources Conserv. Recycling.* 166, 105358. doi: 10.1016/j.resconrec.2020.105358

Liu, C., Guttieri, M. J., Waters, B. M., Eskridge, K. M., Easterly, A., and Baenziger, P. S. (2018). Cadmium concentration in terminal tissues as tools to select low-cadmium wheat. *Plant Soil.* 430 (1), 127–138. doi: 10.1007/s11104-018-3712-8

Mao, X., Cai, T., Olyarchuk, J. G., and Wei, L. (2005). Automated genome annotation and pathway identification using the KEGG orthology (KO) as a controlled vocabulary. *Bioinformatics*. 21 (19), 3787–3793. doi: 10.1093/bioinformatics/bti430

Mei, X., Li, S., Li, Q., Yang, Y., Luo, X., He, B., et al. (2014). Sodium chloride salinity reduces cd uptake by edible amaranth (*Amaranthus mangostanus* L.) *via* competition for Ca channels. *Ecotoxicology Environ. safety.* 105, 59–64. doi: 10.1016/j.ecoenv.2014.04.005

Meng, Y., Jing, H., Huang, J., Shen, R. , and Zhu, X. (2022). The role of nitric oxide signaling in plant responses to cadmium stress. *Int. J. Mol. Sci.* 23(13), 6901. doi: 10.3390/ijms23136901

Milner, M. J., Mitani-Ueno, N., Yamaji, N., Yokosho, K., Craft, E., Fei, Z., et al. (2014). Root and shoot transcriptome analysis of two ecotypes of *Noccaea caerulescens* uncovers the role of *NcNramp1* in Cd hyperaccumulation. *Plant J.* 78 (3), 398–410. doi: 10.1111/tpj.12480

Mitrus, J., and Horbowicz, M. (2020). Impact of short-term exposure to lead and cadmium of common buckwheat (*Fagopyrum esculentum moench*) seedlings grown in hydroponic culture. *J. Elementology.* 25 (2), 633–644. doi: 10.5601/jelem.2019.24.4.1908

Mostofa, M. G., Rahman, A., Ansary, M. D., Uddin, M., Watanabe, A., Fujita, M., et al. (2015). Hydrogen sulfide modulates cadmium-induced physiological and biochemical responses to alleviate cadmium toxicity in rice. *Sci. Rep.* 5 (1), 1–17. doi: 10.1038/srep14078

Mwamba, T. M., Islam, F., Ali, B., Lwalaba, J. L. W., Gill, R. A., Zhang, F., et al. (2020). Comparative metabolomic responses of low-and high-cadmium accumulating genotypes reveal the cadmium adaptive mechanism in *Brassica napus*. *Chemosphere.* 250, 126308. doi: 10.1016/j.chemosphere.2020.126308

Naeem, A., Zafar, M., Khalid, H., Zia-ur-Rehman, M., Ahmad, Z., Ayub, M. A., et al. (2019). "Cadmium-induced imbalance in nutrient and water uptake by plants", in *Cadmium Toxicity and Tolerance in Plants: From Physiology to Remediation.* pp, 299–326. doi: 10.1016/B978-0-12-814864-8.00012-7

Nephali, L., Piater, L. A., Dubery, I. A., Patterson, V., Huyser, J., Burgess, K., et al. (2020). Biostimulants for plant growth and mitigation of abiotic stresses: A metabolomics perspective. *Metabolites.* 10 (12), 505. doi: 10.3390/metabo10120505

Peng, H., He, X., Gao, J., Ma, H., Zhang, Z., Shen, Y., et al. (2015). Transcriptomic changes during maize roots development responsive to cadmium (Cd) pollution using comparative RNAseq-based approach. *Biochem. Biophys. Res. Commun.* 464 (4), 1040–1047. doi: 10.1016/j.bbrc.2015.07.064

Pérez-Patricio, M., Camas-Anzueto, J. L., Sanchez-Alegría, A., Aguilar-González, A., Gutiérrez-Miceli, F., Escobar-Gómez, E., et al. (2018). Optical method for estimating the chlorophyll contents in plant leaves. *Sensors.* 18 (2), 650. doi: 10.3390/s18020650

Peter-Katalinić, J. (2005). Methods in enzymology: O-glycosylation of proteins. *Methods Enzymology.* 405, 139–171. doi: 10.1016/S0076-6879(05)05007-X

Punia, H., Tokas, J., Malik, A., Bajguz, A., El-Sheikh, M. A., and Ahmad, P. (2021). Ascorbate–glutathione oxidant scavengers, metabolome analysis and adaptation mechanisms of ion exclusion in sorghum under salt stress. *Int. J. Mol. Sci.* 22 (24), 13249. doi: 10.3390/ijms222413249

Qin, X., Xia, Y., Hu, C., Yu, M., Shabala, S., Wu, S., et al. (2021). Ionomics analysis provides new insights into the co-enrichment of cadmium and zinc in wheat grains. *Ecotoxicology Environ. Safety.* 223, 112623. doi: 10.1016/j.ecoenv.2021.112623

Rahman, Z., and Singh, V. P. (2019). The relative impact of toxic heavy metals (THMs) (arsenic (As), cadmium (Cd), chromium (Cr)(VI), mercury (Hg), and lead (Pb)) on the total environment: an overview. *Environ. Monit. assessment.* 191 (7), 1–21. doi: 10.1007/s10661-019-7528-7

Ramalingam, A. P., Mohanavel, W., Premnath, A., Muthurajan, R., Prasad, P. V., and Perumal, R. (2021). Large-scale non-targeted metabolomics reveals antioxidant, nutraceutical and therapeutic potentials of sorghum. *Antioxidants.* 10(10), 1511. doi: 10.3390/antiox10101511

Ran, X., Liu, R., Xu, S., Bai, F., Xu, J., Yang, Y., et al. (2015). Assessment of growth rate, chlorophyll a fluorescence, lipid peroxidation and antioxidant enzyme activity in *Aphanizomenon flos-aquae*, pediastrum simplex and synedra acus exposed to cadmium. *Ecotoxicology.* 24 (2), 468–477. doi: 10.1007/s10646-014-1395-3

Rehman, A. U., Nazir, S., Irshad, R., Tahir, K., Islam, R. U., and Wahab, Z. (2021). Toxicity of heavy metals in plants and animals and their uptake by magnetic iron oxide nanoparticles. *J. Mol. Liquids.* 321, 114455. doi: 10.1016/j.molliq.2020.114455

Riaz, M., Kamran, M., Fang, Y., Yang, G., Rizwan, M., Ali, S., et al. (2021). Boron supply alleviates cadmium toxicity in rice (*Oryza sativa* L.) by enhancing cadmium adsorption on cell wall and triggering antioxidant defense system in roots. *Chemosphere.* 266, 128938. doi: 10.1016/j.chemosphere.2020.128938

Romero-Puertas, M. C., Palma, J. M., Gómez, M., Del Rio, L. A., and Sandalio, L. M. (2002). Cadmium causes the oxidative modification of proteins in pea plants. *Plant Cell Environment.* 25 (5), 677–686. doi: 10.1046/j.1365-3040.2002.00850.x

Rui, H., Zhang, X., Shinwari, K. I., Zheng, L., and Shen, Z. (2018). Comparative transcriptomic analysis of two *Vicia sativa* L. varieties with contrasting responses to cadmium stress reveals the important role of metal transporters in cadmium tolerance. *Plant soil.* 423 (1), 241–255. doi: 10.1007/s11104-017-3501-9

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 (2), 277. doi: 10.3390/antiox10020277

Sarwar, N., Malhi, S. S., Zia, M. H., Naeem, A., Bibi, S., and Farid, G. (2010). Role of mineral nutrition in minimizing cadmium accumulation by plants. *J. Sci. Food Agriculture.* 90 (6), 925–937. doi: 10.1002/jsfa.3916

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

Shu, X., Zhang, K., Zhang, Q., and Wang, W. (2019). Ecophysiological responses of jatropha curcas l. seedlings to simulated acid rain under different soil types. *Ecotoxicology Environ. Safety.* 185, 109705. doi: 10.1016/j.ecoenv.2019.109705

Stewart, R. R., and Bewley, J. D. (1980). Lipid peroxidation associated with accelerated aging of soybean axes. *Plant Physiol.* 65 (2), 245–248. doi: 10.1104/pp.65.2.245

Sung, J., Lee, S., Lee, Y., Ha, S., Song, B., Kim, T., et al. (2015). Metabolomic profiling from leaves and roots of tomato (*Solanum lycopersicum* L.) plants grown under nitrogen, phosphorus or potassium-deficient condition. *Plant Science.* 241, 55–64. doi: 10.1016/j.plantsci.2015.09.027

Sytar, O., Kosyan, A., Taran, N., and Smetanska, I. (2014). Anthocyanin's as marker for selection of buckwheat plants with high rutin content. *Gesunde Pflanzen.* 66 (4), 165–169. doi: 10.1007/s10343-014-0331-z

Szymański, J., Bocobza, S., Panda, S., Sonawane, P., Cárdenas, P. D., Lashbrooke, J., et al. (2020). Analysis of wild tomato introgression lines elucidates the genetic basis of transcriptome and metabolome variation underlying fruit traits and pathogen response. *Nat. Genet.* 52 (10), 1111–1121. doi: 10.1038/s41588-020-0690-6

Tang, F., Lao, K., and Surani, M. A. (2011). Development and applications of single-cell transcriptome analysis. *Nat. Methods* 8 (4), S6–S11. doi: 10.1038/nmeth.1557

Thevenot, E. A. (2016). Ropls: PCA, PLS (-DA) and OPLS (-DA) for multivariate analysis and feature selection of omics data. *R Package version.* 1 (0).

Tiryakioglu, M., Eker, S., Ozkutlu, F., Husted, S., and Cakmak, I. (2006). Antioxidant defense system and cadmium uptake in barley genotypes differing in cadmium tolerance. *J. Trace Elements Med. Biol.* 20 (3), 181–189. doi: 10.1016/j.jtemb.2005.12.004

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 (5), 511–515. doi: 10.1038/nbt.1621

Uraguchi, S., Kamiya, T., Sakamoto, T., Kasai, K., Sato, Y., Nagamura, Y., et al. (2011). Low-affinity cation transporter (*OsLCT1*) regulates cadmium transport into rice grains. *Proc. Natl. Acad. Sci.* 108 (52), 20959–20964. doi: 10.1073/pnas.1116531110

Wang, P., Chen, H., Kopittke, P. M., and Zhao, F. J. (2019). Cadmium contamination in agricultural soils of China and the impact on food safety. *Environ. pollution.* 249, 1038–1048. doi: 10.1016/j.envpol.2019.03.063

Wang, W., Pang, J., Zhang, F., Sun, L., Yang, L., Zhao, Y., et al. (2021). Integrated transcriptomics and metabolomics analysis to characterize alkali stress responses in canola (*Brassica napus* L.). *Plant Physiol. Biochem.* 166, 605–620. doi: 10.1016/j.plaphy.2021.06.021

Weckwerth, W. (2003). Metabolomics in systems biology. *Annu. Rev. Plant Biol.* 54 (1), 669–689. doi: 10.1146/annurev.arplant.54.031902.135014

Wei, W., Peng, H., Xie, Y., Wang, X., Huang, R., Chen, H., et al. (2021). The role of silicon in cadmium alleviation by rice root cell wall retention and vacuole compartmentalization under different durations of cd exposure. *Ecotoxicology Environ. Safety.* 226, 112810. doi: 10.1016/j.ecoenv.2021.112810

Wu, Y., Yang, H., Yang, H., Zhang, C., Lyu, L., Li, W., et al. (2022). A physiological and metabolomic analysis reveals the effect of shading intensity on blueberry fruit quality. *Food Chem.* 15, 100367. doi: 10.1016/j.fochx.2022.100367

Xiao, Y., Du, Z., Qi, X., Wu, H., Guo, W., and Zhao, Z. (2019). RNA-Sequencing analysis reveals transcriptional changes in the roots of low-cadmium-accumulating winter wheat under cadmium stress. *Acta Physiologiae Plantarum.* 41 (1), 1–10. doi: 10.1007/s11738-019-2808-y

Xu, L., Wang, Y., Liu, W., Wang, J., Zhu, X., Zhang, K., et al. (2015). *De novo* sequencing of root transcriptome reveals complex cadmium-responsive regulatory networks in radish (*Raphanus sativus* L.). *Plant Science.* 236, 313–323. doi: 10.1016/j.plantsci.2015.04.015

Yang, S., Zu, Y., Li, B., Bi, Y., Jia, L., He, Y., et al. (2019). Response and intraspecific differences in nitrogen metabolism of alfalfa (*Medicago sativa* L.) under cadmium stress. *Chemosphere.* 220, 69–76. doi: 10.1016/j.chemosphere.2018.12.101

Yue, R., Lu, C., Qi, J., Han, X., Yan, S., Guo, S., et al. (2016). Transcriptome analysis of cadmium-treated roots in maize (*Zea mays* L.). *Front. Plant Science.* 71298. doi: 10.3389/fpls.2016.01298

Zeng, T., Fang, B., Huang, F., Dai, L., Tang, Z., Tian, J., et al. (2021). Mass spectrometry-based metabolomics investigation on two different indica rice grains (*Oryza sativa* L.) under cadmium stress. *Food Chem.* 343, 128472. doi: 10.1016/j.foodchem.2020.128472

Zhang, Y., Ali, U., Zhang, G., Yu, L., Fang, S., Iqbal, S., et al. (2019). Transcriptome analysis reveals genes commonly responding to multiple abiotic stresses in rapeseed. *Mol. Breeding.* 39 (10), 1–19. doi: 10.1007/s11032-019-1052-x

Zhang, H., Heal, K., Zhu, X., Tigabu, M., Xue, Y., and Zhou, C. (2021a). Tolerance and detoxification mechanisms to cadmium stress by hyperaccumulator erigeron annuus include molecule synthesis in root exudate. *Ecotoxicology Environ. Safety.* 219, 112359. doi: 10.1016/j.ecoenv.2021.112359

Zhang, L., Li, X., Ma, B., Gao, Q., Du, H., Han, Y., et al. (2017). The tartary buckwheat genome provides insights into rutin biosynthesis and abiotic stress tolerance. *Mol. Plant* 10 (9), 1224–1237. doi: 10.1016/j.molp.2017.08.013

Zhang, J., Zhang, C., Huang, S., Chang, L., Li, J., Tang, H., et al. (2021b). Key cannabis salt-responsive genes and pathways revealed by comparative transcriptome and physiological analyses of contrasting varieties. *Agronomy* 11 (11), 2338. doi: 10.3390/agronomy11112338

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 20 March 2023
DOI 10.3389/fpls.2023.1155018

OPEN ACCESS

EDITED BY
Meng Kou,
Sweet Potato Research Institute (CAAS), China

REVIEWED BY
Shaopei Gao,
China Agricultural University, China
Qi Wu,
Chengdu University, China

*CORRESPONDENCE
Zengzhi Si
sizengzhi@163.com

SPECIALTY SECTION
This article was submitted to Crop and Product Physiology, a section of the journal Frontiers in Plant Science

RECEIVED 31 January 2023
ACCEPTED 06 March 2023
PUBLISHED 20 March 2023

CITATION
Si Z, Wang L, Ji Z, Zhao M, Zhang K and Qiao Y (2023) Comparative analysis of the *MYB* gene family in seven *Ipomoea* species.
Front. Plant Sci. 14:1155018.
doi: 10.3389/fpls.2023.1155018

Comparative analysis of the *MYB* gene family in seven *Ipomoea* species

Zengzhi Si[1]*, Lianjun Wang[2], Zhixin Ji[1], Mingming Zhao[1], Kai Zhang[1] and Yake Qiao[1]

[1]Hebei Key Laboratory of Crop Stress Biology, Hebei Normal University of Science and Technology, Qinghuangdao, Hebei, China, [2]Institute of Food Corps, Hubei Academy of Agricultural Sciences, Wuhan, Hubei, China

The MYB transcription factors regulate plant growth, development, and defense responses. However, information about the *MYB* gene family in *Ipomoea* species is rare. Herein, we performed a comprehensive genome-wide comparative analysis of this gene family among seven *Ipomoea* species, sweet potato (*I. batatas*), *I. trifida*, *I. triloba*, *I. nil*, *I. purpurea*, *I. cairica*, and *I. aquatic*, and identified 296, 430, 411, 291, 226, 281, and 277 *MYB* genes, respectively. The identified *MYB* genes were classified into five types: *1R-MYB* (*MYB*-related), *2R-MYB* (*R2R3-MYB*), *3R-MYB* (*R1R2R3-MYB*), *4R-MYB*, and *5R-MYB*, and the *MYB-related* or *R2R3-MYB* type was the most abundant *MYB* genes in the seven species. The *Ipomoea MYB* genes were classed into distinct subgroups based on the phylogenetic topology and the classification of the *MYB* superfamily in Arabidopsis. Analysis of gene structure and protein motifs revealed that members within the same phylogenetic group presented similar exon/intron and motif organization. The identified *MYB* genes were unevenly mapped on the chromosomes of each *Ipomoea* species. Duplication analysis indicated that segmental and tandem duplications contribute to expanding the *Ipomoea MYB* genes. Non-synonymous substitution (Ka) to synonymous substitution (Ks) [Ka/Ks] analysis showed that the duplicated *Ipomoea MYB* genes are mainly under purifying selection. Numerous *cis*-regulatory elements related to stress responses were detected in the *MYB* promoters. Six sweet potato transcriptome datasets referring to abiotic and biotic stresses were analyzed, and *MYB* different expression genes' (DEGs') responses to stress treatments were detected. Moreover, 10 sweet potato *MYB* DEGs were selected for qRT-PCR analysis. The results revealed that four responded to biotic stress (stem nematodes and *Ceratocystis fimbriata* pathogen infection) and six responded to the biotic stress (cold, drought, and salt). The results may provide new insights into the evolution of *MYB* genes in the *Ipomoea* genome and contribute to the future molecular breeding of sweet potatoes.

KEYWORDS

***Ipomoea* species, MYB gene family, phylogenetic analysis, duplication analysis, gene expression, stresses response, sweet potato**

1 Introduction

The MYB transcription factor (TF) family, as one of the most prominent TF families in the plant, is composed of three conservative functional domains: DNA-binding domain (DBD), transcriptional activation domain (TAD), and incompletely defied negative regulatory region (NRD). Among the three, DBD is the most conservative and is generally known as the MYB domain (Frampton, 2004). The MYB domain consists of 51–52 conserved amino acids and spacing coding sequences, and every 18 amino acids has a regular spacing of tryptophan residues. This amino acid residue expands the MYB domain into a helix (HTH) structure (Frampton, 2004). According to the total number of adjacent MYB repetitions, MYB TF can be divided into four categories, namely, 1R-MYB (MYB-related), 2R-MYB, 3R-MYB, and 4R-MYB, which contain the corresponding number of MYB repeats, respectively (Dubos et al., 2010).

1R-MYB, also called MYB-related, generally but not always include a single or partial MYB iteration (Klempnauer et al., 1982). Based on the highly conserved motif, the MYB-related TFs can be divided into subgroups (Liu et al., 2020b; Arce-Rodríguez et al., 2021; Yan et al., 2022). The MYB-related TFs play essential roles in plant seed germination and root growth (Zhao et al., 2019), leaf senescence and epidermal cell patterning (Lee and Schiefelbein, 1999; Machado et al., 2009; Park et al., 2018), biosynthesis (Sinaga et al., 2021), the regulation of circadian rhythms (Alabadí et al., 2001; Mizoguchi et al., 2002; Okada et al., 2009; Kamioka et al., 2016), and stress responses (Dubos et al., 2010; Park et al., 2018; Zhao et al., 2019; Hoseinpour et al., 2021).

The *2R-MYB* genes, named R2R3-MYB, were usually the largest *MYB* gene subfamily in green plants (Martin and Paz-Ares, 1997). They play essential roles in plant lifetime, including (1) primary and secondary metabolites, such as phenylpropanoid metabolism (Liu et al., 2015; Tang et al., 2021), flavonoid biosynthesis (Czemmel et al., 2012; Lin et al., 2021; Zhao et al., 2022b), anthocyanin biosynthesis (Naing and Kim, 2018; Upadhyaya et al., 2021; Wang et al., 2022c), chlorogenic acid biosynthesis (Tang et al., 2021), indolic and aliphatic glucosinolate biosynthesis (Gigolashvili et al., 2009), and apigenin biosynthesis (Wang et al., 2022b); (2) cell fate and identity, for instance, epidermis and root formation (Du et al., 2009); (3) developmental processes, like embryogenesis (Wang et al., 2009), pollen development (Liu et al., 2021), male sterility (Yu et al., 2021), and plant trichome development (Shangguan et al., 2021); and (4) responses to biotic and abiotic stresses, including diseases (Shen et al., 2017; Gu et al., 2020; Hawku et al., 2022), high salinity (Wang et al., 2021a; Du et al., 2022), cold (Dong et al., 2021), heat (Wu et al., 2021b), and drought (Lv et al., 2021; Wu et al., 2021a).

The 3R-MYB TFs in plants recognize mitosis-specific activator (MSA) elements (Ito et al., 1998; Ma et al., 2009) and play an essential role in cell-cycle regulation. The plant 3R-MYB TFs regulate the G2/M transition (Ito et al., 2001). The recognized DNA element (MSA) was in the upstream promoter region of G2/M-phase-specific genes, which is both necessary and sufficient for driving G2/M-phase-specific gene expression (Ito et al., 2001; Kato et al., 2009). The plant 3R-MYB TFs were often classed into three subgroups: A, B, and C (Ito, 2005). The 3R-MYB TFs in the A and B subgroups were reported to be involved in cell-cycle regulation (Ito et al., 2001; Kato et al., 2009; Haga et al., 2011), and the ones in the C group participated in both cell-cycle and abiotic stresses (Dai et al., 2007; Ma et al., 2009). It was also found that 3R-MYB inhibited plant trichome development and inhibited flavonoid biosynthetically (Dubos et al., 2008; Matsui et al., 2008). The 3R-MYB gene, *PhMYBx* in petunia, downregulated anthocyanin synthesis (Koes et al., 2005; Chezem and Clay, 2016). The 4R-MYB TFs, representing the smallest subfamily of MYB TFs, contain four R1/R2-like repeats. Little is known about their functions in plants. Recently, the 4R-MYB protein SNAPc4 in Arabidopsis was reported as part of the SNAP complex involved in snRNA gene transcription, and *AtSNAPc4* is proved to be an essential gene in gametophyte and zygote development (Thiedig et al., 2021).

Since the first plant *MYB* gene (*ZmMYBC1*) was cloned and reported to be involved in pigment biosynthesis in *Zea maize* (Paz-Ares et al., 1987), the characterization and analysis of *MYB* genes have been conducted in different plant species. In most cases, the *2R-MYB* genes were the most abundant, followed by *MYB-related* genes and *3R-MYB* genes, and the *4R-MYB* gene group was the smallest (Dubos et al., 2010). However, the number of *MYB* genes varied widely among species, ranging from less than 100 to more than 500. For instance, 198 *MYB* genes in Arabidopsis (Chen et al., 2006), 183 in *Oryza sativa* (Chen et al., 2006), 475 in *Brassica rapa* ssp. *Pekinensis* (Saha et al., 2016), 524 in *Gossypium hirsutum* (Salih et al., 2016), 174 in *Morella rubra* (Cao et al., 2021), 217 in *Solanum tuberosum* (Li et al., 2021c), 182 in *Casuarina equisetifolia* (Wang et al., 2021b), and 54 in *Mangifera indica* (Zhang et al., 2022) were identified and analyzed. The number of *MYB* genes greatly varied in the species of the same genus. For example, in the *Solanum* genus, 127 *MYB* genes were reported in *S. lycopersicum* (Li et al., 2016) whereas 217 in *S. tuberosum* (Li et al., 2021c); in the *Musa* genus, 305 *MYB* genes were identified in *M. acuminate* whereas 251 in *M. balbisiana* (Tan et al., 2020). To date, information about the comparative analysis of *MYB* genes in closely related species (the same genus) is still limited.

The genus *Ipomoea* (family Convolvulaceae) contains 600–700 species (Nimmakayala et al., 2011), many of which are of considerable importance either as medicinal plants or as ornamental plants (Nimmakayala et al., 2011). The sweet potato (*I. batatas*), serving as the seventh most important crop in the world, is an essential source of calories, proteins, vitamins, and minerals for humanity (Bovell-Benjamin, 2007; Yang et al., 2017). *I. trifida* and *I. triloba*, two diploid relatives of sweet potatoes, can be used as model plants to facilitate sweet potato breeding (Wu et al., 2018). *I. nil* and *I. purpurea* are ideal plants for researching photoperiodic flowering and flower coloration (Hoshino et al., 2016). *I. cairica* (L.) is a perennial vine plant that blooms all year round and has been widely introduced to subtropical, subtropical, and temperate regions as ornamental plants for the landscape. It also has the characteristics of medicinal value because it contains a large number of bioactive compounds. In many countries, a decoction of the whole plant is used to treat tuberculosis, cough, asthma, liver cirrhosis, and jaundice (Jiang et al., 2022). Water spinach (*I. aquatica*) is one of Asia's most popular green leafy

vegetables, with both aquatic and terrestrial characteristics (Hao et al., 2021).

Related information is limited compared to the importance of the *MYB* gene family and *Ipomoea* species, particularly for sweet potatoes. For the sweet potato, the researchers mainly focused on researching a few *2R-MYB* genes in the past years. *IbMYB1* increases the anthocyanin content and the resistance ability of potatoes (Cheng et al., 2013); the *IbMYB1a* gene induces anthocyanin accumulation in Arabidopsis (Chu et al., 2013); and overexpression of the *IbMYB1* gene in an orange-fleshed sweet potato cultivar produces a dual-pigmented transgenic sweet potato with improved antioxidant activity (Park et al., 2015). *IbMYB116* was reported to enhance the drought tolerance of Arabidopsis (Zhou et al., 2019). A single amino acid mutant in the EAR motif of *IbMYB44.2* could reduce the inhibition of anthocyanin accumulation in the purple-fleshed sweet potato (Li et al., 2021a). *IbMYB48* (a sweet potato R2R3-MYB gene) confers enhanced tolerance to salt and drought stresses in transgenic Arabidopsis (Zhao et al., 2022a). *IbMYB308* improves salt stress tolerance in transgenic tobacco (Wang et al., 2022a). However, there is no comparative analysis of the *MYB* gene family of *Ipomoea* species, and knowledge of *MYB* genes of this genus is limited.

This study conducted a genome-wide comparative analysis of the *MYB* gene family in seven *Ipomoea* species. There were 296, 430, 411, 291, 226, 281, and 277 *MYB* genes identified from sweet potato (*I. batatas*), *I. trifida*, *I. triloba*, *I. nil*, *I. purpurea*, *I. cairica*, and *I. aquatic*, respectively. The identified *MYB* genes were then subjected to phylogenetic analysis, gene structure investigation, chromosome location, syntenic analysis, non-synonymous substitution (Ka) to synonymous substitution (Ks) [Ka/Ks] calculation, *cis*-regulatory element (CRE) detection, and expression profile analysis. Then, 10 sweet potato *MYB* different expression genes (DEGs) were selected for qRT-PCR analysis. The results showed that four responded to biotic stress (stem nematodes and *Ceratocystis fimbriata* pathogen infection) and six responded to abiotic stress (cold, drought, and salt). These results are likely to give a new view on the evolution of the *MYB* gene in *Ipomoea* species. They are conducive to the development of molecular breeding in sweet potatoes in the future.

2 Materials and methods

2.1 Data resources

The whole genome sequences and annotated files of the seven *Ipomoea* species were obtained from the following open-access databases: sweet potato from the *I. batatas* Genome Browser (http://public-genomes-ngs.molgen.mpg.de/SweetPotato/), *I. trifida* and *I. triloba* from GenBank BioProject (accessions numbers PRJNA428214 and PRJNA428241), *I. nil* from the *I. nil* web (http://viewer.shigen.info/asagao/index.php), *I. purpurea* from the CoGe platform (https://genomevolution.org/coge/GenomeInfo.pl?gid=58735), *I. cairica* from AGIS (ftp://ftp.agis.org.cn/~fanwei/Ipomoea_cairica_genome_v1), and *I. aquatic* from BIGD (PRJCA002216). The MYB protein sequences of Arabidopsis were downloaded from the TAIR database (https://www.arabidopsis.org) (Chen et al., 2006).

2.2 Identification of *MYB* genes in *Ipomoea* species

To identify the *MYB* genes in *Ipomoea* species, two strategies were adopted. First, the protein sequences of the seven *Ipomoea* species were searched for the MYB domain (Pfam accession number: PF00249) using HMMsearch (ver. 3.1b2) with default parameters. Second, the MYB protein sequences of Arabidopsis were used as queries to a search against the protein database of each *Ipomoea* species by BLASTP (ver. 2.10.0+) with an E-value cutoff of 1e-10 (Altschul et al., 1990). The protein sequences obtained from the two strategies were merged, and the redundant ones were removed. Then, all of the MYB proteins were subjected to a search against the Pfam database (release 35; adopted Pfam A) using PfamScan.pl (v1.6) with default parameters, search against the Conserved Domain Database (CDD) database (ver. 3.20) employing Reverse Position-Specific BLAST (RPS–BLAST) (ver. 2.10.0+) with an e-value cutoff of 1e-10, and search against the Simple Modular Architecture Research Tool (SMART) database (http://smart.embl-heidelberg.de/) by performing SMART_batch.pl (http://smart.embl-heidelberg.de/help/SMART_batch.pl) with default parameters. The proteins confirmed by all three databases were considered candidate MYB TFs.

2.3 Molecular weight, isoelectric point, and subcellular localization analysis of *Ipomoea* MYB proteins

The molecular weight (MW) and isoelectric point (pI) for each MYB protein were analyzed by using ExPASy (http://www.expasy.ch/tools/pi_tool.html). The online website WoLF PSORT (https://wolfpsort.hgc.jp/) was employed for predicting the subcellular localization of the *Ipomoea* proteins.

2.4 Identification of conserved motifs of the *Ipomoea MYB* genes

The MYB protein sequence motif analysis was performed using the online MEME Suite (https://meme-suite.org/meme/) (Bailey et al., 2009). The maximum detecting number of motifs was designed to identify 20 motifs and the site distribution was set as any, whereas other parameters were set as default.

2.5 Sequence alignment and phylogenetic analysis of *Ipomoea* MYB proteins

The *Ipomoea* and Arabidopsis MYB protein sequences were aligned using Clustal Omega (Sievers et al., 2011; Sievers and Higgins, 2018). The obtained aligned sequences were submitted to

IQ-TREE for phylogenetic analysis using the maximum likelihood approach (Nguyen et al., 2015). The branch support values were calculated by SH-aLRT and UFBoot2 with 1,000 bootstrap replicates (Anisimova et al., 2011).

2.6 Protein motif compositions and gene structures of *Ipomoea MYB* genes

Based on the results of the online MEME Suite, phylogenetic analysis, and the gff3 files of the genomes, the *Ipomoea MYB* genes identified above were submitted to TBtools for protein motif compositions and gene structure analysis and picture drawing (Chen et al., 2020).

2.7 Chromosome distribution and duplication pattern analysis of the *Ipomoea MYB* genes

The *MYB* genes with chromosomal positions were mapped on the chromosomes of seven *Ipomoea* species with MapChart (ver. 2.30) (Voorrips, 2002). The potential duplicated *MYB* genes in the *Ipomoea* genome were analyzed with MCScanX software (Wang et al., 2012). During this stage, the protein sequences of *Ipomoea* species were compared against themselves by running the BLASTP (ver. 2.10.0+) program with an E-value of 1e−10 (Altschul et al., 1990). The final output of the duplicated analysis was visualized with the CIRCOS software (ver. 0.66) (Krzywinski et al., 2009).

2.8 Syntenic analysis of *MYB* genes in the seven *Ipomoea* genomes

Syntenic block in the genomes of the seven *Ipomoea* species was analyzed using MCScan software (Python version3) (Tang et al., 2008) with the default parameters (VanBuren et al., 2018; Li et al., 2021b). The gene models were aligned with LAST, and hits were filtered to locate the best 1:1 syntenic blocks (pairs) and were visualized in the dot-plot script using the JCVI package (Tang et al., 2008).

2.9 Ka/Ks analysis of duplicated and syntenic *MYB* genes

Both duplicated and syntenic *MYB* gene pairs of the *Ipomoea* species were selected for the non-synonymous substitution (Ka) to synonymous substitution (Ks) [Ka/Ks] calculation with TBtools (Chen et al., 2020).

2.10 Promoter analysis of *Ipomoea MYB* genes

The 1.5-kb promoter sequences of the *Ipomoea MYB* genes were submitted into PLANT CARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/, accessed on 21 March 2021) for identification of the putative *cis*-elements (Magali et al., 2002).

2.11 Expression profile of *MYB Ipomoea* genes

Six transcriptome bio project datasets referring to abiotic and biotic stresses were selected for expression profile analysis of the sweet potato *MYB* genes. Three bio project datasets (PRJNA341328 for cold, PRJNA413661 for drought, and PRJNA631585 for salt) referring to abiotic stress and one bio project (PRJNA429283 for root-knot nematode) referring to biotic stress were obtained from the NCBI database. The other two were our in-house transcriptome datasets (unpublished) for sweet potato stem nematodes and *C. fimbriata* resistance of four sweet potato cultivars or lines (sweet potato stem nematode-susceptible cultivar, "Tengfei," sweet potato stem nematode-resistant line, "JK20," *C. fimbriata*-susceptible cultivar, "Santiandao," and *C. fimbriata*-resistant line, "JK274"). In each comparing case, reads were treated as DEGs if $|\log2FC|>1$ and FDR $\leq$ 5%. Thus, the mean log2FC value for each DEG was calculated. The heat map was constructed to visualize the distribution of the expression level of genes using the fragments per kilobase per million (FPKM) value in MeV software (Howe et al., 2011). The gene FPKM data of *I. trifida* and *I. triloba* were downloaded from the sweet potato database (http://sweetpotato.plantbiology.msu.edu/).

2.12 RNA isolation and quantitative qRT-PCR analysis

Xu32 (susceptible cultivar) and JK328 (resistant line) were selected for cold, salt, and drought stress treatments. The cuttings about 25 cm in length from 6-week-olds of them grown in a field were cultured in the Hoagland solution for 3 days to survive: for cold stress treatment, the cuttings were then placed at 28°C (control) and 16°C (cold stress), respectively; for salt stress treatment, the cuttings were cultured in the Hoagland solution with 0 and 86 mM NaCl, respectively; and for drought stress treatments, the cuttings were cultured in Hoagland solution with 0% and 30% PEG 6000. Samples were collected at seven time points (0, 2, 4, 6, 12, 24, and 48 h) after the treatments. The cultivars Tengfei (susceptible cultivar) and JK20 (resistant line) were inoculated with sweet potato stem nematodes (Gao et al., 2011), and the cultivars Santiandao (susceptible cultivar) and JK274 (resistant line) were inoculated with *C. fimbriata* (Muramoto et al., 2012). Samples were collected at seven time points (0 h, 6 h, 12 h, 1 day, 2 days, 4 days, and 6 days) after the injection. Root samples without injection were used as a control. Then, the total RNA of the samples was isolated using RNAprep Pure Plant Kit (Tiangen Biotech, Beijing, China) and first-strand cDNA was synthesized by QuantScript Reverse Transcriptase Kit (Tiangen Biotech, Beijing, China). The sweet potato β-actin gene (GenBank AY905538) was used as a control and to normalize the relative quantities of the three individual targeted DEGs based on its consistency across the different time points of each treatment.

Three biological replicates were performed at each time point, and the gene expression changes were calculated using the $2^{-\Delta\Delta Ct}$ method for each sample (Schmittgen and Livak, 2008). Using the primers (Table S1) generated with Primer-BLAST software (Ye et al., 2006), qRT-PCR was performed as described previously (Zhai et al., 2016).

3 Results

3.1 Identification of the *MYB* genes in the seven *Ipomoea* species

A total of 2,212 *MYB* genes were identified from the seven *Ipomoea* species: 296 from sweet potato (*I. batatas*), 430 from *I. trifida*, 411 from *I. triloba*, 291 from *I. nil*, 226 from *I. purpurea*, 281 from *I. cairica*, and 277 from *I. aquatic* (Tables S2, S3). The number of MYB TFs in the seven investigated *Ipomoea* species varied. According to the total number of adjacent MYB repetitions, the identified MYB TFs were classed into five types: MYB-related (1R-MYB), R2R3-MYB (2R-MYB), 3R-MYB, 4R-MYB, and 5R-MYB (Tables S2, S3). Among these types, the MYB-related or R2R3-MYB was the most abundant in the seven *Ipomoea* species. At the same time, the percentages of 3R-MYB, 4R-MYB, and 5R-MYB were comparatively small and no 4R-MYB TFs were detected in *I. nil*, *I. purpurea*, *I. cairica*, and *I. aquatic* (Tables S2, S3). In the sweet potato (*I. batatas*), *I. nil*, *I. cairica*, and *I. aquatic*, the number of R2R3-MYB TFs was more than the MYB-related ones, whereas in *I. trifida*, *I. triloba*, and *I. purpurea*, the MYB-related TFs were more dominant than the R2R3-MYB ones, especially in *I. purpurea*; the proportion of MYB-related TFs (67.26%) was equivalent to 2.17 times of that of R2R3-MYB ones (Tables S2, S3).

3.2 Molecular weight, isoelectric point, and subcellular localization analysis of *Ipomoea* MYB proteins

The average length of the *Ipomoea* MYB TFs was 377.43 amino acids, with an average exon number of 4.41(Tables S2, S3). The average of the calculated molecular weight (MW) of the *Ipomoea* MYB TFs was 41.96 kDa, ranging from 8.66 to 219.26 kDa. In all of the seven *Ipomoea* species, the average length, the average exon number, and the average MW of the *MYB* genes were increased with the increase in the number of adjacent MYB repetitions, except for that of the *R2R3-MYB* genes which was smaller than that of *MYB-related* genes (Tables S2, S3). The predicted isoelectric point (pI) of the *Ipomoea* MYB TFs ranged from 4.15 to 11.23, with an average pI of 7.15. Moreover, the average pI of 4R-MYB and 5R-MYB TFs was larger than that of other types of MYB TFs in the seven *Ipomoea* species (Tables S2, S3). The grand average of hydropathicity of the *Ipomoea* MYB TFs was -0.70, ranging from -1.53 to 0.18 (Tables S2, S3). Subcellular localization analysis showed that the majority of the *Ipomoea* MYB TFs (>90%) were predicted to be localized in the nucleus, with a small set of them predicted to localize in other subcellular locations, such as chloroplast, cytosol, and mitochondrial (Tables S4, S5).

3.3 Conserved motif analysis of the *Ipomoea* MYB proteins

To understand the conserved domains of the *Ipomoea* MYB proteins, the Motif Elicitation (MEME) analysis was performed on each type of MYB TFs identified above. Twenty conserved motifs were found in the *Ipomoea* MYB-related TFs, R2R3-MYB TFs, 3R-MYB TFs, and the other (4R- and 5R-MYB) TFs, respectively (Figure S1). As shown in Figure 1, in addition to the number of adjacent MYB repetitions, the MYB domain of each type of MYB TFs varied. At the same time, many conserved amino acids were also detected, especially for the tryptophan residues (W) (Figure 1). In the R2 and R3 domains of the R2R3-MYB TFs, there were three conserved W in R2 and only two conserved W in R3; the first W of R3 was replaced by phenylalanine (F) in this study (Figure 1).

3.4 Phylogenetic analysis of *Ipomoea* MYB proteins

Phylogenetic trees were constructed with MYB-related TFs, R2R3-MYB TFs, 3R-MYB TFs, and the other (4R- and 5R-MYB) TFs of *Ipomoea* species and Arabidopsis, respectively. Based on the tree's topology, the MYB-related TFs of *Ipomoea* species and Arabidopsis were classed into 27 subgroups, from A1 to A27 (Figure 2). The Arabidopsis MYB-related TFs dispersed in 10 of the 27 subgroups (i.e., A4, A12, A13, A14, A15, A20, A21, A22, A23, and A25) (Figure 2). According to the classification of Arabidopsis MYB-related TFs (Chen et al., 2006), we identified the CCA1-like subfamily (A4 and A13), R-R-type subfamily (A12), I-box-binding-like subfamily (A14 and A25), CPC-like subfamily (A15), TBP-like subfamily (A20, A21and A23), and TRF-like subfamily (A22) (Figure 2). The R2R3-MYB TFs of *Ipomoea* species and Arabidopsis were divided into 46 subgroups, from B1 to B46 (Figure 3). Based on the previous study of Arabidopsis R2R3-MYB TFs (Dubos et al., 2010), the S1 to S25 subfamilies and potential functions of the genes in each subfamily are indicated in Figure 3; for instance, B11 and B23 were indicated as S11 and S2, respectively, and the gene in the two subgroups might be involved in defense function (Figure 3). The 3R-MYB TFs were classed into 10 subgroups (C1 to C10). All of the 3R-MYB TFs of sweet potato, except for *IbMYB275*, were clustered together in subgroup C10 (Figure S2). The 3R-MYB TFs belonging to the same species tended to cluster together, for example, subgroups C3, C4, C7, and C10. In comparison, the genes clustered in the subgroups C2, C6, and C8 were always from different species (Figure S2). The results indicated that the members of *Ipomoea* 3R-MYB TFs might experience different evolutions after the species differentiation: some kept the initial copy number of their ancestors, whereas others increased their copies (Figure S2). The 4R- and 5R-MYB TFs of the seven *Ipomoea* species were divided into four subgroups, named D1 to D4 (Figure S3).

10.3389/fpls.2023.1155018

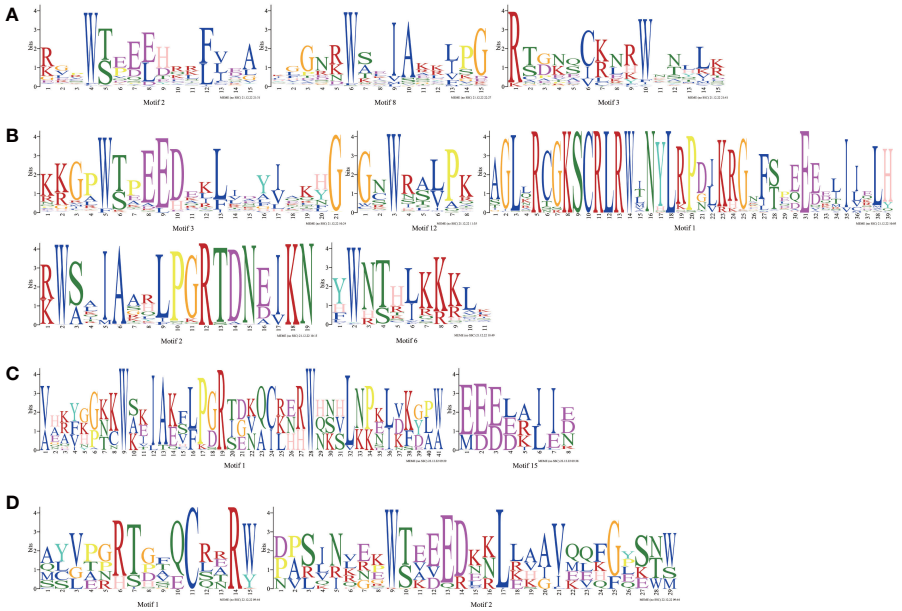


FIGURE 1
MYB repeats of the proteins in *Ipomoea* MYB TFs. **(A)** MYB repeats of *Ipomoea* MYB-related TFs; **(B)** MYB repeats of *Ipomoea* R2R3-MYB TFs; **(C)** MYB repeats of *Ipomoea* 3R-MYB TFs; **(D)** MYB repeats of *Ipomoea* 4R- and 5R-MYB TFs.

3.5 Gene structures and motif composition of *MYB* genes in *Ipomoea* species

As shown in Figure S4, the gene's structural diversity might also be a reflection of the evolutional divergence among homologous *MYB* genes. In other words, the *Ipomoea MYB* genes clustered in the same subgroups shared a roughly similar exon/intron structure, exon number, and gene length (Figure S4). The *MYB-related* genes harbored more exons and longer sequence lengths than R2R3-MYB genes (Figure S4, Table S3). The distributions of motifs of *MYB* proteins in the same subgroup were also roughly similar, suggesting that the motif distributions might be related to functions (Figure S4). The motifs that consist of the core *MYB* domain were the most conservative: for the *MYB-related* genes, motif 2 was the most conserved, followed by motif 3, motif 1, and motif 8; for the *R2R3-MYB* genes, motif 3 was the most conserved, followed by motif 2, motif 6, and motif 12; 19 of the 23 *3R-MYB* genes contained motif 1 and motif 15; and all of the *4R-* and *5R-MYB* genes contained motif 1 and motif 2 (Figure S4, Table S6).

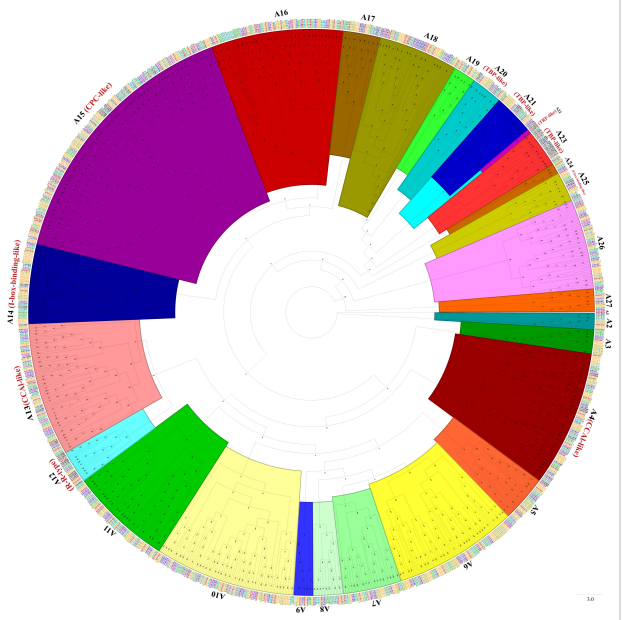


FIGURE 2
The phylogenetic tree of *Ipomoea* species and Arabidopsis MYB-related proteins. The proteins were grouped into 27 subgroups, each given a number (e.g., A1 to A27). The "CCA1-like," "CPC-like," "I-box-binding-like," "R-R-type," "TBP-like," and "TRF-like" subfamilies were indicated with red name, respectively (Chen et al., 2006).

3.6 Chromosomal location and duplication analysis of the *Ipomoea MYB* genes

To understand the genomic distribution of the *Ipomoea MYB* genes, chromosomal location analysis was performed. As shown in Figure 4, the *Ipomoea MYB* genes were unevenly distributed across the chromosomes of the *Ipomoea* species (Figure 4). Duplication analysis showed that 106, 105, and 20 pairs of tandemly duplicated *MYB* genes were found in *I. trifida*, *I. triloba*, and *I. nil*, respectively. In contrast, no duplicated *MYB* genes were detected in sweet potato (*I. batatas*), *I. purpurea*, *I. cairica*, and *I. aquatic* (Figure 4, Table S7). Meanwhile, a total of six, one, one, four, three, three, and seven pairs of segmentally duplicated genes were found in sweet potato (*I. batatas*), *I. trifida*, *I. triloba*, *I. nil*, *I. purpurea*, *I. cairica*, and *I. aquatic*, respectively (Figure 4, Table S7). These results suggested that both segmental and tandem duplications have played essential

10.3389/fpls.2023.1155018

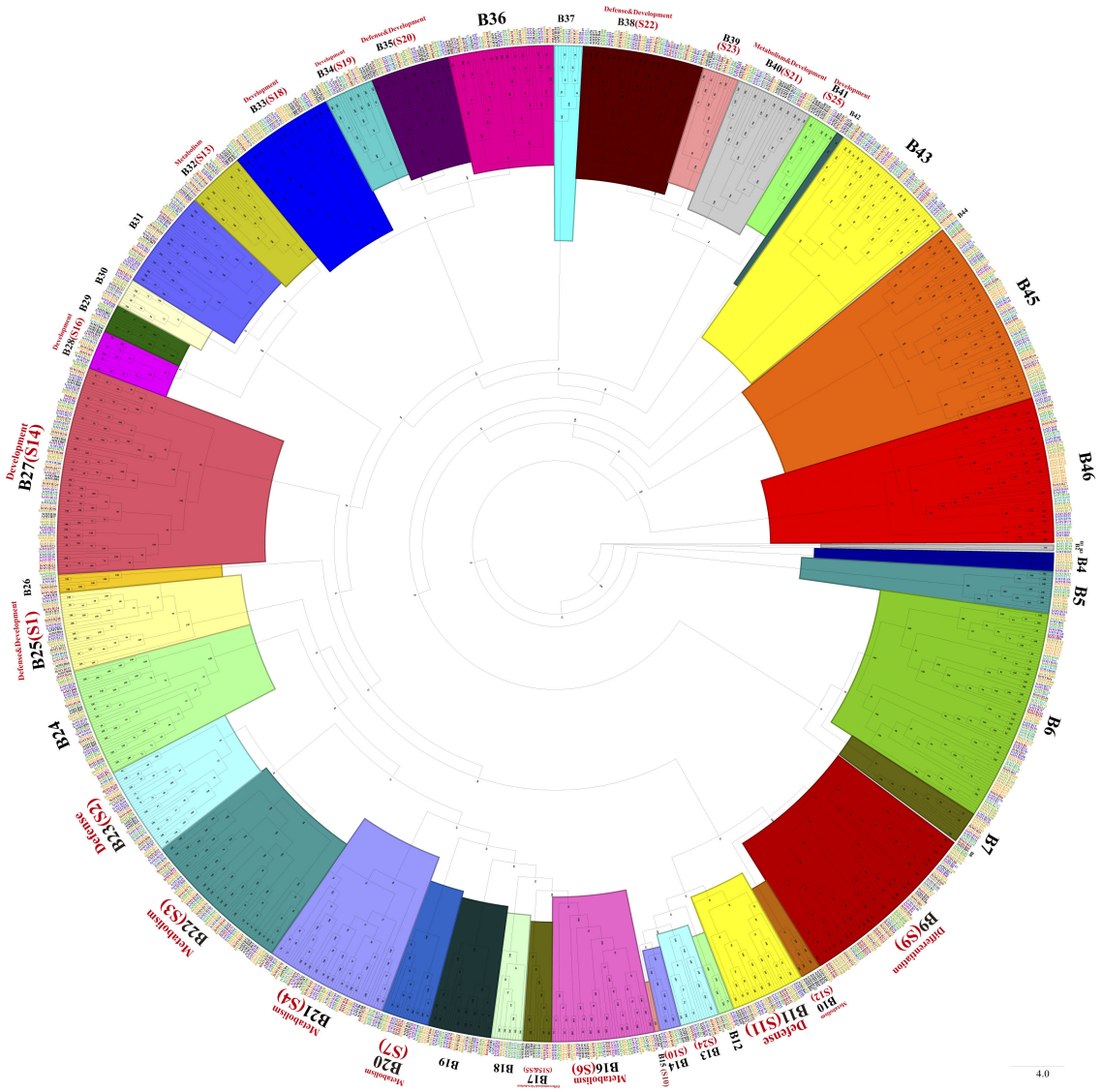


FIGURE 3
The phylogenetic tree of *Ipomoea* species and Arabidopsis R2R3-MYB proteins. The proteins were grouped into 46 subgroups, and each group has given a number (e.g., B1 to B46). S1–S25 indicated the Arabidopsis subfamily identified previously and were written in red words, as well as the prospective functions of the Arabidopsis MYB TFs in each subfamily (Dubos et al., 2010).

roles in *MYB* gene expansion in *I. trifida*, *I. triloba*, and *I. nil*, and tandem duplications were predominant, whereas in sweet potato (*I. batatas*), *I. purpurea*, *I. cairica*, and *I. aquatic*, segmental duplications were predominant.

3.7 Syntenic analysis of *MYB* genes in the genomes of *Ipomoea* species

To determine the evolutionary mechanism of *Ipomoea MYB* genes, comparative synteny maps of the seven *Ipomoea* species were constructed (Figure 5). A total of 1,703 *Ipomoea MYB* genes that formed 4,050 ortholog pairs were detected in the seven *Ipomoea* species (Tables S8, S9). Of these genes, *R2R3-MYB* genes were the most (890), followed by *MYB-relative* (782), *3R-MYB* (24), and *5R-MYB* (7) genes; when referring to subfamilies, the A15 members were the most (115), followed by the A10 (77), A13 (71), A16 (67), B27 (66), A4 (66), B6 (61), and other subfamily (<60) members (Tables S8, S9). *I. trifida* and *I. triloba* harbored the most ortholog *MYB* gene pairs (248 pairs), followed by *I. triloba* and *I. cairica* (240 pairs), *I. nil* and *I. cairica* (223 pairs), and *I. trifida*, and *I. cairica* (221 pairs) (Table S8). The sweet potato showed more than 200 ortholog *MYB* gene pairs with *I. triloba* (214 pairs), *I. cairica* (214 pairs), and *I. trifida* (208 pairs) (Table S8). The ortholog *MYB* genes were distributed in all of the subfamilies except for A1, A22, and D1 (Table S8). In most cases (3,636 of 4,150, 87.61%), two of the ortholog *MYB* genes were from the same subfamily (Table S10). The A13 subfamily held the most *MYB* gene pairs (166 pairs), followed by A10 (161 pairs), A4 (161 pairs), A16 (156 pairs), B6 (145 pairs), and B45 (134 pairs) (Table S10). A total of 621 *MYB*

10.3389/fpls.2023.1155018

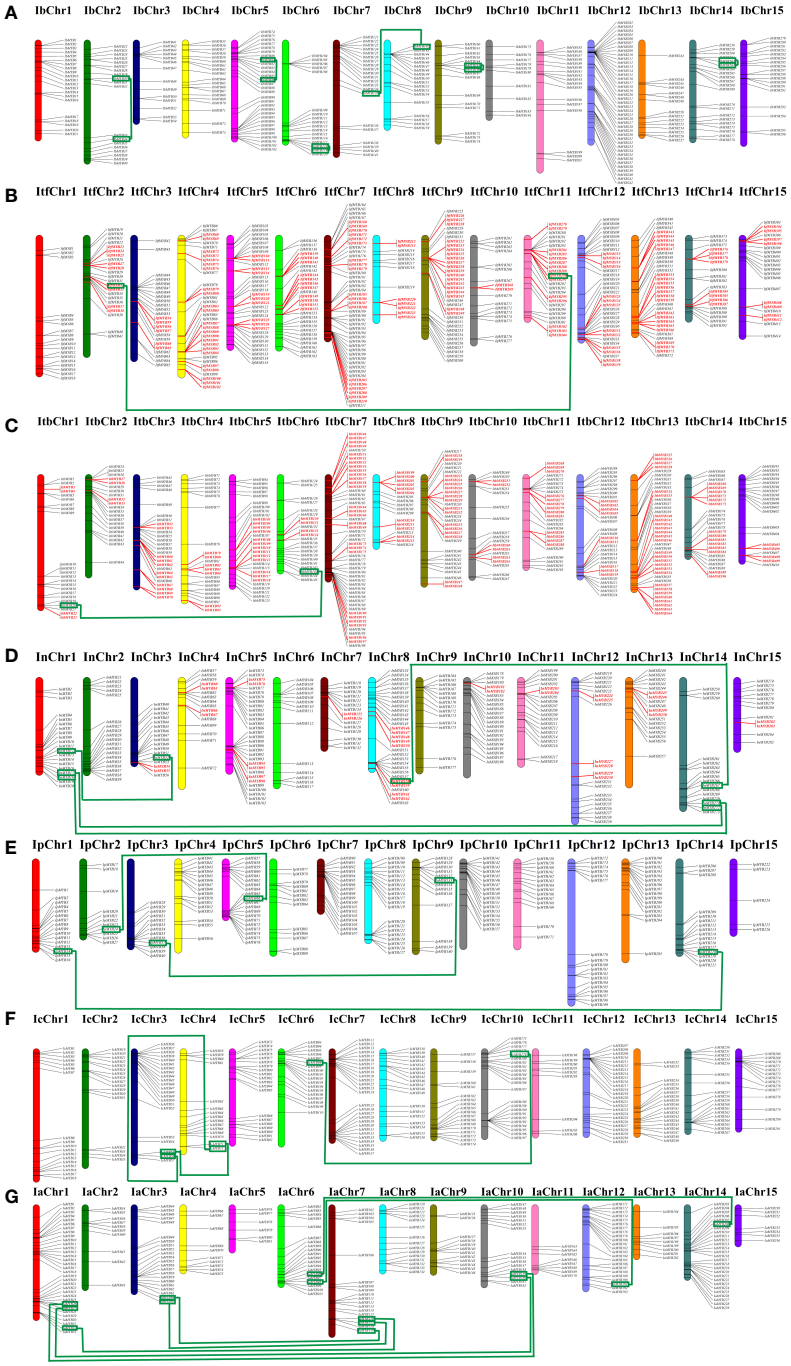


FIGURE 4
Distribution of *MYB* genes across the chromosomes of the seven *Ipomoea* species. **(A)** Distribution in sweet potato (*I. batatas*). **(B)** Distribution in (*I.*) *trifida*. **(C)** Distribution in (*I.*) *triloba*. **(D)** Distribution in (*I.*) *nil*. **(E)** Distribution in (*I.*) *purpurea*. **(F)** Distribution in (*I.*) *cairica*. **(G)** Distribution in (*I.*) *aquatic*. The red color indicates the tandemly duplicated *MYB* genes on the chromosomal positions; the green line connected green square frame indicated segmentally duplicated genes.

genes (90 from sweet potato, 86 from *I. trifida*, 90 from *I. triloba*, 91 from *I. nil*, 87 from *I. purpurea*, 89 from *I. cairica*, and 88 from *I. aquatic*) can be traced to form ortholog gene pairs between any two of the seven *Ipomea* species (Figure 5, Table S11). Of the 621 *MYB* genes, the A4 subfamily members were the most abundant (40), followed by the A16 (36), B6 (36), A15 (33), A10 (30), B45 (29), B43 (24), A13 (23), and other (<20) subfamily members (Table S11).

3.8 Ka/Ks analysis of duplicated and syntenic *MYB* genes

To detect whether some *MYB* genes are under positive selection, Ka/Ks analysis was performed on duplicated and syntenic *MYB* genes within or between the seven *Ipomoea* species. Within the seven *Ipomoea* species, all of the segmentally

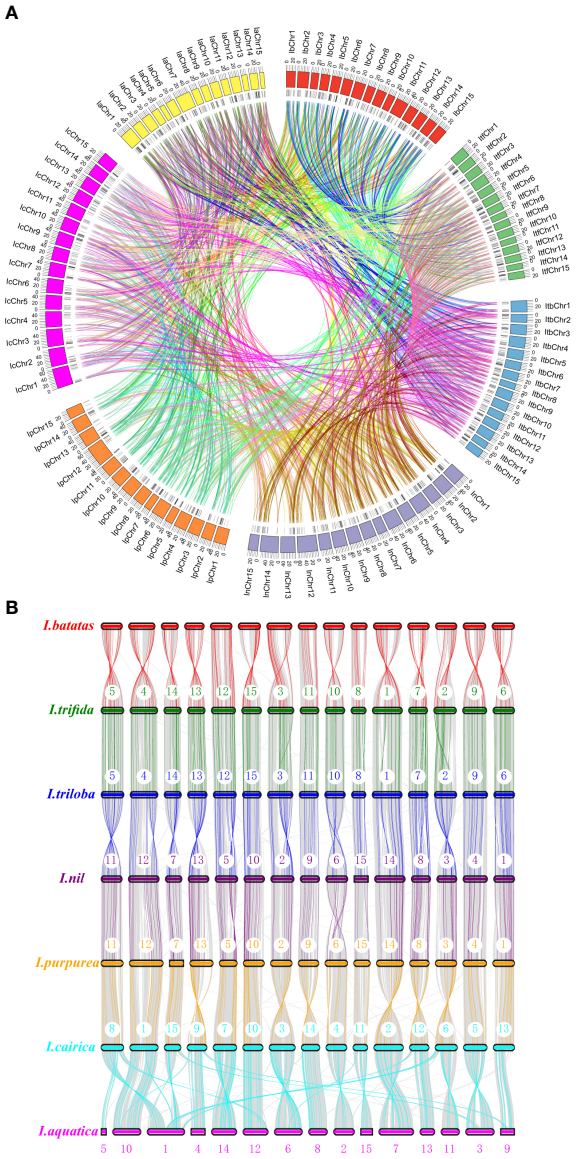


FIGURE 5
Syntenic analyses of *MYB* genes in the genomes of *Ipomoea* species. **(A)** Chromosomal distribution in the seven *Ipomoea* species. The outer circle represents the haploid chromosomes of sweet potato (*I. batatas*) (red), (*I.*) *trifida* (green), *(I.) triloba* (cornflower blue), (*I.*) *nil* (medium purple), (*I.*) *purpurea* (orange), (*I.*) *cairica* (magenta), and (*I.*) *aquatic* (yellow), respectively. The second circle (black) represents the matches of *MYB* genes with the genome of the *Ipomoea* species. Colorful lines show the collinear *MYB* gene pairs in the whole genome of the *Ipomoea* species. **(B)** Schematic representation of syntenic genes among sweet potato (*I. batatas*), (*I*) *trifida*, *(I) triloba*, (*I*) *nil*, (*I*) *purpurea*, (*I*) *cairica*, and (*I*) *aquatic*. The chromosomes of the seven Ipomoea species were reordered. Gray lines in the background indicate the collinear blocks within *Ipomoea* genomes, with *MYB* gene pairs highlighted in chromatic color. The black lines indicated synteny gene pairs between the seven *Ipomea* species.

duplicated genes (23 pairs) harbored a Ka/Ks ratio <1, suggesting that these genes underwent purifying (negative) evolutionary selection. Fifty-one pairs of tandemly duplicated genes' Ka/Ks <1 underwent purifying (negative) evolutionary selection, and 21 pairs of tandemly duplicated genes possessed Ka/Ks ratio >1 and were subjected to positive selection (Table S12). Between the seven *Ipomoea* species, 99.5% (4,116 pairs) of the syntenic *MYB* genes possessed Ka/Ks ratio <1 and the other ones (0.5%, 22 pairs) possessed Ka/Ks ratio >1 (Table S12). These results suggest that the majority of duplicated and syntenic *MYB* genes were subjected to purifying selection inside duplicated genomic elements during speciation. In contrast, a smaller number of such genes were subjected to positive selection.

3.9 *Cis*-regulatory elements (CREs) in putative promoter regions of the *Ipomoea MYB* genes

A 1.5-kb upstream promoter region of *Ipomoea MYB* genes was scanned for CRE analysis. A lot of CREs related to plant growth and development (light response, cell cycle, endosperm expression, meristem expression, root-specific and seed-specific regulation, and zein metabolism regulation) and stress response (ABA response, anaerobic induction, defense and stress responsiveness, drought stress, flavonoid biosynthetic genes regulation, gibberellin responsiveness, SA responsiveness, low-temperature responsiveness, and MeJA response) were detected (Figure S5, Table S13). Of the CREs related to plant growth and development, the Box 4 motif was the most abundant, followed by the G-Box motif and GT1-motif, and all of them contributed to light response. Of the CREs related to stress response, ABRE motif (ABA response) was the most abundant, followed by ARE (anaerobic induction), TGACG-motif (MeJA response), MBS (drought stress), TCA-element (SA responsive), LTR (low-temperature responsiveness), TGA-element (anaerobic induction), and TC-rich repeats (defense and stress responsiveness) (Figure S5, Table S13).

The entire CREs shown in the head of Table S13 appeared in both *MYB-related* and *R2R3-MYB* promoter regions, with Box 4 being the most abundant, followed by G-Box, ABRE, and ARE (Table S14). However, some CREs (3-AF1-binding site, AAAC-motif, chs-Unit 1 m1, GATT-motif, motif I, RY-element, TGA-box, and RY-element were detected in *3R-MYB* promoter regions and the AAAC-motif, AT1-motif, ATC-motif, ATCT-motif, CAG-motif, chs-CMA2a, chs-Unit 1 m1, GA-motif, Gap-box, GATT-motif, GTGGC-motif, LS7, MRE, MSA-like, motif I, RY-element, TGA-box, TC-rich repeats, and MBSI) were absent in *3R-*, *4R-*, and *5R-MYB* promoter regions (Table S14). The CGTCA-motif, ABRE, ARE, LTR, and TGACG-motif were the most conserved in the *Ipomoea MYB-related* promoter regions, and they could be detected in all of the subfamilies of this type of genes (Table S14). Box 4 was the most conserved in the *Ipomoea R2R3-MYB* promoter regions, followed by TCT-motif, CGTCA-motif, ARE, and TGACG-motif, and they were absent in no more than three of the subfamilies of this type of genes (Table S14). Box 4 and ARE were the most conserved in the subfamilies of *3R-MYB*, and chs-CMA1a, G-Box, and ABRE were the most conserved in the subfamilies of *4R-* and *5R-MYB* (Table S14). However, some CREs showed subfamily specificity; for example, the GATT motif was only detected in the A13, A14, A15, A18, B20, B22, and B37 subfamilies (Table S14). In each subfamily of the MYB promoter regions, the distribution of the CREs was not conservative as the gene-coding sequences; the distribution even showed gene specificity (Figure S6). The same distribution of the CREs

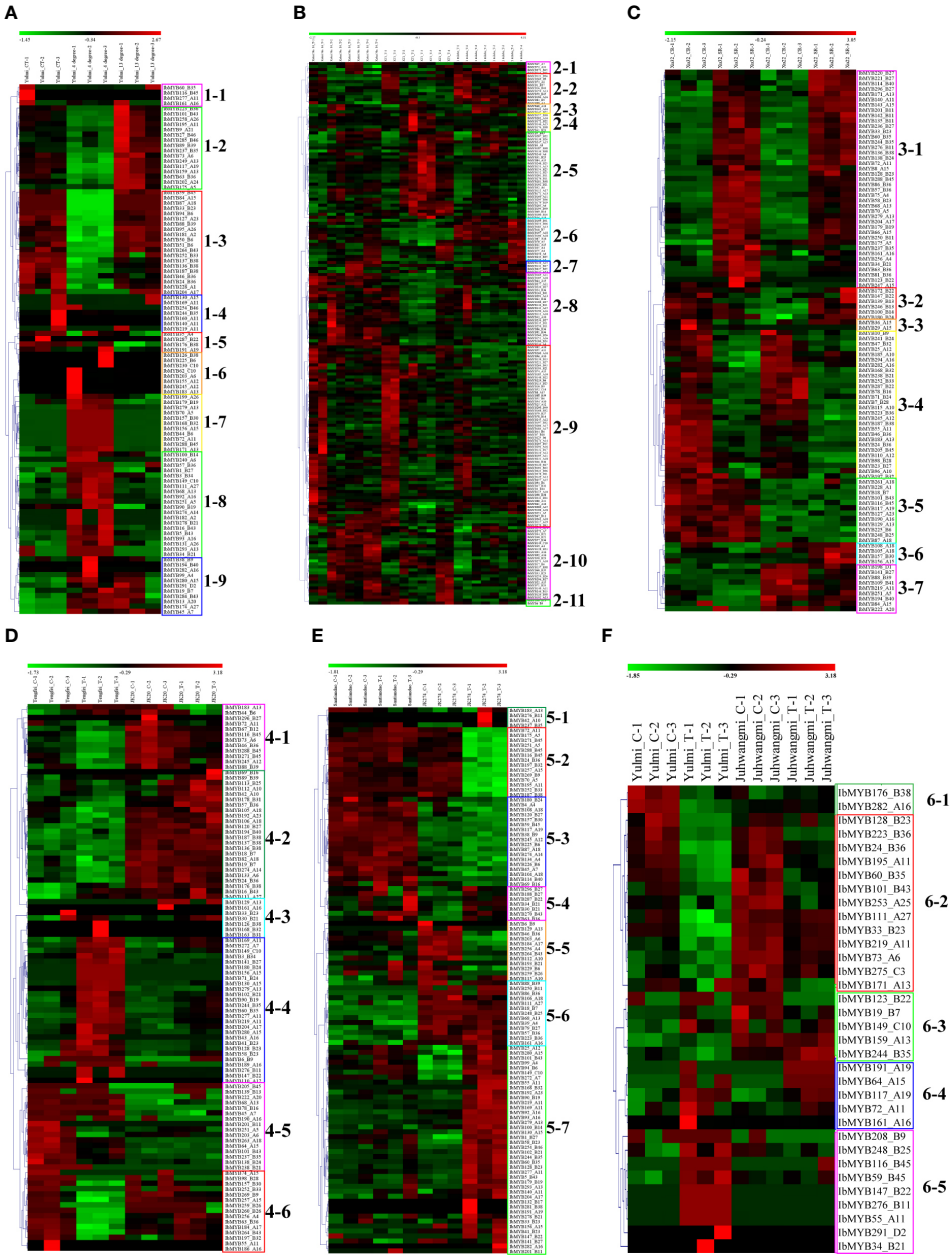


FIGURE 6
Heatmap of the expression profiles of sweet potato *MYB* differentially expressed genes (DEGs) in response to abiotic and biotic stresses. **(A)** *MYB* DEGs in "Yulmi" under control (CT), 4°C, and 13°C. **(B)** *MYB* DEGs in "Kokei 14," "KT 1," and "*I. triloba*" during different times of drought treatments. **(C)** *MYB* DEGs in "Xu 32" and "Xu 22" under control and salt treatments. CR, control; SR, treatment. **(D)** *MYB* DEGs in "Tengfei" and "JK20" under control and sweet potato stem nematode inoculation. C, control; T, treatment. **(E)** *MYB* DEGs in "Santiandao" and "JK274" under control and *Ceratocystis fimbriata* inoculation. C, control; T, treatment. **(F)** *MYB* DEGs in "Yulmi" and "Juhwangmi" under control and root-knot nematode inoculation. C, control; T, treatment. The letters after the MYB gene name indicated its phylogenetic subgroup.

occasionally appeared in tandem-duplicated genes; for example, *ItbMYB210*, *ItbMYB211*, and *ItbMYB212* were tandem-duplicated genes detected in *I. triloba*, and they shared the same CRE distribution. However, the MYB genes in the same family may show similarity in the number and type of CREs; for instance, *ItfMYB68*, *IbMYB39*, and *ItfMYB69* in the A4 subfamily, and *ItbMYB45*, *IbMYB141*, and *ItfMYB42* in the B27 subfamily (Table S13; Figure S6).

3.10 Expression profiles of *MYB* genes in sweet potato

To explore *MYB* genes related to stress response, six transcriptome datasets referring to abiotic stresses (cold, drought, and salt) and biotic stresses (sweet potato stem nematodes, *Ceratocystis fimbriata*, and root-knot nematodes) were analyzed. A total of 101 *MYB* differentially expressed genes (DEGs) were detected in the transcriptome dataset of chilling injury at low-

10.3389/fpls.2023.1155018

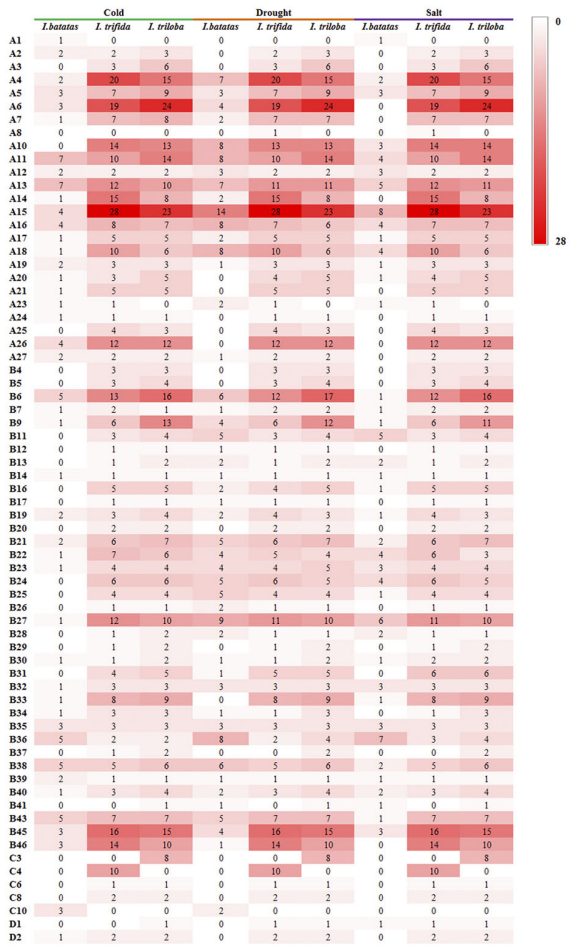


FIGURE 7
Comparison of subfamily-based gene numbers under abiotic stresses in the sweet potato, *I. trifida*, and *I. triloba*.

temperature storage, which was further classed into nine subclasses (1–1 to 1–9) based on their expression (Figure 6A). Compared with the control temperature (CT), the *MYB* genes in subclass 1-2 were upregulated at 13°C and downregulated at 4°C; the ones in subclass 1-3 were unchanged at 13°C and downregulated at 4°C; the ones in subclasses 1-7 and 1-8 were all unchanged at 13°C and upregulated at 4°C (Figure 6A). A total of 180 *MYB* DEGs were found in the drought stress transcriptome analysis, and they were classed into 11 subclasses (2-1 to 2-11) (Figure 6B). With the extension of drought time, the *MYB* genes in subclass 2-5 mainly maintained high expression in drought sense cultivar "Kokei 14," whereas they were comparably downregulated in both "KT 1" and "*I. triloba*" (drought-tolerant types) (Figure 6B). A total of 102 MYB DEGs were found in the salt stress transcriptome analysis, and they were classed into seven subclasses (3-1 to 3-7) (Figure 6C). Compared with the control, the *MYB* genes in subclass 3-1 were upregulated in both "Xu 32" (salt sense cultivar) and "Xu 22" (salt-tolerant cultivar), whereas the *MYB* genes in subclass 3-4 were downregulated in the two investigated cultivars; compared with the salt-sensitive cultivar "Xu 32," the *MYB* genes in subclass 3-5 were kept downregulated under control and salt treatment of "Xu 22"; conversely, the *MYB* genes in subclass 3-7 were mainly kept upregulated under control and salt treatment on "Xu 22" (Figure 6C).

A total of 110 *MYB* DEGs were found in the dataset of sweet potato stem nematode inoculation transcriptome analysis, and they were further classed into six subclasses (4-1 to 4-6) (Figure 6D). Compared with the sense cultivar "Tengfei," the *MYB* genes in subclass 4-2 were mainly kept upregulated under control and sweet potato stem nematode inoculation of "JK20" (sweet potato line); compared to the control, the *MYB* genes in subclass 4-4 were upregulated in sweet potato stem nematode inoculation of "Tengfei," whereas they remained downregulated in the same treatment of "JK20" (Figure 6D). There were 101 *MYB* DEGs found in the dataset of *Ceratocystis fimbriata* inoculation transcriptome analysis, and they were classed into seven subclasses (5-1 to 5-7) (Figure 6E). Compared with the control, the *MYB* genes in subclass 5-2 were downregulated in the *Ceratocystis fimbriata*-resistant line "JK274" after inoculation, whereas they kept a high expression in the sense cultivar "Santiandao"; the *MYB* genes in subclass 5-7 were upregulated in the *Ceratocystis fimbriata*-resistant line "JK274" after inoculation, whereas they kept a low expression in the sensitive cultivar "Santiandao"; the *MYB* genes in subclass 5-3 kept a high expression in "Santiandao" and a low expression in "JK274"; and the *MYB* genes in subclass 5-6 kept a low expression in "Santiandao" and a high expression in "JK274" (Figure 6E). There were 34 *MYB* DEGs found in the dataset of root-knot nematode inoculation transcriptome analysis, and they were classed into five subclasses (6-1 to 6-5) (Figure 6F).

During the drought and salt treatments of sweet potato, the DEGs belonging to the A15, B27, and B36 phylogenetic subfamily were the most, whereas in cold stress, the DEGs belonging to the A11, A13, B38, B43, B6, and B36 subfamilies were the most (Figures 6 ,7). During the biotic stress treatments, the DEGs belonging to the A11, A13, A16, and B23 subfamilies were abundant (Figures 6, 7). Both the abiotic and biotic stress treatments detected DEGs in the B11, B23, B25, B35, and B38 subfamilies, the members of which were considered to have the function of defense (Figures 3, 6, 7). To compare the transcriptional changes of various *MYB* subfamilies between species, the *I. trifida* and *I. triloba MYB* genes that responded to abiotic stresses were analyzed. During the cold, drought, and salt treatments, 365, 359, and 366 *I. trifida MYB* DEGs and 370, 369, and 368 *I. triloba MYB* DEGs were found, respectively (Tables S15, S16). When comparing the results with that of sweet potato, we discovered that the *MYB* DEGs of the three *Ipomoea* species dispersed in 68 phylogenetic subfamilies (Figure 7). Of these subfamilies, 41, 45, and 38 were common in the three *Ipomoea* species and contained DEG response to cold, drought, and salt treatments, respectively (Figure 7, Table S17). Further analysis revealed that the *MYB* DEG response to abiotic stresses generally belonged to MYB-related and R2R3-MYB gene types, involving 28 phylogenetic subfamilies, such as A4, A5, A13, A15, B23, B35, and B38 (Figure 7, Table S18).

3.11 Expression analysis of sweet potato *MYB* genes by quantitative reverse-transcription polymerase chain reaction (qRT-PCR)

Based on transcriptome results, *IbMYB68*, *IbMYB237*, *IbMYB39*, *IbMYB128*, *IbMYB127*, *IbMYB201*, *IbMYB82*, *IbMYB187*, *IbMYB4*, and *IbMYB18* were selected for further analysis using qRT-PCR. Under cold treatments, compared with the control condition (0 h), the transcripts of *IbMYB68* in "JK328" were upregulated and peaked at 12 h after treatments, whereas there was no obvious change in "Xu 32"; the transcripts of *IbMYB237* in both "Xu 32" and "JK328" were upregulated and peaked at 6 and 24 h after treatments, respectively (Figure 7). Under drought treatments, compared with the control condition (0 h), the transcripts of *IbMYB39* and *IbMYB128* in "Xu 32" were upregulated and peaked at 12 and 4 h after treatments, respectively, whereas their transcripts showed no obvious change in "JK328" (Figure 7). Under salt treatments, compared with the control condition (0 h), the transcripts of *IbMYB127* in "Xu 32" were upregulated and peaked at 12 h after treatments, whereas the transcripts of *IbMYB127* in "JK328" were downregulated and decreased to valley at 12 h after treatments; the transcripts of *IbMYB201* in both "Xu 32" and "JK328" were upregulated and peaked at 6 and 12 h after treatments, respectively (Figure 7). Under sweet potato stem nematode infection treatments, compared with the control condition (0 h), the transcripts of *IbMYB82* in both "Tengfei" and "JK20" were upregulated and decreased to the valley at 2 days and 12 h after treatments, respectively; the transcripts of *IbMYB187* in "Tengfei" were downregulated and peaked at 1 day after treatments, whereas the transcripts of *IbMYB187* in "JK20" were upregulated and peaked at 12 h after treatments (Figure 7). Under *Ceratocystis fimbriata* inoculation treatments, compared with the control condition (0 h), the transcripts of *IbMYB4* and *IbMYB18* in "JK274" were upregulated and peaked at 2 days and 1 day hours after treatments, respectively, whereas their transcripts showed no obvious changing in "JK274" (Figure 7).

4 Discussion

The MYB transcription factor gene family is one of plants' most prominent transcription factor families. It is involved in various essential life processes, including plant growth (Zhao et al., 2019), development (Liu et al., 2021; Yu et al., 2021), biosynthesis (Sinaga et al., 2021), and defense responses (Dubos et al., 2010; Lv et al., 2021; Hawku et al., 2022). Therefore, this gene family attracted more and more attention from various plant researchers and has been extensively investigated in many plant species, such as Arabidopsis (Chen et al., 2006), rice (Chen et al., 2006), soybean (Du et al., 2012b), maize (Du et al., 2012a), and potato (Li et al., 2021c). The *Ipomoea* contains 600–700 species, many of which are of considerable importance as medicinal or ornamental plants (Nimmakayala et al., 2011). The sweet potato is a highly heterozygous hexaploid (2n = 6x = 90) with a large, complex genome approximately 2–3 Gb in size (Ozias-Akins and Jarret, 1994). It is the world's seventh most important food crop by production quantity (https://www.fao.org/faostat/, accessed 2022 November). *I. trifida* and *I. triloba* are two diploid species closely related to the hexaploid sweet potato, with an estimated genome size of 526.4 and 495.9 Mb, respectively (Wu et al., 2018). Due to their smaller genome size and ploidy, they were generally used as model plants to facilitate sweet potato breeding. *I. nil* is an annual climbing herb-producing blue flower capable of self-pollination (Hoshino et al., 2016). With an estimated genome size of 750 Mb, it has been a model plant for studying photoperiodic flowering and flower coloration. *I. purpurea*, with an estimated haploid genome size of 814 Mb, is a common agricultural weed exhibiting varying levels of herbicide resistance (Gupta et al., 2021). *I. cairica* is a perennial creeper introduced as a garden ornamental, and it has an estimated haploid genome size of 730 Mb (Jiang et al., 2022). *I. aquatica*, with an estimated genome size of 550.03 Mb, is usually called water spinach (Hao et al., 2021). It is one of the most popular green leafy vegetables, with aquatic and terrestrial characteristics (Hao et al., 2021). However, little is known about the MYB gene family in *Ipomoea* species.

In the present study, a total of 2,212 MYB genes were identified from the seven *Ipomoea* species, and the method used to identify the *Ipomoea MYB* genes was reliable and has been reported by previous studies (Zhou et al., 2020; Li et al., 2021c; Wang et al., 2022). Significant variations in *MYB* family members were observed in our analyzed *Ipomoea* species. There were 296, 430, 411, 291, 226, 281, and 277 *MYB* genes identified from sweet potato (*I. batatas*), *I. trifida*, *I. triloba*, *I. nil*, *I. purpurea*, *I. cairica*, and *I. aquatic*, respectively (Tables S1, S2). The number of MYB TFs in the seven investigated *Ipomoea* species differed, suggesting that *MYBs* in different *Ipomoea* plants have expanded to different degrees (Cao et al., 2021). By comparing the genome size and the number of identified *MYB* genes in this study, it is found that the number of *MYB* genes in each *Ipomoea* species was not highly correlated with the genome size (Yang et al., 2019; Liu et al., 2020a). Moreover, it might also correlate with the quality of the assembled *Ipomoea* genomes, and the MYB gene number will change along with the improvement of *Ipomoea* genome quality (Yang et al., 2019; Liu et al., 2020a). In various plants, the number of *MYB* genes is different; for instance, 198 MYBs were identified in Arabidopsis (Chen et al., 2006), 253 in *Hedychium coronarium* (Abbas et al., 2021), 235 in *Capsicum* spp. (Arce-Rodríguez et al., 2021), 174 in *M. rubra* (Cao et al., 2021), and 155 in *Petunia axillaris* (Chen et al., 2021). A similar phenomenon has also been reported in other closely related species, for instance, *S. lycopersicum* (127) and *S. tuberosum* (217) (Li et al., 2016; Li et al., 2021c), *M. acuminate* (305), and *M. balbisiana* (251) (Tan et al., 2020). The identified *Ipomoea* MYB TFs were classed into five types: MYB-related (1R-MYB), R2R3-MYB (2R-MYB), 3R-MYB, 4R-MYB, and 5R-MYB. Among these types, the MYB-related or R2R3-MYB was the largest in *Ipomoea* species, and they contain the highest quantity of genes within the MYB family. The R2R3-MYB was usually the largest type in green plants (Martin and Paz-Ares, 1997; Abbas et al., 2021; Arce-Rodríguez et al., 2021; Wang et al., 2021b; Li et al., 2021c), whereas the MYB-related type was also reported as the largest subfamily in some species (Mmadi et al., 2017; Pu et al., 2020).

In order to investigate the evolutionary relationships of *Ipomoea MYB* genes, a systematic analysis was performed on

MYB-related TFs, R2R3-MYB TFs, 3R-MYB TFs, and the other (4R- and 5R-MYB) TFs of *Ipomoea* species and Arabidopsis, and four trees of them were constructed, respectively. Based on the topology, the trees were further divided into several subgroups (Figures 2, 3, Figures S2, S3). Supposed that genes belonging to the same branch may experience a standard evolutionary process and possess a conserved function (Li et al., 2021c), this study compared the *MYB* genes of *Ipomoea* species and Arabidopsis. Subfamilies containing the *MYB* members from the *Ipomoea* species, Arabidopsis, and some species-specific subfamilies were detected. Similar results have been reported by phylogenetic analysis of other species (Wang et al., 2021b; Li et al., 2021c). The phylogenetic trees of MYB protein constructed in this study would help to predict the function of *Ipomoea MYB* genes with Arabidopsis *MYB* genes as a reference (Chen et al., 2006; Dubos et al., 2010).

The *Ipomoea MYB* genes identified in this study were different in amino acid sequence length, molecular weight, and isoelectric point, reflecting the complexity and functional diversity of the *Ipomoea MYB* genes to some degrees centigrade (Tables S2-5). The structure analysis of the *Ipomoea MYB* genes was consistent with phylogenetic analysis. Most genes in the same subgroup exhibited similar exon–intron structures (Figure S4). The *Ipomoea* MYB proteins within the same subgroup showed similar motif compositions, whereas the ones in different subgroups were of high variance (Figure S4). This is consistent with previous reports of other species, such as *G. hirsutum* (Salih et al., 2016), *Prunus salicina* (Liu et al., 2020a), and *Helianthus annuus* (Li et al., 2020). It was also found that the motifs of the core *MYB* domain were the most conservative (Figure 1, Figure S4). The *Ipomoea MYB* genes with a close evolutionary relationship might have similar functions (Wang et al., 2021b). Conserved motifs within the same TF family, especially the ones in the same subfamilies, may play essential roles in protein-specific functions (Chen et al., 2021).

The chromosomal location showed that the *Ipomoea MYB* genes were unevenly distributed across the chromosomes of the *Ipomoea* species (Figure 4). The uneven distributions of *MYB* genes have also been detected in chromosomes of other species, such as *Jatropha curcas* (Zhou et al., 2015), *M. rubra* (Cao et al., 2021), and *C. equisetifolia* (Wang et al., 2021b). The gene duplication events, segmental and tandem duplication, played essential roles in gene family expansion and distribution of genes in plants (Cannon et al., 2004; Kong et al., 2007; Jiang et al., 2013). In the present study, segmental and tandem duplications contributed to the *MYB* gene expansion in *I. trifida*, *I. triloba*, and *I. nil*; tandem duplications were predominant, whereas in sweet potato (*I. batatas*), *I. purpurea*, *I. cairica*, and *I. aquatic*, segmental duplications were predominant (Figure 4, Table S7). The predominant types of duplications might be species-specific performance, for instance, in *H. coronarium* (Abbas et al., 2021), *Prunus salicina* (Liu et al., 2020a), and *Helianthus annuus* (Li et al., 2020), segmental duplications were dominant, whereas in watermelon (Qing et al., 2018) and palm (Zhou et al., 2020), tandem duplications were predominant.

To further explore the evolutionary relationships of *Ipomoea MYB* genes, the syntenic maps of the seven *Ipomoea* species were constructed and compared. There were 1,703 *Ipomoea MYB* genes that formed 4,050 ortholog pairs detected in the seven *Ipomoea* species (Tables S8, S9). The results showed that many *MYB* genes from *Ipomoea* species presented high levels of collinearity and suggested that these *MYB* genes probably come from a common ancestor (Liu et al., 2020a; Abbas et al., 2021). It was found that some *MYB* genes contained more than one counterpart in the *Ipomoea* species, suggesting that the MYB gene family might undergo different amplification in *Ipomoea* genomes (Liu et al., 2020a). The orthologous events occurred in all phylogenetic subfamilies except for the A1, A22, and D1 (Table S8). The A1 and A22 subfamily contained only one *Ipomoea MYB* gene (*IbMYB228* and *IbMYB290*) (Figure 2); the D1 subfamily consisted of three *MYB* members (*IbMYB198*, *ItfMYB414*, and *ItbMYB404*) (Figure S3). In most cases, the two *MYB* genes of orthologous pairs were generally from the same subfamily. The A13 subfamily held the most orthologous members, followed by A10, A4, A16, B6, and B45. A total of 621 *Ipomoea MYB* genes could form orthologous gene pairs between any two of the seven *Ipomea* species (Figure 5, Table S11).

Ka/Ks analysis of duplicated and syntenic *MYB* genes within or between the seven *Ipomoea* species showed that most duplicated

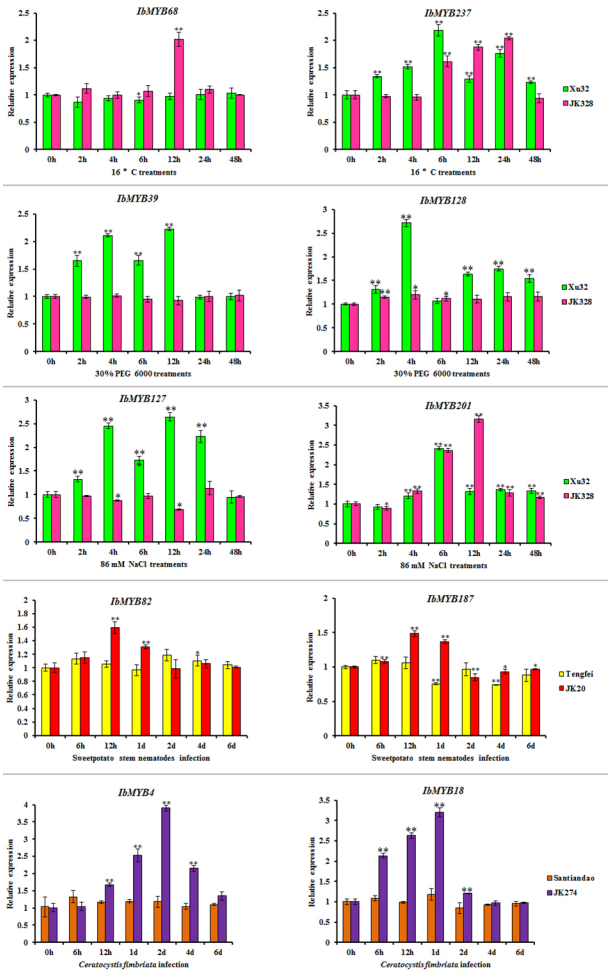


FIGURE 8
Expression analyses of *IbMYB* genes in sweet potato cultivars or lines. Denoted the significance of expression levels compared with control were ∗<0.05, ∗∗<0.01. h, hours; d, day(s).

and syntenic *MYB* genes were subjected to purifying selection inside duplicated genomic elements during speciation. In contrast, fewer such genes were subjected to positive selection (Table S12). These results indicated that *Ipomoea MYB* genes underwent a robust purifying selection during evolution with slight variation after duplication (Zhou et al., 2020). The phenomenon of the high proportion of purifying selection and low proportion of positive selection has been reported in previous MYB gene family studies, such as *G. hirsutum* (Salih et al., 2016), *Elaeis guineensis* (Zhou et al., 2020), and *H. coronarium* (Abbas et al., 2021).

The CREs in promoters of genes are usually involved in transcription initiation and regulation (Ho and Geisler, 2019; Mulat and Sinha, 2021). The present study scanned a 1.5-kb upstream promoter region of *Ipomoea MYB* genes for *cis*-regulatory elements analysis. Various CREs related to plant growth, development, and stress response were detected (Figure S5, Table S13). CREs related to stress response, such as ABRE, ARE, TGACG-motif, MBS, TCA-element, LTR, TGA-element, and TC-rich repeats, were abundant in the promoters of the *Ipomoea MYB* gene. Similar results have been reported by other species, such as *Medicago sativa* (Zhou et al., 2019), *H. coronarium* (Abbas et al., 2021), and *Petunia axillaris* (Chen et al., 2021). It was also found that some CREs were detected in *MYB-related* and *R2R3-MYB* promoter regions, whereas they were absent in *3R*-, *4R*-, and *5R-MYB* promoter regions, suggesting that different transcription initiation or regulation might occur between them (Ho and Geisler, 2019; Mulat and Sinha, 2021). In the subfamilies of *MYB-related* and *R2R3-MYB* genes, the CGTCA-motif, ABRE, ARE, and TGACG-motif were the most conserved. Moreover, some subfamily-specific CREs were also detected (Table S14). The distribution of the CREs was not conservative as the gene-coding sequences in each subfamily, and the non-conservative distribution of CREs has been reported by other TF gene families, such as NAC (Diao et al., 2018), WRKY (Zheng et al., 2021), and HSP20 (Cui et al., 2021). In this study, the exact distribution of the CREs occasionally appeared in tandemly duplicated genes. However, the *MYB* genes in the same family may show similarities in the number and type of CREs (Table S13; Figure S6).

Focused on the stress resistance of sweet potatoes, six transcriptome datasets referring to abiotic and biotic stresses were analyzed, and DEGs in each transcriptome dataset were detected. For comparative analysis, the *I. trifida* and *I. triloba MYB* genes that responded to abiotic stresses were also detected and analyzed in this study (Figure 7). The results showed that the *Ipomoea MYB* DEGs' response to abiotic stresses usually belonged to MYB-related and R2R3-MYB types and was referred to various phylogenetic subfamilies (Figure 7, Table S18). In these subfamilies, some were considered as the defense function subfamilies with Arabidopsis MYB members as the reference (Chen et al., 2006), such as B11, B23, B25, B35, and B38 (Figure 3). During the abiotic stress treatments in the sweet potato, *I. trifida* and *I. triloba*, 28 phylogenetic subfamilies contained the *MYB* members that responded in all the abiotic stress treatments of the three *Ipomoea* species (Figure 7, Table S18). Based on the results of transcriptome datasets analysis, *IbMYB68*, *IbMYB237*, *IbMYB39*, *IbMYB128*, *IbMYB127*, *IbMYB201*, *IbMYB82*, *IbMYB187*, *IbMYB4*, and *IbMYB18* were selected for further analyzing using qRT-PCR (Figure 8). The results of qRT-PCR of the selected *IbMYB* genes were consistent with the transcriptome analysis. *IbMYB68* and *IbMYB237* responded to cold treatments, *IbMYB39* and *IbMYB128* responded to drought treatments, *IbMYB127* and *IbMYB201* responded to salt treatments, *IbMYB82* and *IbMYB187* responded to sweet potato stem nematode infection, and *IbMYB4* and *IbMYB18* responded to *Ceratocystis fimbriata* infection. These results indicated that the *IbMYB* genes were essential in biotic and abiotic stress responses.

5 Conclusions

This study is the first to systematically analyze the *MYB* gene family in *Ipomoea* species. There were 296, 430, 411, 291, 226, 281, and 277 *MYB* genes were identified from sweet potato (*I. batatas*), *I. trifida*, *I. triloba*, *I. nil*, *I. purpurea*, *I. cairica*, and *I. aquatic*, respectively. The identified *MYB* genes were classified into five types and were subjected to molecular weight (MW), isoelectric point (pI), and subcellular localization analysis. Based on the topology, phylogenetic analysis classed the *Ipomoea MYB* genes into several subgroups using the Arabidopsis *MYB* genes as a reference. Gene structure, motif organization, duplicated and syntenic analyses, and *cis*-regulatory element investigation were further performed. Focus on sweet potato, six transcriptome datasets referring to different abiotic and biotic stresses were analyzed and DEGs related to stress responses were detected. Based on the expression profiles of *IbMYBs*, 10 of the detected DEGs were selected for qRT-PCR analysis under five abiotic and biotic stress treatments. The results were consistent with the transcriptome analysis. This study may help to study the evolution of *MYB* genes in the *Ipomoea* species genomes and provide helpful information for the resistance breeding of sweet potatoes.

Data availability statement

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

Author contributions

Design of the study: ZS and LW; identification of *MYB* genes in the seven *Ipomoea* species: ZS and LW; molecular weight (MW) analysis: ZS, LW, and ZJ; isoelectric point (pI) analysis: ZS, LW, and MZ; subcellular localization analysis: ZS, ZJ, and YQ; motif analysis: ZS and YQ; gene structure analysis: ZS, LW, and ZJ; phylogenetic analysis: ZS, LW, and ZJ; chromosome location: ZS, KZ, and ZJ; duplication pattern analysis: ZS, LW, MZ, and ZJ; syntenic analysis: ZS, LW, MZ, and KZ; Ka/Ks analysis: ZS, KZ, MZ, and ZJ; expression profile analysis: ZS, KZ, MZ, and YQ; qRT-PCR analysis: ZS, LW, KZ, and YQ; manuscript preparation: ZS, LW, ZJ, MZ, KZ, and YQ. All authors contributed to the article and approved the submitted version.

Funding

This work was funded by the Department of Science and Technology of Hebei Province (grant number: 19226335D) and Science Research Foundation of Hebei Normal University of Science and Technology (grant number: 2021JK01).

Acknowledgments

We would like to thank all the editors and the reviewers for their efforts and their valuable comments.

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1155018/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Sequence logos for the 20 conserved motifs of the MYB-related, R2R3-MYB, 3R-MYB, and 4R- and 5R-MYB proteins in the seven *Ipomoea* species, respectively.

SUPPLEMENTARY FIGURE 2
The phylogenetic tree of *Ipomoea* species and Arabidopsis 3R-MYB proteins. The proteins were grouped into 10 subgroups and each group has given a number (C1 to C10).

SUPPLEMENTARY FIGURE 3
The phylogenetic tree of *Ipomoea* species and Arabidopsis 4R- and 5R-MYB proteins.

SUPPLEMENTARY FIGURE 4
Motif and gene structure analyses of the *MYB* genes in the seven *Ipomoea* species. **(A)** Motif and gene structure analyses of the *MYB-related* genes in the seven *Ipomoea* species. **(B)** Motif and gene structure analyses of the *R2R3-MYB* genes in the seven *Ipomoea* species. **(C)** Motif and gene structure analyses of the *3R-MYB* genes in the seven *Ipomoea* species. **(D)** Motif and gene structure analyses of the *4R-* and *5R-MYB* genes in the seven *Ipomoea* species.

SUPPLEMENTARY FIGURE 5
The summary heatmap of *cis*-elements involved in plant growth and development and stress responses in *Ipomoea MYB* promoters.

SUPPLEMENTARY FIGURE 6
Cis-elements in the promoter region of *MYB* genes in *Ipomoea* species. **(A)** *Cis*-elements in *MYB-related* genes promoter; **(B)** *Cis*-elements in *R2R3-MYB* genes promoter; **(C)** *Cis*-elements in *3R-MYB* genes promoter; **(D)** *Cis*-elements in *4R-* and *5R-MYB* genes promoter.

References

Abbas, F., Ke, Y., Zhou, Y., Yu, Y., Waseem, M., Ashraf, U., et al. (2021). Genome-wide analysis reveals the potential role of MYB transcription factors in floral scent formation in *Hedychium coronarium*. *Front. Plant Sci.* 12, 623742. doi: 10.3389/fpls.2021.623742

Alabadí, D., Oyama, T., Yanovsky, M. J., Harmon, F. G., Más, P., and Kay, S. A. (2001). Reciprocal regulation between TOC1 and LHY/CCA1 within the arabidopsis circadian clock. *Science* 293 (5531), 880–883. doi: 10.1126/science.106132

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215 (3), 403–410. doi: 10.1016/S0022-2836(05)80360-2

Anisimova, M., Gil, M., Dufayard, J. F., Dessimoz, C., and Gascuel, O. (2011). Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. *Syst. Biol.* 60 (5), 685–699. doi: 10.1093/sysbio/syr041

Arce-Rodríguez, M. L., Martínez, O., and Ochoa-Alejo, N. (2021). Genome-wide identification and analysis of the MYB transcription factor gene family in chili pepper (*Capsicum* spp.). *Int. J. Mol. Sci.* 22 (5), 2229. doi: 10.3390/ijms22052229

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., et al. (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 37 (Web Server issue), W202–W208. doi: 10.1093/nar/gkp335

Bovell-Benjamin, A. C. (2007). Sweet potato: a review of its past, present, and future role in human nutrition. *Adv. Food Nutr. Res.* 52, 1–59. doi: 10.1016/S1043-4526(06)52001-7

Cannon, S. B., Mitra, A., Baumgarten, A., Young, N. D., and May, G. (2004). The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. *BMC Plant Biol.* 4, 10. doi: 10.1186/1471-2229-4-10

Cao, Y., Jia, H., Xing, M., Jin, R., Grierson, D., Gao, Z., et al. (2021). Genome-wide analysis of MYB gene family in Chinese bayberry (*Morella rubra*) and identification of members regulating flavonoid biosynthesis. *Front. Plant Sci.* 12, 1244. doi: 10.3389/fpls.2021.691384

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020). TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 13 (8), 1194–1202. doi: 10.1016/j.molp.2020.06.009

Chen, G., He, W., Guo, X., and Pan, J. (2021). Genome-wide identification, classification and expression analysis of the MYB transcription factor family in *Petunia*. *Int. J. Mol. Sci.* 22 (9), 4838. doi: 10.3390/ijms22094838

Chen, Y., Yang, X., He, K., Liu, M., Li, I., Gao, Z., et al (2006) The MYB transcription factor superfamily of arabidopsis: expression analysis and phylogenetic comparison with the rice MYB family *Plant Mol. Biol.* 60 (1), 107–124.

Cheng, Y.-J., Kim, M.-D., Deng, X.-P., Kwak, S.-S., and Chen, W. (2013). Enhanced salt stress tolerance in transgenic potato plants expressing IbMYB1, a sweet potato transcription factor. *J. Microbiol. Biotechnol.* 23 (12), 1737–1746. doi: 10.4014/jmb.1307.07024

Chezem, W. R., and Clay, N. K. (2016). Regulation of plant secondary metabolism and associated specialized cell development by MYBs and bHLHs. *Phytochemistry* 131, 26–43. doi: 10.1016/j.phytochem.2016.08.006

Chu, H., Jeong, J. C., Kim, W. J., Chung, D. M., Jeon, H. K., Ahn, Y. O., et al. (2013). Expression of the sweetpotato R2R3-type IbMYB1a gene induces anthocyanin accumulation in arabidopsis. *Physiologia Plantarum* 148 (2), 189–199. doi: 10.1111/j.1399-3054.2012.01706.x

Cui, F., Taier, G., Wang, X., and Wang, K. (2021). Genome-wide analysis of the *HSP20* gene family and expression patterns of *HSP20* genes in response to abiotic stresses in *Cynodon transvaalensis*. *Front. Genet.* 12, 732812. doi: 10.3389/fgene.2021.732812

Czemmel, S., Heppel, S. C., and Bogs, J. (2012). R2R3 MYB transcription factors: key regulators of the flavonoid biosynthetic pathway in grapevine. *Protoplasma* 249 (2), 109–118. doi: 10.1007/s00709-012-0380-z

Dai, X., Xu, Y., Ma, Q., Xu, W., Wang, T., Xue, Y., et al. (2007). Overexpression of an R1R2R3 MYB gene, OsMYB3R-2, increases tolerance to freezing, drought, and salt stress in transgenic arabidopsis. *Plant Physiol.* 143 (4), 1739–1751. doi: 10.1104/pp.106.094532

Diao, W., Snyder, J. C., Wang, S., Liu, J., Pan, B., Guo, G., et al. (2018). Genome-wide analyses of the NAC transcription factor gene family in pepper (*Capsicum annuum* l.): chromosome location, phylogeny, structure, expression patterns, *cis*-elements in the promoter, and interaction network. *Int. J. Mol. Sci.* 19 (4), 1028. doi: 10.3390/ijms19041028

Dong, J., Cao, L., Zhang, X., Zhang, W., Yang, T., Zhang, J., et al. (2021). An R2R3-MYB transcription factor RmMYB108 responds to chilling stress of rosa multiflora and conferred cold tolerance of arabidopsis. *Front. Plant Sci.* 12, 696919. doi: 10.3389/fpls.2021.696919

Du, H., Feng, B.-R., Yang, S.-S., Huang, Y.-B., and Tang, Y.-X. (2012a). The R2R3-MYB transcription factor gene family in maize. *PloS One* 7 (6), e37463. doi: 10.1371/journal.pone.0037463

Du, B., Liu, H., Dong, K., Wang, Y., and Zhang, Y. (2022). Over-expression of an R2R3 MYB gene, *MdMYB108L*, enhances tolerance to salt stress in transgenic plants. *Int. J. Mol. Sci.* 23 (16), 9428. doi: 10.3390/ijms23169428

Du, H., Yang, S. S., Liang, Z., Feng, B. R., Liu, L., Huang, Y. B., et al. (2012b). Genome-wide analysis of the MYB transcription factor superfamily in soybean. *BMC Plant Biol.* 12 (1), 1–22. doi: 10.1186/1471-2229-12-106

Du, H., Zhang, L., Liu, L., Tang, X.-F., Yang, W.-J., Wu, Y.-M., et al. (2009). Biochemical and molecular characterization of plant MYB transcription factor family. *Biochemistry* 74 (1), 1–11. doi: 10.1186/1471-2229-12-106

Dubos, C., Le Gourrierec, J., Baudry, A., Huep, G., Lanet, E., Debeaujon, I., et al. (2008). MYBL2 is a new regulator of flavonoid biosynthesis in *Arabidopsis thaliana*. *Plant J.* 55 (6), 940–953. doi: 10.1111/j.1365-313X.2008.03564.x

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., and Lepiniec, L. (2010). MYB transcription factors in arabidopsis. *Trends Plant Sci.* 15 (10), 573–581. doi: 10.1016/j.tplants.2010.06.005

Frampton, J. (2004). *Myb transcription factors: Their role in growth, differentiation and disease* (Netherlands: Springer).

Gao, S., Yu, B., Yuan, L., Zhai, H., He, S.-z., and Liu, Q.-c. (2011).). production of transgenic sweetpotato plants resistant to stem nematodes using oryzacystatin-I gene. *Scientia Hortic.* 128 (4), 408–414. doi: 10.1016/j.scienta.2011.02.015

Gigolashvili, T., Berger, B., and Flügge, U.-I. (2009). Specific and coordinated control of indolic and aliphatic glucosinolate biosynthesis by R2R3-MYB transcription factors in *Arabidopsis thaliana*. *Phytochem. Rev.* 8 (1), 3–13. doi: 10.1007/s11101-008-9112-6

Gu, K.-D., Zhang, Q.-Y., Yu, J.-Q., Wang, J.-H., Zhang, F.-J., Wang, C.-K., et al. (2020). R2R3-MYB transcription factor MdMYB73 confers increased resistance to the fungal pathogen botryosphaeria dothidea in apples *via* the salicylic acid pathway. *J. Agric. Food Chem.* 69 (1), 447–458. doi: 10.21203/rs.3.rs-107830/v1

Gupta, S., Harkess, A., Soble, A., Van Etten, M., Leebens-Mack, J., and Baucom, R. (2021). Inter-chromosomal linkage disequilibrium and linked fitness cost loci influence the evolution of nontarget site herbicide resistance in an agricultural weed. *bioRxiv* 2021.04.04.438381. doi: 10.1101/2021.04.04.438381

Haga, N., Kobayashi, K., Suzuki, T., Maeo, K., Kubo, M., Ohtani, M., et al. (2011). Mutations in MYB3R1 and MYB3R4 cause pleiotropic developmental defects and preferential down-regulation of multiple G2/M-specific genes in arabidopsis. *Plant Physiol.* 157 (2), 706–717. doi: 10.1104/pp.111.180836

Hao, Y., Bao, W., Li, G., Gagoshidze, Z., Shu, H., Yang, Z., et al. (2021). The chromosome-based genome provides insights into the evolution in water spinach. *Scientia Hortic.* 289, 110501. doi: 10.1016/j.scienta.2021.110501

Hawku, M. D., He, F., Bai, X., Islam, M. A., Huang, X., Kang, Z., et al. (2022). A R2R3 MYB transcription factor, *TaMYB391*, is positively involved in wheat resistance to *Puccinia striiformis* f. sp. tritici. *Int. J. Mol. Sci.* 23 (22), 14070. doi: 10.3390/ijms232214070

Ho, C.-L., and Geisler, M. (2019). Genome-wide computational identification of biologically significant cis-regulatory elements and associated transcription factors from rice. *Plants* 8 (11), 441. doi: 10.3390/plants8110441

Hoseinpour, F., Darvishzadeh, R., Abdollahi Mandoulakani, B., Habibi, N., Dardan, E., and Ranjbar, A. (2021). Evaluation of relative changes in the expression level of SOS2, MYB-related and HD-ZIP genes in oil seed sunflower lines under salinity stress. *J. Crop Breed.* 13 (39), 152–165.

Hoshino, A., Jayakumar, V., Nitasaka, E., Toyoda, A., Noguchi, H., Itoh, T., et al. (2016). Genome sequence and analysis of the Japanese morning glory *Ipomoea nil*. *Nat. Commun.* 7 (1), 1–10. doi: 10.1038/ncomms13295

Howe, E. A., Sinha, R., Schlauch, D., and Quackenbush, J. (2011). RNA-Seq analysis in MeV. *Bioinformatics* 27 (22), 3209–3210. doi: 10.1093/bioinformatics/btr490

Ito, M. (2005). Conservation and diversification of three-repeat myb transcription factors in plants. *J. Plant Res.* 118 (1), 61–69. doi: 10.1007/s10265-005-0192-8

Ito, M., Araki, S., Matsunaga, S., Itoh, T., Nishihama, R., Machida, Y., et al. (2001). G2/M-phase–specific transcription during the plant cell cycle is mediated by c-Myb–like transcription factors. *Plant Cell* 13 (8), 1891–1905.

Ito, M., Iwase, M., Kodama, H., Lavisse, P., Komamine, A., Nishihama, R., et al. (1998). A novel cis-acting element in promoters of plant B-type cyclin genes activates M phase–specific transcription. *Plant Cell* 10 (3), 331–341.

Jiang, S. Y., González, J. M., and Ramachandran, S. (2013). Comparative genomic and transcriptomic analysis of tandemly and segmentally duplicated genes in rice. *PloS One* 8 (5), e63551. doi: 10.1371/journal.pone.0063551

Jiang, F., Wang, S., Wang, H., Wang, A., Xu, D., Liu, H., et al. (2022). A chromosome-level reference genome of a convolvulaceae species *Ipomoea cairica*. *G3* 12 (9), jkac187. doi: 10.1093/g3journal/jkac187

Kamioka, M., Takao, S., Suzuki, T., Taki, K., Higashiyama, T., Kinoshita, T., et al. (2016). Direct repression of evening genes by CIRCADIAN CLOCK-ASSOCIATED1 in the arabidopsis circadian clock. *Plant Cell* 28 (3), 696–711. doi: 10.1105/tpc.15.00737

Kato, K., Gális, I., Suzuki, S., Araki, S., Demura, T., Criqui, M.-C., et al. (2009). Preferential up-regulation of G2/M phase-specific genes by overexpression of the hyperactive form of NtmybA2 lacking its negative regulation domain in tobacco BY-2 cells. *Plant Physiol.* 149 (4), 1945–1957. doi: 10.1104/pp.109.135582

Klempnauer, K. H., Gonda, T. J., and Bishop, J. M. (1982). Nucleotide sequence of the retroviral leukemia gene v-myb and its cellular progenitor c-myb: the architecture of a transduced oncogene. *Cell* 31 (2), 453–463. doi: 10.1016/0092-8674(82)90138-6

Koes, R., Verweij, W., and Quattrocchio, F. (2005). Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. *Trends Plant Sci.* 10 (5), 236–242. doi: 10.1016/j.tplants.2005.03.002

Kong, H., Landherr, L. L., Frohlich, M. W., Leebens-Mack, J., Ma, H., and dePamphilis, C. W. (2007). Patterns of gene duplication in the plant SKP1 gene family in angiosperms: evidence for multiple mechanisms of rapid gene birth. *Plant J.* 50 (5), 873–885. doi: 10.1111/j.1365-313X.2007.03097.x

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., et al. (2009). Circos: an information aesthetic for comparative genomics. *Genome Res.* 19 (9), 1639–1645. doi: 10.1101/gr.092759.109

Lee, M. M., and Schiefelbein, J. (1999). WEREWOLF, a MYB-related protein in arabidopsis, is a position-dependent regulator of epidermal cell patterning. *Cell* 99 (5), 473–483. doi: 10.1016/S0092-8674(00)81536-6

Li, Y., Leveau, A., Zhao, Q., and Feng, Q. (2021b). Subtelomeric assembly of a multi-gene pathway for antimicrobial defense compounds in cereals. *Nat. Commun.* 12 (1), 2563. doi: 10.1038/s41467-021-22920-8

Li, Y., Liang, J., Zeng, X., Guo, H., Luo, Y., Kear, P., et al. (2021c). Genome-wide analysis of MYB gene family in potato provides insights into tissue-specific regulation of anthocyanin biosynthesis. *Hortic. Plant J.* 7 (2), 129–141. doi: 10.1016/j.hpj.2020.12.001

Li, J., Liu, H., Yang, C., Wang, J., Yan, G., Si, P., et al. (2020). Genome-wide identification of MYB genes and expression analysis under different biotic and abiotic stresses in *Helianthus annuus* l. *Ind. Crops Products* 143, 111924. doi: 10.1016/j.indcrop.2019.111924

Li, Z., Peng, R., Tian, Y., Han, H., Xu, J., and Yao, Q. (2016). Genome-wide identification and analysis of the MYB transcription factor superfamily in *Solanum lycopersicum*. *Plant Cell Physiol.* 57 (8), 1657–1677. doi: 10.1093/pcp/pcw091

Li, L.-X., Wei, Z.-Z., Zhou, Z.-L., Zhao, D.-L., Tang, J., Yang, F., et al. (2021a). A single amino acid mutant in the EAR motif of IbMYB44. 2 reduced the inhibition of anthocyanin accumulation in the purple-fleshed sweetpotato. *Plant Physiol. Biochem.* 167, 410–419. doi: 10.1016/j.plaphy.2021.08.012

Lin, S., Singh, R. K., Moehninsi, M., and Navarre, D. A. (2021). R2R3-MYB transcription factors, StmiR858 and sucrose mediate potato flavonol biosynthesis. *Horticulture Res.* 8 (2021), 25. doi: 10.1038/s41438-021-00463-9

Liu, C., Hao, J., Qiu, M., Pan, J., and He, Y. (2020a). Genome-wide identification and expression analysis of the MYB transcription factor in Japanese plum (*Prunus salicina*). *Genomics* 112 (6), 4875–4886. doi: 10.1016/j.ygeno.2020.08.018

Liu, J., Osbourn, A., and Ma, P. (2015). MYB transcription factors as regulators of phenylpropanoid metabolism in plants. *Mol. Plant* 8 (5), 689–708. doi: 10.1016/j.molp.2015.03.012

Liu, X., Wu, Z., Feng, J., Yuan, G., He, L., Zhang, D., et al. (2021). A novel R2R3-MYB gene LoMYB33 from lily is specifically expressed in anthers and plays a role in pollen development. *Front. Plant Sci.* 12, 730007. doi: 10.3389/fpls.2021.730007

Liu, Y., Zeng, Y., Li, Y., Liu, Z., Lin-Wang, K., Espley, R. V., et al. (2020b). Genomic survey and gene expression analysis of the MYB-related transcription factor superfamily in potato (*Solanum tuberosum* l.). *Int. J. Biol. Macromolecules* 164, 2450–2464. doi: 10.1016/j.ijbiomac.2020.08.062

Lv, K., Wei, H., and Liu, G. (2021). A R2R3-MYB transcription factor gene, BpMYB123, regulates BpLEA14 to improve drought tolerance in *Betula platyphylla*. *Front. Plant Sci.* 12, 791390. doi: 10.3389/fpls.2021.791390

Ma, Q., Dai, X., Xu, Y., Guo, J., Liu, Y., Chen, N., et al. (2009). Enhanced tolerance to chilling stress in OsMYB3R-2 transgenic rice is mediated by alteration in cell cycle and ectopic expression of stress genes. *Plant Physiol.* 150 (1), 244–256. doi: 10.1104/pp.108.133454

Machado, A., Wu, Y., Yang, Y., Llewellyn, D. J., and Dennis, E. S. (2009). The MYB transcription factor GhMYB25 regulates early fibre and trichome development. *Plant J.* 59 (1), 52–62. doi: 10.1111/j.1365-313X.2009.03847.x

Magali, L., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., et al. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* 30 (1), 325–327. doi: 10.1093/nar/30.1.325

Martin, C., and Paz-Ares, J. (1997). MYB transcription factors in plants. *Trends Genet.* 13 (2), 67–73. doi: 10.1016/S0168-9525(96)10049-4

Matsui, K., Umemura, Y., and Ohme-Takagi, M. (2008). AtMYBL2, a protein with a single MYB domain, acts as a negative regulator of anthocyanin biosynthesis in arabidopsis. *Plant J.* 55 (6), 954–967. doi: 10.1111/j.1365-313X.2008.03565.x

Mizoguchi, T., Wheatley, K., Hanzawa, Y., Wright, L., Mizoguchi, M., Song, H.-R., et al. (2002). LHY and CCA1 are partially redundant genes required to maintain circadian rhythms in arabidopsis. *Dev. Cell* 2 (5), 629–641. doi: 10.1016/S1534-5807(02)00170-3

Mmadi, M. A., Dossa, K., Wang, L., Zhou, R., Wang, Y., Cisse, N., et al. (2017). Functional characterization of the versatile MYB gene family uncovered their important roles in plant development and responses to drought and waterlogging in sesame. *Genes* 8 (12), 362. doi: 10.3390/genes8120362

Mulat, M. W., and Sinha, V. B. (2021). Distribution and abundance of CREs in the promoters depicts crosstalk by WRKYs in tef [*Eragrostis tef* (Zucc.) troetter]. *Gene Rep.* 23, 101043. doi: 10.1016/j.genrep.2021.101043

Muramoto, N., Tanaka, T., Shimamura, T., Mitsukawa, N., Hori, E., Koda, K., et al. (2012). Transgenic sweet potato expressing thionin from barley gives resistance to black rot disease caused by *Ceratocystis fimbriata* in leaves and storage roots. *Plant Cell Rep.* 31 (6), 987–997. doi: 10.1007/s00299-011-1217-5

Naing, A. H., and Kim, C. K. (2018). Roles of R2R3-MYB transcription factors in transcriptional regulation of anthocyanin biosynthesis in horticultural plants. *Plant Mol. Biol.* 98 (1), 1–18. doi: 10.1007/s11103-018-0771-4

Nguyen, L. T., Schmidt, H. A., Von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32 (1), 268–274. doi: 10.1093/molbev/msu300

Nimmakayala, P., Vajja, G., and Reddy, U. K. (2011). "Ipomoea," in *Wild crop relatives: Genomic and breeding resources*. Ed. C. Kole (Berlin, Heidelberg: Springer). doi: 10.1007/978-3-642-21102-7_7

Okada, R., Kondo, S., Satbhai, S. B., Yamaguchi, N., Tsukuda, M., and Aoki, S. (2009). Functional characterization of *CCA1/LHY* homolog genes, *PpCCA1a* and *PpCCA1b*, in the moss *Physcomitrella patens*. *Plant J.* 60 (3), 551–563. doi: 10.1111/j.1365-313X.2009.03979.x

Ozias-Akins, P., and Jarret, R. L. (1994). Nuclear DNA content and ploidy levels in the genus *Ipomoea*. *J. Am. Soc. Hortic. Sci.* 119, 110–115. doi: 10.21273/JASHS.119.1.110

Park, S.-C., Kim, Y.-H., Kim, S. H., Jeong, Y. J., Kim, C. Y., Lee, J. S., et al. (2015). Overexpression of the *IbMYB1* gene in an orange-fleshed sweet potato cultivar produces a dual-pigmented transgenic sweet potato with improved antioxidant activity. *Physiologia Plantarum* 153 (4), 525–537. doi: 10.1111/ppl.12281

Park, D.-Y., Shim, Y., Gi, E., Lee, B.-D., An, G., Kang, K., et al. (2018). The MYB-related transcription factor RADIALIS-LIKE3 (OsRL3) functions in ABA-induced leaf senescence and salt sensitivity in rice. *Environ. Exp. Bot.* 156, 86–95. doi: 10.1016/j.envexpbot.2018.08.033

Paz-Ares, J., Ghosal, D., Wienand, U., Peterson, P., and Saedler, H. (1987). The regulatory c1 locus of zea mays encodes a protein with homology to myb proto-oncogene products and with structural similarities to transcriptional activators. *EMBO J.* 6 (12), 3553–3558. doi: 10.1002/j.1460-2075.1987.tb02684.x

Pu, X., Yang, L., Liu, L., Dong, X., Chen, S., Chen, Z., et al. (2020). Genome-wide analysis of the MYB transcription factor superfamily in physcomitrella patens. *Int. J. Mol. Sci.* 21 (3), 975. doi: 10.3390/ijms21030975

Qing, X., Jie, H., Jianhui, D., Xiaojin, H., and Zhang, X. (2018). Genomic survey and expression profiling of the MYB gene family in watermelon. *Hortic. Plant J.* 4 (1), 1–15. doi: 10.1016/j.hpj.2017.12.001

Saha, G., Park, J.-I., Ahmed, N. U., Kayum, M. A., Kang, K.-K., and Nou, I.-S. (2016). Characterization and expression profiling of MYB transcription factors against stresses and during male organ development in Chinese cabbage (*Brassica rapa* ssp. pekinensis). *Plant Physiol. Biochem.* 104, 200–215. doi: 10.1016/j.plaphy.2016.03.021

Salih, H., Gong, W., He, S., Sun, G., Sun, J., and Du, X. (2016). Genome-wide characterization and expression analysis of MYB transcription factors in *Gossypium hirsutum*. *BMC Genet.* 17 (1), 1–12. doi: 10.1186/s12863-016-0436-8

Schmittgen, T. D., and Livak, K. J. (2008). Analyzing real-time PCR data by the comparative CT method. *Nat. Protoc.* 3 (6), 1101–1108. doi: 10.1038/nprot.2008.73

Shangguan, X., Yang, Q., Wu, X., and Cao, J. (2021). Function analysis of a cotton R2R3 MYB transcription factor GhMYB3 in regulating plant trichome development. *Plant Biol.* 23 (6), 1118–1127. doi: 10.1111/plb.13299

Shen, X., Guo, X., Guo, X., Zhao, D., Zhao, W., Chen, J., et al. (2017). PacMYBA, a sweet cherry R2R3-MYB transcription factor, is a positive regulator of salt stress tolerance and pathogen resistance. *Plant Physiol. Biochem.* 112, 302–311. doi: 10.1016/j.plaphy.2017.01.015

Sievers, F., and Higgins, D. G. (2018). Clustal omega for making accurate alignments of many protein sequences. *Protein Sci.* 27 (1), 135–145. doi: 10.1002/pro.3290

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega. *Mol. Syst. Biol.* 7, 539. doi: 10.1038/msb.2011.75

Sinaga, D. S., Ho, S.-L., Lu, C.-A., Yu, S.-M., and Huang, L.-F. (2021). Knockdown expression of a MYB-related transcription factor gene, *OsMYBS2*, enhances production of recombinant proteins in rice suspension cells. *Plant Methods* 17 (1), 1–10. doi: 10.1186/s13007-021-00799-2

Tan, L., Ijaz, U., Salih, H., Cheng, Z., Ni Win Htet, N., Ge, Y., et al. (2020). Genome-wide identification and comparative analysis of MYB transcription factor family in *Musa acuminata* and *Musa balbisiana*. *Plants* 9 (4), 413. doi: 10.3390/plants9040413

Tang, N., Cao, Z., Yang, C., Ran, D., Wu, P., Gao, H., et al. (2021). A R2R3-MYB transcriptional activator LmMYB15 regulates chlorogenic acid biosynthesis and phenylpropanoid metabolism in lonicera macranthoides. *Plant Sci.* 308, 110924. doi: 10.1016/j.plantsci.2021.110924

Tang, H., Wang, X., Bowers, J. E., Ming, R., Alam, M., and Paterson, A. H. (2008). Unraveling ancient hexaploidy through multiply-aligned angiosperm gene maps. *Genome Res.* 18 (12), 1944–1954. doi: 10.1101/gr.080978.108

Thiedig, K., Weisshaar, B., and Stracke, R. (2021). Functional and evolutionary analysis of the arabidopsi s 4R-MYB protein SNAPc4 as part of the SNAP complex. *Plant Physiol.* 185 (3), 1002–1020. doi: 10.1093/plphys/kiaa067

Upadhyaya, G., Das, A., and Ray, S. (2021). A rice R2R3-MYB (OsC1) transcriptional regulator improves oxidative stress tolerance by modulating anthocyanin biosynthesis. *Physiologia Plantarum* 173 (4), 2334–2349. doi: 10.1111/ppl.13583

VanBuren, R., Wai, C. M., Keilwagen, J., and Pardo, J. (2018). A chromosome-scale assembly of the model desiccation tolerant grass *Oropetium thomaeum*. *Plant Direct* 2 (11), e00096. doi: 10.1002/pld3.96

Voorrips, R. E. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. *J. Hered* 93 (1), 77–78. doi: 10.1093/jhered/93.1.77

Wang, W., Fang, H., Aslam, M., Du, H., Chen, J., Luo, H., et al. (2022). *MYB* gene family in the diatom *Phaeodactylum tricornutum* revealing their potential functions in the adaption to nitrogen deficiency and diurnal cycle. *J. Phycology* 58 (1), 121–132. doi: 10.1111/jpy.13217

Wang, H., Liu, J.-X., Feng, K., Li, T., Duan, A.-Q., Liu, Y.-H., et al. (2022b). AgMYB12, a novel R2R3-MYB transcription factor, regulates apigenin biosynthesis by interacting with the AgFNS gene in celery. *Plant Cell Rep.* 41 (1), 139–151. doi: 10.1007/s00299-021-02792-4

Wang, X., Niu, Q.-W., Teng, C., Li, C., Mu, J., Chua, N.-H., et al. (2009). Overexpression of PGA37/MYB118 and MYB115 promotes vegetative-to-embryonic transition in arabidopsis. *Cell Res.* 19 (2), 224–235. doi: 10.1038/cr.2008.276

Wang, S., Shi, M., Zhang, Y., Xie, X., Sun, P., Fang, C., et al. (2021a). FvMYB24, a strawberry R2R3-MYB transcription factor, improved salt stress tolerance in transgenic arabidopsis. *Biochem. Biophys. Res. Commun.* 569, 93–99. doi: 10.1016/j.bbrc.2021.06.085

Wang, Y., Tang, H., Debarry, J. D., Tan, X., Li, J., Wang, X., et al. (2012). MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 40 (7), e49. doi: 10.1093/nar/gkr1293

Wang, C., Wang, L., Lei, J., Chai, S., Jin, X., Zou, Y., et al. (2022a). *IbMYB308*, a sweet potato R2R3-MYB gene, improves salt stress tolerance in transgenic tobacco. *Genes* 13 (8), 1476. doi: 10.3390/genes13081476

Wang, Y., Zhang, Y., Fan, C., Wei, Y., Meng, J., Li, Z., et al. (2021b). Genome-wide analysis of MYB transcription factors and their responses to salt stress in *Casuarina equisetifolia*. *BMC Plant Biol.* 21 (1), 1–17. doi: 10.1186/s12870-021-03083-6

Wang, Y., Zhou, L.-J., Wang, Y., Geng, Z., Ding, B., Jiang, J., et al. (2022c). An R2R3-MYB transcription factor CmMYB21 represses anthocyanin biosynthesis in color fading petals of chrysanthemum. *Scientia Hortic.* 293, 110674. doi: 10.1016/j.scienta.2021.110674

Wu, S., Lau, K. H., Cao, Q., Hamilton, J. P., Sun, H., Zhou, C., et al. (2018). Genome sequences of two diploid wild relatives of cultivated sweetpotato reveal targets for genetic improvement. *Nat. Commun.* 9 (1), 1–12. doi: 10.1038/s41467-018-06983-8

Wu, Y., Li, T., Cheng, Z., Zhao, D., and Tao, J. (2021a). R2R3-MYB transcription factor PlMYB108 confers drought tolerance in herbaceous peony (*Paeonia lactiflora* pall.). *Int. J. Mol. Sci.* 22 (21), 11884. doi: 10.3390/ijms222111884

Wu, Z., Li, T., Liu, X., Yuan, G., Hou, H., and Teng, N. (2021b). A novel R2R3-MYB transcription factor LlMYB305 from lilium longiflorum plays a positive role in thermotolerance *via* activating heat-protective genes. *Environ. Exp. Bot.* 184, 104399. doi: 10.1016/j.envexpbot.2021.104399

Yan, X., Ding, W., Wu, X., Wang, L., Yang, X., and Yue, Y. (2022). Insights into the MYB-related transcription factors involved in regulating floral aroma synthesis in sweet osmanthus. *Front. Plant Sci.* 13, 765213–765213. doi: 10.3389/fpls.2022.765213

Yang, K., Li, Y., Wang, S., Xu, X., Sun, H., Zhao, H., et al. (2019). Genome-wide identification and expression analysis of the MYB transcription factor in moso bamboo (*Phyllostachys edulis*). *PeerJ* 6, e6242. doi: 10.7717/peerj.6242

Yang, J., Moeinzadeh, M., Kuhl, H., Helmuth, J., Xiao, P., Haas, S., et al. (2017). Haplotype-resolved sweet potato genome traces back its hexaploidization history. *Nat. Plants* 3 (9), 696–703. doi: 10.1038/s41477-017-0002-z

Ye, J., McGinnis, S., and Madden, T. L. (2006). BLAST: improvements for better sequence analysis. *Nucleic Acids Res.* 34 (Web Server issue), W6–W9. doi: 10.1093/nar/gkl164

Yu, J., Zhao, G., Li, W., Zhang, Y., Wang, P., Fu, A., et al. (2021). A single nucleotide polymorphism in an R2R3 MYB transcription factor gene triggers the male sterility in soybean *ms6* (Ames1). *Theor. Appl. Genet.* 134 (11), 3661–3674. doi: 10.1007/s00122-021-03920-0

Zhai, H., Wang, F., Si, Z., Huo, J., Xing, L., An, Y., et al. (2016). A *myo*-inositol-1-phosphate synthase gene, *IbMIPS1*, enhances salt and drought tolerance and stem

nematode resistance in transgenic sweet potato. *Plant Biotechnol. J.* 14 (2), 592–602. doi: 10.1111/pbi.12402

Zhang, H., Liu, Z., Luo, R., Sun, Y., Yang, C., Li, X., et al. (2022). Genome-wide characterization, identification and expression profile of MYB transcription factor gene family during abiotic and biotic stresses in mango (*Mangifera indica*). *Plants* 11 (22), 3141. doi: 10.3390/plants11223141

Zhao, P., Hou, S., Guo, X., Jia, J., Yang, W., Liu, Z., et al. (2019). A MYB-related transcription factor from sheepgrass, LcMYB2, promotes seed germination and root growth under drought stress. *BMC Plant Biol.* 19 (1), 1–15. doi: 10.1186/s12870-019-2159-2

Zhao, Y., Zhang, G., Tang, Q., Song, W., Gao, Q., Xiang, G., et al. (2022b). EbMYBP1, a R2R3-MYB transcription factor, promotes flavonoid biosynthesis in *Erigeron breviscapus*. *Front. Plant Sci.* 13, 946827. doi: 10.3389/fpls.2022.946827

Zhao, H., Zhao, H., Hu, Y., Zhang, S., He, S., Zhang, H., et al. (2022a). Expression of the sweet potato MYB transcription factor IbMYB48 confers salt and drought tolerance in arabidopsis. *Genes* 13 (10), 1883. doi: 10.3390/genes13101883

Zheng, J., Zhang, Z., Tong, T., Fang, Y., Zhang, X., Niu, C., et al. (2021). Genome-wide identification of WRKY gene family and expression analysis under abiotic stress in barley. *Agronomy* 11 (3), 521. doi: 10.3390/agronomy11030521

Zhou, C., Chen, Y., Wu, Z., Lu, W., Han, J., Wu, P., et al. (2015). Genome-wide analysis of the MYB gene family in physic nut (*Jatropha curcas* l.). *Gene* 572 (1), 63–71. doi: 10.1016/j.gene.2015.06.072

Zhou, L., Yarra, R., Jin, L., and Cao, H. (2020). Genome-wide identification and expression analysis of MYB gene family in oil palm (*Elaeis guineensis* jacq.) under abiotic stress conditions. *Environ. Exp. Bot.* 180, 104245. doi: 10.1016/j.envexpbot.2020.104245

Zhou, Y., Zhu, H., He, S., Zhai, H., Zhao, N., Xing, S., et al. (2019). A novel sweetpotato transcription factor gene *IbMYB116* enhances drought tolerance in transgenic arabidopsis. *Front. Plant Sci.* 10, 1025. doi: 10.3389/fpls.2019.01025

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 30 May 2023
DOI 10.3389/fpls.2023.1158329

OPEN ACCESS

EDITED BY
Kaixuan Zhang,
Chinese Academy of Agricultural Sciences, China

REVIEWED BY
Aziz Khan Khan,
Guangxi University, China
Zhenhua Wang,
Shihezi University, China
Abdul Rehman,
Zhengzhou University, China

*CORRESPONDENCE
Wangfeng Zhang
zhwf_agr@shzu.edu.cn

RECEIVED 03 February 2023
ACCEPTED 04 May 2023
PUBLISHED 30 May 2023

CITATION
Zuo W, Wu B, Wang Y, Xu S, Tian J, Jiu X, Dong H and Zhang W (2023) Optimal planting pattern of cotton is regulated by irrigation amount under mulch drip irrigation. *Front. Plant Sci.* 14:1158329. doi: 10.3389/fpls.2023.1158329

Optimal planting pattern of cotton is regulated by irrigation amount under mulch drip irrigation

Wenqing Zuo[1], Baojian Wu[1], Yuxuan Wang[1], Shouzhen Xu[1], Jingshan Tian[1], Xingli Jiu[2], Hengyi Dong[2] and Wangfeng Zhang[1]*

[1]Key Laboratory of Oasis Eco–Agriculture, Xinjiang Production and Construction Corps, College of Agronomy, Shihezi University, Shihezi, Xinjiang, China, [2]Regimental Farm 149, Division Eight, Xinjiang Production and Construction Corps, Shihezi, China

Objective: It is of great importance to explore agronomic management measures for water conservation and cotton yield in arid areas.

Methods: A four–year field experiment was conducted to evaluate cotton yield and soil water consumption under four row spacing configurations (high/low density with 66+10 cm wide, narrow row spacing, RS_{66+10H} and RS_{66+10L}; high/low density with 76 cm equal row spacing, $RS_{76}H$ and $RS_{76}L$) and two irrigation amounts (CI:conventional drip irrigation; LI:limited drip irrigation) during the growing seasons in Shihezi, Xinjiang.

Results: A quadratic relationship was observed between the maximum LAI (LAI_{max}) and seed yield. Canopy apparent transpiration rate(CAT), daily water consumption intensity (DWCI) and crop evapotranspiration (ET_C) were positively and linearly correlated with LAI. The seed yields, lint yields, and ET_C under CI were 6.6–18.3%,7.1–20.8% and 22.9–32.6%higher than those observed under LI, respectively. The RS_{66+10H} under CI had the highest seed and lint yields. $RS_{76}L$ had an optimum LAI_{max} range, which ensured a higher canopy apparent photosynthesis and daily dry matter accumulation and reached the same yield level as RS_{66+10H}; however, soil water consumption in $RS_{76}L$ was reduced ET_C by 51–60 mm at a depth of 20–60 cm at a radius of 19–38 cm from the cotton row, and water use efficiency increased by 5.6–8.3%compared to RS_{66+10H} under CI.

Conclusion: A $5.0<LAI_{max}<5.5$ is optimum for cotton production in northern Xinjiang, and $RS_{76}L$ under CI is recommended for high yield and can further reduce water consumption. Under LI, the seed and lint yield of RS_{66+10H} were 3.7–6.0% and 4.6–6.9% higher than those of $RS_{76}L$, respectively. In addition, high-density planting can exploit the potential of soil water to increase cotton yields under water shortage conditions.

KEYWORDS

mulch drip irrigation, row spacing configuration, cotton, leaf area index, soil water consumption

1 Introduction

Cotton (*Gossypium hirsutum* L.) is the most widely cultivated and vital fiber crop worldwide (Dai and Dong, 2014). It is mainly planted in the arid areas of China, the United States, Australia, Pakistan, and India (Tian et al., 2017; Tabashnik and Carrière, 2019; Anwar et al., 2020). Insufficiency or deficit irrigation is the main obstacle to the sustainable development of cotton in arid regions (Forouzani and Karami, 2011; Wei et al., 2022). However, agricultural irrigation water accounts for more than 60% of total land water consumption (Qin et al., 2016; Rafiee and Kalhor, 2016). The increasing demand for crop yield resulting from population growth further exacerbates water shortages in arid areas (Neumann et al., 2011). Therefore, it is crucial to ensure that agricultural development measures in arid areas consider the regulation of field management, appropriate irrigation methods to improve the rational use of water resources, and high crop yield.

Augmenting planting density is an important cultivation practice for increasing crop yield (Zhang et al., 2004; Feng et al., 2017), because it can increase leaf area index (LAI) and the interception of light energy, resulting in higher canopy photosynthetic capacity (Zhang et al., 2004; Liao et al., 2022). As the largest cotton-producing region in China, Xinjiang has a favorable ecological environment for producing high quality and yield cotton, due to abundant sunshine, a dry climate, and large diurnal temperature differences (Li et al., 2017). The widespread application of mulch drip irrigation technology since the late 1990s has effectively improved the water resource efficiency of crop production in arid and semi–arid regions in northern China (Dai and Dong, 2014; Guo et al., 2021). In recent years, cotton planting density in Xinjiang has been stable at around 22.5×10^4 plant hm^{-2} (Hu et al., 2021) due to the breeding of new varieties (Wang et al., 2021a) and the rational use of growth regulators (Shi et al., 2022). A higher LAI combined with sufficient light in Xinjiang improved the effective interception of photosynthetic radiation (Feng et al., 2017; Wu et al., 2017), canopy photosynthesis, and biomass accumulation, reduces the number of bolls per plant, increases the number of bolls per population (Zhang et al., 2004), thus increasing cotton yield (Dong et al., 2006; Araus et al., 2021). Therefore, increasing planting density to improve above–ground LAI is an important measure to obtain high crop yield. However, a higher LAI may lead to mutual shading within the cotton canopy, thus affecting population photosynthetic productivity (Hu et al., 2021; Paul et al., 2021). Appropriate planting density can also improve cotton yield by improving dry matter accumulation and potassium fertilizer absorption (Khan et al., 2017a; Khan et al., 2017b). Studies on cotton LAI in the Americas and other major cotton–growing countries in Asia have shown that the optimum LAI for a higher cotton yield is between 4.0 and 5.0 (Kerby et al., 1990; Heitholt, 1994; Bilal et al., 2019). However, there is still no definite conclusion regarding the optimum cotton LAI range after machine–harvested planting was implemented in Xinjiang.

A high crop yield in arid areas should be accompanied by efficient utilization of water resources. Crop evapotranspiration (ET_C), which includes soil evaporation and crop transpiration, is a key component of water consumption in agricultural fields, accounting for more than 90% of agricultural water use (Hou et al., 2022). Moreover, a high planting density significantly increases the ET_C (Cui et al., 2018; Wang et al., 2021b). Therefore, higher irrigation volumes are required to meet the demand for higher yields under high-density planting (Kodur, 2017; Wu et al., 2017; Hernandez et al., 2021). Some studies have also shown that ET_C is related to leaf area, but a larger LAI did not lead to higher soil water consumption because of shading between leaves, although more bare land areas were covered (Rahman et al., 2018; Di et al., 2019). Therefore, it is necessary to adjust the planting density and irrigation amount to regulate the aboveground LAI in a suitable range to improve the photosynthetic rate and increase the dry population accumulation (Yao et al., 2017; Chen et al., 2019). The adoption of new irrigation practices, such as sub–membrane drip irrigation (Zou et al., 2020), deficit irrigation (Paul et al., 2021) and limited irrigation (Chen et al., 2019) are common irrigation practices for improving water use efficiency (WUE) in arid areas. Many scholars have shown that, proper irrigation kept the crop root system in the irrigated wet zone (Chen et al., 2018), improves root morphology and physiological activity (Luo et al., 2014), and facilitates rapid water uptake by the root system for upward transport through the main stem to supply upper ground growth (Chen et al., 2019).

The combination of mulch drip irrigation and high-density planting is an important technical measure for high cotton yields in Xinjiang (Sui et al., 2018; Guo et al., 2021). However, the increasing shortage and unbalanced distribution of water have severely restricted cotton production in this area (Li et al., 2021). Therefore, it is important to explore water saving strategies and high cotton yields by conducting research on agronomic technical measures based on drip irrigation projects. We hypothesize that under water deficit conditions, increasing planting density could maintain the photosynthetic productivity of cotton populations by maximizing the use of soil water and increasing the population LAI to ensure high cotton yields. While under water-sufficient conditions, low-density planting has the potential to optimize canopy LAI to achieve high photosynthetic productivity while reducing soil water consumption. We hypothesized that there would be an optimal planting pattern under different water supply conditions to achieve a combination of water savings and cotton yield. The objectives of this study were (a) to determine the effects of irrigation amount and row spacing configuration on LAI dynamics and population photosynthesis capacity during the cotton reproductive period, and (b) to clarify the population transpiration water consumption and soil water consumption and provide suitable field management measures for high yields and water conservation in arid areas.

2 Materials and methods

2.1 Study site and experimental design

A four–year field experiment was conducted at the 13th Company (45°12′N, 86°05′E, 380m a. s. l.) and 11th Company of

the 149th Regimental Farm of Sihezi (45°12′N, 86°06′E, 380m, a. s. l.) and the experimental sites of Shihezi University (45°19′N, 86°03′E, 482m a. s. l.) and Wulanwusu Agrometeorological Experiment Station of Shihezi (44°17′N, 85°49′E, 520m a. s. l.) during the 2016–2019 growing seasons, respectively. The four test sites are typical of temperate continental climates. The locations in which this trial was conducted were in accordance with conventional tillage. Weather data for the sites were obtained from the nearest meteorological station. Daily maximum temperature, minimum temperature, and rainfall from planting until harvest (April to October) for 4 years are shown in Figure 1. The soil texture, soil moisture content, soil bulk density, and soil nutrient content of the test area before sowing are shown in Table 1.

The experiment (two irrigation amount and four row spacing configurations) was arranged in a randomized complete block design with four replicates. Two irrigation amounts were applied to the main plots: the local conventional irrigation amount (CI) ranging from 510 to 600 mm adopting one film with three drip tapes, and limited irrigation (LI) adopting one film with two drip tapes (70% of the CI amount). This design was selected because large–scale drip irrigation cotton fields in Xinjiang adopt a rotation irrigation system, and the arrangement of one film with three-tapes as irrigation method has a greater water output per unit time, saving irrigation time and shortening the rotation cycle. The four row spacing configurations combined with planting density and row spacing were as follows: $RS_{66+10}H$ (66 + 10 cm row spacing with 26 plants m^{-2}; high density), $RS_{76}H$ (76 cm row spacing with 26 plants m^{-2}; high density), $RS_{76}L$ (76 cm row spacing with 13 plants m^{-2}; low density), and $RS_{66+10}L$ (66 + 10 cm row spacing with 13 plants m^{-2}; low density). Each subplot consisted of 12 (66 + 10 cm row spacing configuration) or 6 (76 cm row spacing configuration) ×10 m cotton plant rows with two 2.28 m – wide sheets of transparent plastic film (Figure 2). The diameter of the labyrinth drip irrigation tape was 12.5 mm; dripper flow rate was 2.2 $m^3\ h^{-1}$, and dripper spacing was 20 cm.

The test sites were set up in the fields of local farmers, and each treatment consisting of one film with two tapes had a fertilizer amount consistent with that of the treatment consisting of one film with three tapes by adding fertilization tanks. Except for uniform seedling watering, all other irrigation periods were applied following a rotational irrigation system. Drip irrigation was applied 8–10 times during the growth period (Table 2).

2.2 Soil water content and crop water use

Soil sample were excavated at 20 cm intervals (up to 80 cm deep soil profiles) by using a soil corer for the soil water content

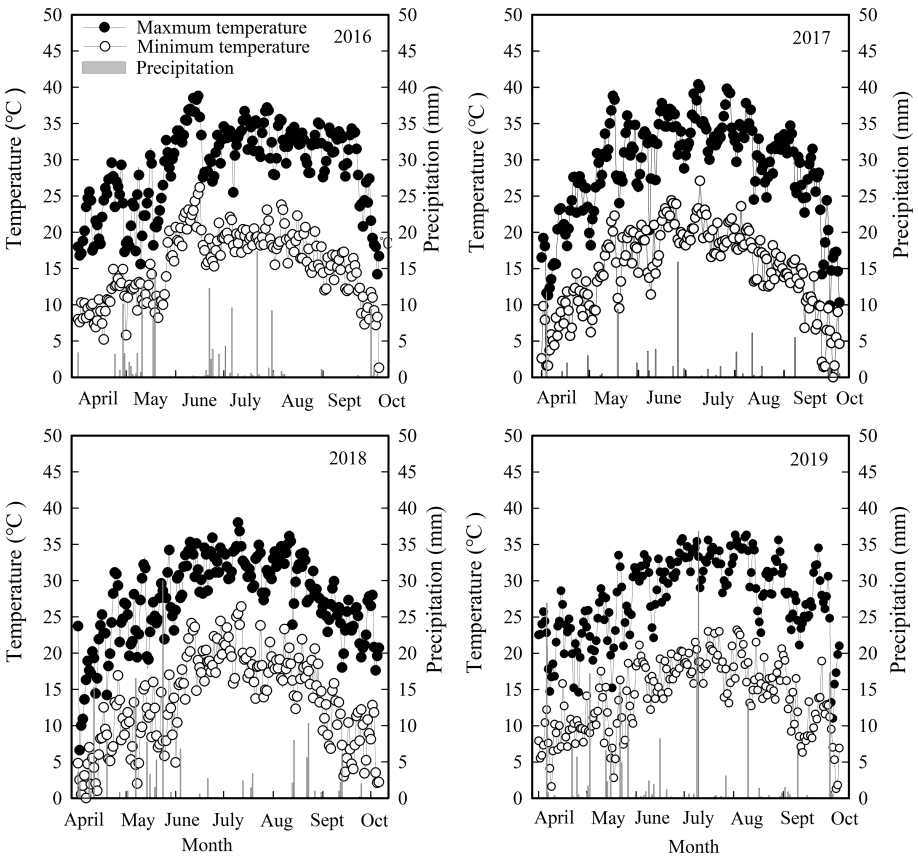


FIGURE 1
Daily maximum temperature, minimum temperature, and rainfall from planting until harvest (April to October) in 2016, 2017, 2018, and 2019.

TABLE 1 Soil texture and soil water content before sowing at different test sites used for evaluating the optimal planting pattern of cotton (*Gossypium hirsutum)* regulated by water amount under mulch drip irrigation in Xinjiang, China.

Station	Soil depth (cm)	Soil water content (%)	Bulk density (g·cm^{-3})	Texture	Alkali-hydro nitrogen (mg·kg^{-1})	Available phosphorus (mg·kg^{-1})	Available potassium (mg·kg^{-1})	Organic matter (g·kg^{-1})
13th Production unit^{2} # 149th Regimental Farm, Shihezi city, Xinjiang (2016)	0–20	8.9	1.35	Sandy loam	53.8	18.8	207.6	15.4
	20–40	11.3	1.29					
	40–60	12.8	1.28					
	60–80	10.2	1.31					
11th Production unit^{2} # 149th Regimental Farm, Shihezi city, Xinjiang (2017)	0–20	9.2	1.33	Loam	61.7	21.8	213.2	16.8
	20–40	12.1	1.28					
	40–60	13.5	1.32					
	60–80	11.2	1.35					
Experimental farm # Shihezi University, Shihezi city, Xinjiang (2018)	0–20	12.9	1.42	Gray desert	54.9	19.1	194.2	15.6
	20–40	13.8	1.28					
	40–60	13.3	1.41					
	60–80	13.4	1.43					
Wulanwusu Agrometeorological Experiment Station # Shihezi city, Xinjiang (2019)	0–20	13.5	1.38	Loam	58.9	21.1	188.7	15.3
	20–40	14.4	1.31					
	40–60	15	1.29					
	60–80	13.8	1.38					

(SWC) measurement (n = 3) in each experimental plot. The measurements were executed with a horizontal distance of 0, 19, and 38 cm from the cotton row at 1 d before sowing, 1 d before irrigation, 2 d after irrigation, and maturity. The samples were immediately weighed and then baked at 80°C in an oven to determine the soil moisture content (SMC). The specific calculation formula for the soil accumulation water consumption (SAWC, mm) for different soil layers as follows:

$$SAWC = \sum(SWC_{i+1} - SWC_i) \quad (1)$$

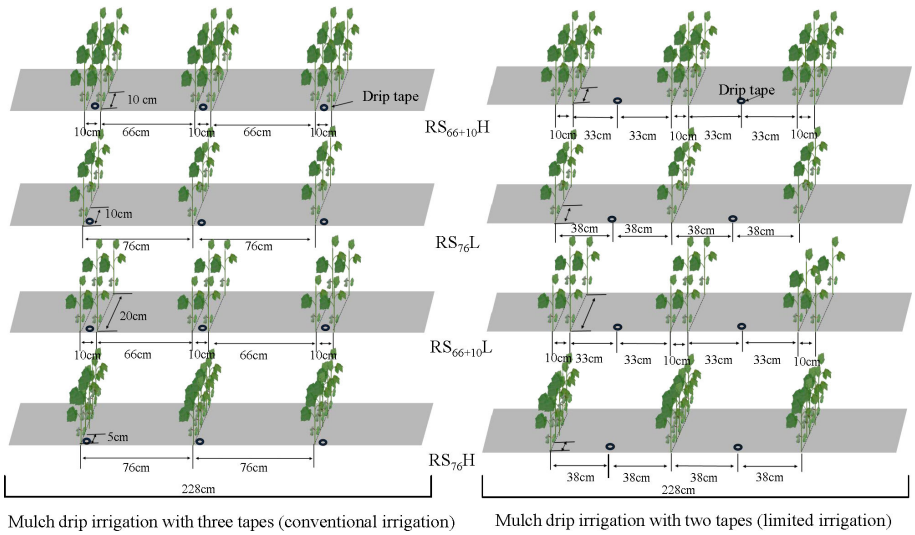


FIGURE 2
Schematic of four different planting densities of cotton (*Gossypium hirsutum*) evaluated during this study on cotton cultivation in Xinjiang, China. Where $RS_{66+10}H$ has 66+10 cm row spacing with 26 plants m^{-2} (high density), $RS_{76}H$, where 76 cm row spacing with 26 plants m^{-2} (high density), $RS_{76}L$ has 76 cm row spacing with 13 plants m^{-2} (low density), and $RS_{66+10}L$ (66+10 cm row spacing with 13 plants m^{-2} (low density).

TABLE 2 Sowing date, harvest period, total irrigation amount, and fertilization amount for different drip tape configurations in 2016, 2017, 2018, and 2019.

Year	Treatment	Seeding date (m/d)	Harvest date (m/d)	Irrigation amount (mm)	Fertilizer application (kg hm^{-2})		
					N	P	K
2016	CI	4/9	10/2	600	282.0	68.3	86.2
	LI	4/9	10/2	430	282.0	68.3	86.2
2017	CI	4/17	10/7	580	272.6	62.6	79
	LI	4/17	10/7	424	272.6	62.6	79
2018	CI	4/20	10/15	510	244.4	56.9	71.8
	LI	4/20	10/15	350	244.4	56.9	71.8
2019	CI	4/21	10/12	518	270.3	63.6	80.4
	LI	4/21	10/12	355	270.3	63.6	80.4

$$SWC = H \times SMC \times P \quad (2)$$

$$SMC = (M_0 - M_1)/M_1 \times 100\,\% \quad (3)$$

SWC_{i+1} and SWC_i means the soil water storage of one day before next irrigation and two days after irrigation respectively, in an irrigation cycle. H (mm) means the thickness of the soil layer; P (g·cm^{-3}) means the soil bulk density. M_0 and M_1means the wet weight of soil sample and the dried weight of soil sample, respectively.

The daily water consumption intensity (DWCI, mm d^{-1}) was determined by using Eq. (3) according to Wang et al. (2021b), which was used to identify the water consumption in different stages of cotton growth as follows:

$$DWCI = ET/\Delta T \quad (3)$$

where ET is the phase water consumption (mm) during a given growth period, and ΔT is the duration (d) of a given growth period.

Total crop water consumption, namely the actual evapotranspiration (ETc, mm), was calculated during the growing season as follows:

$$ETc = R + I - F - Q + \Delta W \quad (4)$$

where ETc is the crop evapotranspiration; where ETc (mm), R (mm), I (mm), F, and Q are the crop evapotranspiration, precipitation, irrigation amount, surface runoff, and capillary rise, respectively; ΔW is the change in SWC (mm). Q is the capillary rising to root zone, which is negligible due to the groundwater table of over 8 m at the experimental site. F could also be ignored at the experimental site.

2.3 Canopy apparent photosynthesis/ transpiration rate

The canopy apparent photosynthesis (CAP) and canopy apparent transpiration rates (CAT) were simultaneously measured using the assimilation chamber method (Reddy et al., 1995; Xie et al., 2010). The CO_2 and H_2O concentrations in the chamber were measured using a Li–840A Soil CO_2 Flux System (LI–COR Inc., Lincoln, NE, USA). The measurements were made between 11:00 and 14:00 h on clear days immediately after determining PAR. The assimilation chamber (85 cm long × 75 cm wide × 125 cm high) was covered with acrylic film that transmitted more than 95% of the solar radiation. Two fans were installed inside the chamber to mix the air. Gas exchange rates in each plot were measured during at least three 60 s intervals. We began to record the values when the CO_2 concentrations inside the chamber began to drop steadily. Measurements were repeated three times for each treatment. The CAP and CAT calculation formula is as follows:

$$CAP = \Delta C_1/10^{-6} \times V \times 360/\Delta M \times 273/(273 + T) \times 44/22.4 \times 1000/L \quad (5)$$

$$CAT = \Delta C_2/10^{-6} \times V \times 360/\Delta M \times 273/(273 + T) \times 44/22.4 \times 1000/L \quad (6)$$

where ΔC_1 represents net photosynthetic assimilation CO_2 concentration in a given time interval (s); ΔC_2 represents net photosynthetic assimilation H_2O concentration in a given time interval (s); V is the assimilation chamber volume (m^3); Δm is the measured time interval; T is the air temperature (°C); and L is the land area of the measured cotton canopy population.

2.4 Leaf area index

The leaf area index (LAI) was measured using the LAI–2200C canopy analyzer (Li–COR Inc., Lincoln, NE, USA) at 7–10 d intervals starting from early July, which referring to the method of Malone et al. (2002). Four to six readings were made in each plot. One measurement

was made above the canopy, and then four measurements were made perpendicular to the cotton rows at the soil surface.

2.5 Dry matter accumulation and yield

For individual plant measurements, cotton plants were randomly selected in each plot at the initial flowering stage, full flowering stage, boll stage, and boll opening stage. On each sample date, four plants at each plot were randomly selected to obtain an average value. Plants were divided into various organs including stems, leaves, buds, and bolls. These segments were subsequently placed in paper bags, dried at 80 °C in an oven until constant weight, and the dry weight was measured. Daily dry matter accumulation was calculated as follows:

$$DDMA = (DMA_{i+1} - DMA_i)/\Delta T \quad (7)$$

where DMA_{i+1} is the dried matter accumulation taken in the next growth period ($g{\cdot}m^{-2}{\cdot}d^{-1}$), and ΔT is the interval time for selecting dry matter accumulation. Seed cotton was hand harvested at 3 × 2.28 m^2 area (n=4) in each plot at maturity. All mature cotton bolls in the 2.28*3 area are collected before harvest to facilitate data veracity. Seed cotton yield (kg hm^{-2}) was determined for each plot after sun–drying for fifteen days, and then weighed after ginning to obtain the lint yield (kg hm^{-2}).

2.6 Statistical analysis

Random block analysis of variance (ANOVA) was used to assess the effects of irrigation amount and row spacing configurations on LAI, CAP, CAT, ET_C, DDMA, DWCI and seed/lint yield. Duncan's multiple range tests were used to separate the treatment means at P< 0.05. Correlation analysis was conducted among LAI and CAP, CAT, DWCI, ET_C, seed yield; DDMA and seed yield. Figures were constructed using the "lme4" and "ggplot2" packages in R 4.0.5 software (R Core Team 2021) and Sigmaplot 12.0 (Aspire Software Intl., Ashburn, USA).

3 Results

3.1 Cotton yield, and water use efficiency

The irrigation amount and row spacing configuration significantly affected daily dry matter accumulation (DDMA), seed yield, ET_C, and WUE of cotton (P< 0.05; Tables 3; 4). The DDMA

TABLE 3 Daily dry matter accumulation (DDMA) characteristics as affected by the combination of irrigation amount and row spacing configuration of cotton under mulch drip irrigation.

Irrigation amount	Row spacing	Daily dry matter accumulation (DDMA, $g{\cdot}m^{-2}\ d^{-1}$)											
		2016			2017			2018			2019		
		BFF-FF	FF-FB	FB-BO	BFF-FF	FF-FB	FB-BO	BFF-FF	FF-FB	FB-BO	BFF-FF	FF-FB	FB-BO
CI	$RS_{66+10}H$	24.5a	37.2a	27.9a	29.9a	38.0a	30.5a	27.3a	37.8a	30.2a	29.7a	34.8a	28.1a
	$RS_{76}H$	24.2a	36.5a	28.1a	28.3a	39.6a	27.6ab	28.6a	36.8a	29.0a	28.5a	32.9a	27.1a
	$RS_{76}L$	22.4bc	32.0b	24.2bc	26.4b	35.5b	24.2c	24.3b	34.3b	26.5b	26.9b	30.2b	24.9b
	$RS_{66+10}L$	23.3b	34.5ab	25.5b	25.5b	36.9b	24.9c	25.2b	35.1b	26.6b	27.3b	30.9b	25.3b
LI	$RS_{66+10}H$	17.9de	26.9c	22.7cd	26.4b	29.1c	22.4d	23.4bc	34.0bc	24.3bc	26.1b	28.7bc	22.0c
	$RS_{76}H$	19.2d	27.5c	23.1c	25.9b	30.1c	21.1d	23.0bc	33.3bc	24.6bc	26.3b	27.7c	21.8c
	$RS_{76}L$	14.8f	21.7de	17.9e	23.0c	27.2cb	18.1e	21.7d	30.6d	21.8d	23.3c	25.8d	20.9d
	$RS_{66+10}L$	15.9f	23.0d	17.1e	23.4c	26.8c	19.0e	21.1d	31.8d	21.5d	25.0c	27.0d	21.0d
I (Irrigation amount)		**	**	**	**	**	**	**	**	**	**	**	**
R (row)		NS	**	NS	**	**	**	**	NS	**	**	**	**
D (Density)		**	**	**	**	**	**	**	**	**	**	**	**
I×R		NS	NS	NS	**	**	**	**	**	**	**	**	**
I×D		NS	*	*	**	**	**	**	NS	**	**	**	**
R×D		**	**	**	**	**	**	**	**	**	**	**	**
I×R×D		NS	*	**	**	NS	**	**	NS	**	**	**	**

BFF, FF, FB, and BO means before full flowering, full flowering, full boll and boll opening of cotton growth stage, respectively. I means irrigation amount; R means row spacing configuration; D means plant density; CI means conventional irrigation; LI means limited irrigation; $RS_{66+10}H$ and $RS_{66+10}L$ mean high/low-density planting with 66 + 10 cm row spacing configuration, respectively; $RS_{76}H$ and $RS_{76}L$ mean high/low-density planting with 76 cm row spacing configuration, respectively. Values are means ± SD (n=4).* Significant at P ≤ 0.05; ** Significant at P ≤ 0.01; NS, not significant.

Values followed by different lowercase letters are significantly different at the 0.05 probability level.

TABLE 4 Seed yield, lint yield, crop evapotranspiration, and water use sufficiency as affected by the combination of irrigation amount and row spacing configuration of cotton under mulch drip irrigation.

Irrigation amount	Row spacing	Seed yield (kg·hm^{-2})				Crop evapotranspiration (ETc, mm)				Water use efficiency (kg·m^{-3})				Lint yield (kg·hm^{-2})
		2016	2017	2018	2019	2016	2017	2018	2019	2016	2017	2018	2019	
CI	$RS_{66+10}H$	7358 ± 109a	7814 ± 85a	7576 ± 87a	7362 ± 80a	683 ± 10a	627 ± 12a	623 ± 13a	619 ± 6a	1.08 ± 0.02c	1.25 ± 0.01c	1.22 ± 0.01c	1.19 ± 0.01d	3389 ± 178a
	$RS_{76}H$	7173 ± 175ab	7683 ± 63ab	7490 ± 73a	7289 ± 73a	681 ± 12a	631 ± 11a	624 ± 10a	621 ± 10a	1.05 ± 0.03c	1.22 ± 0.01cd	1.20 ± 0.01cd	1.17 ± 0.01d	3216 ± 203ab
	$RS_{76}L$	7249 ± 14a	7711 ± 79a	7470 ± 20a	7276 ± 59a	623 ± 13b	571 ± 21bc	572 ± 17bc	568 ± 11cd	1.16 ± 0.00b	1.34 ± 0.01b	1.29 ± 0.00b	1.27 ± 0.01b	3315 ± 182a
	$RS_{66+10}L$	7135 ± 156b	7573 ± 75b	7369 ± 30b	7207 ± 92a	639 ± 17b	591 ± 12b	595 ± 13b	582 ± 12c	1.12 ± 0.03b	1.26 ± 0.01c	1.23 ± 0.01c	1.24 ± 0.02bc	3173 ± 109ab
LI	$RS_{66+10}H$	6846 ± 37c	6866 ± 35c	6898 ± 38c	6881 ± 62c	531 ± 12c	484 ± 13d	477 ± 11d	467 ± 11e	1.25 ± 0.01a	1.40 ± 0.01b	1.44 ± 0.02a	1.47 ± 0.01a	3021 ± 81c
	$RS_{76}H$	6689 ± 138cd	6732 ± 62cd	6930 ± 95c	6841 ± 127c	528 ± 18c	486 ± 13d	477 ± 10d	472 ± 9e	1.27 ± 0.03a	1.39 ± 0.01b	1.45 ± 0.01a	1.45 ± 0.02a	2986 ± 156c
	$RS_{76}L$	6468 ± 48d	6589 ± 74d	6731 ± 96d	6606 ± 42d	507 ± 8d	455 ± 17 e	460 ± 13de	450 ± 11f	1.28 ± 0.01a	1.46 ± 0.01a	1.46 ± 0.02a	1.47 ± 0.01a	2826 ± 88d
	$RS_{66+10}L$	6519 ± 57cd	6318 ± 85e	6728 ± 14d	6623 ± 74d	520 ± 12cd	469 ± 12de	467 ± 14d	452 ± 15f	1.25 ± 0.01a	1.35 ± 0.02b	1.44 ± 0.00a	1.47 ± 0.02a	2810 ± 20
I (Irrigation amount)		**	**	**	**	**	**	**	**	**	**	**	**	**
R (row)		*	*	NS	NS	**	**	**	**	*	**	**	*	*
D (Density)		NS	**	NS	NS	**	**	**	**	*	**	**	*	NS
I×R		NS	NS	*	NS	NS	NS	**	**	*	NS	**	**	*
I×D		NS	NS	NS	NS	NS	NS	*	NS	NS	*	NS	NS	*
R×D		**	**	**	**	**	**	**	**	**	**	**	**	**
I×R×D		NS	*	NS	NS	**	**	**	**	**	**	**	**	NS

Lint yield data presented is a four-year average; CI means conventional irrigation; LI means limited irrigation; $RS_{66+10}H$ and $RS_{66+10}L$ mean high/low-density planting with 66 + 10 cm row spacing configuration, respectively; $RS_{76}H$ and $RS_{76}L$ mean high/low-density planting with 76 cm row spacing configuration, respectively. I means irrigation amount; R means row spacing configuration; D means plant density. Values are means ± SD (n=4).* Significant at $P \leq 0.05$; ** Significant at $P \leq 0.01$; NS, not significant.
Values followed by different lowercase letters are significantly different at the 0.05 probability level.

under CI was 9.0–51.4% [before full flowering (BFF) – full flowering (FF)], 10.5–50.0% [FF – full boll stage (FB)] and 19.1–49.1% [FB –boll opening stage (BO)] greater than that of the same treatment under LI ($P< 0.05$). The seed yields, lint yields, and ET_C under CI were 6.6–18.3%, 7.8–14.3% and 22.9–32.6% higher, respectively, but the WUE was 6.2–19.0% lower than that of the same treatment under LI. The seed and lint yields of $RS_{66+10}H$ were 3.7–6.0% and 4.6–6.9% higher than those of $RS_{76}L$ under LI, respectively, while there was no significant difference in WUE ($P > 0.05$). Notably, under CI, the seed and lint yields of $RS_{76}L$ were not different from those of $RS_{66+10}H$ ($P > 0.05$), but ET_C was reduced by 51–60 mm and WUE was increased by 5.6–8.3% compared to $RS_{66+10}H$

3.2 Soil accumulated water consumption in different soil layers

Under CI (Figure 3), the drip tapes were placed close to the cotton rows, and the trend of SAWC variation in each soil layer at a horizontal distance of 0–38 cm from the cotton row was in the following order:0 cm > 19 cm> 38 cm. Under LI, the trend variation of SAWC in each soil layer at 0–38 cm horizontal distance from the cotton rows was19 cm > 0 cm > 38 cm because the drip tapes were placed in the middle of the wide rows. The SAWC of the 0 –60 cm soil layer within 0, 19, and 38 cm radius from the horizontal distance of cotton rows under CI increased by 40.7–45.4%, 15.2–18.1% and 27.5–32.7% when compared to the same row spacing configuration under LI ($P< 0.05$). The highest SAWC was found in each soil layer in $RS6_{6+10}H$ under the same irrigation amount (no difference between $RS_{66+10}H$ and $RS_{76}H$). Under LI, the SAWC of $RS_{66+10}H$ was 4.2–6.9% and 6.8–11.4% greater than that of $RS_{76}L$ in the 20 – 40 cm and 40 – 60 cm soils at 19 cm from the horizontal distance of the cotton rows, respectively, and 10.6–11.1% and 13.1–13.2% greater than $RS_{76}L$ in the 20–40 cm and 40–60 cm soil layers at 38 cm from the horizontal distance of the cotton rows, respectively. Under CI, the SAWC of $RS_{76}L$ was 9.8–12.4% and 10.8–15.7% lower than $RS_{66+10}H$ in the 20–40 cm and 40–60 cm

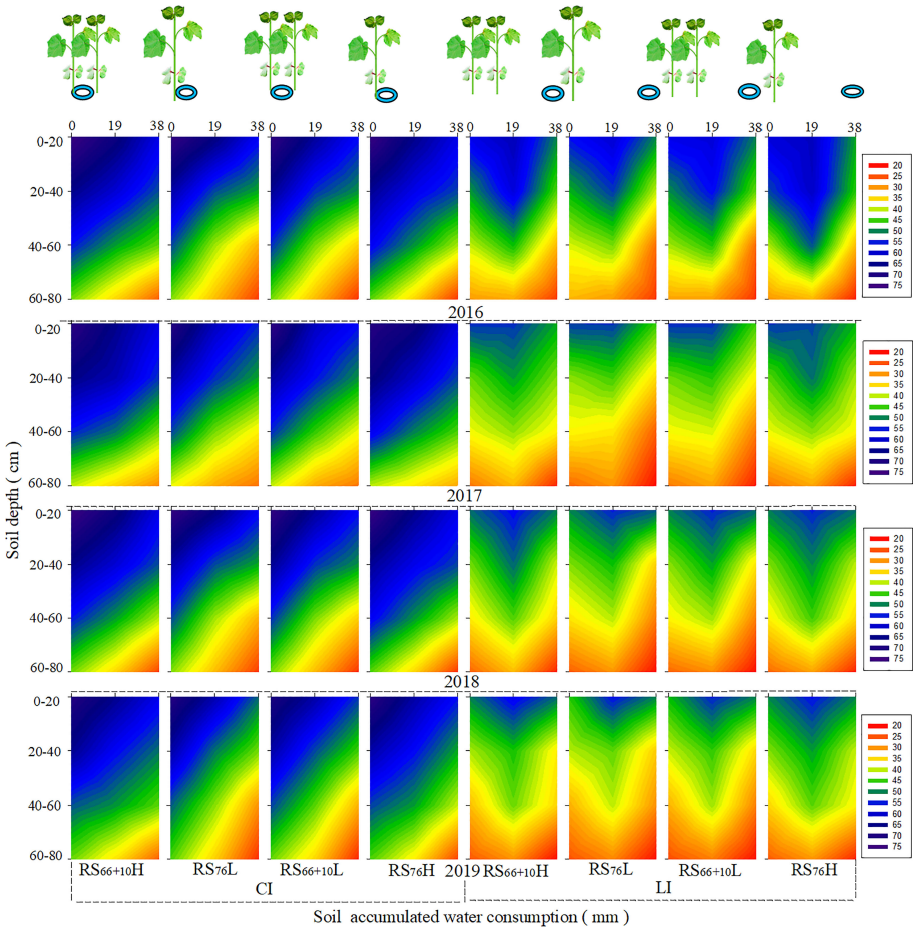


FIGURE 3
Soil accumulated water consumption (SAWC) in different soil layers under a combination of irrigation amount and row spacing configuration. The color changes from red to dark blue indicate a gradual increase in water consumption.

soil layers at a horizontal distance of 19 cm from the cotton rows, respectively: 7.7–16.8% and 9.1%–14.8% lower than $RS_{66+10}H$ in the 20–40 cm and 40–60 cm soil layers at a horizontal distance of 38 cm from the cotton rows, respectively.

3.3 Daily water consumption intensity and canopy apparent transpiration rate

The DWCI of cotton was significantly affected by the irrigation amount and row spacing configuration (Table 5, $P< 0.05$). The DWCI under CI was 11.5–30.2% (FF), 6.1–50.1%(FB) and 56.0–106.1%(BO) higher than that of the same row spacing under LI. The DWCI of $RS_{76}L$ was significantly lower than that of $RS_{66+10}H$ and $RS_{76}H$ under the same irrigation level, especially the largest difference during BO, where $RS_{76}L$ was 6.5–20.5% and 3.1–12.6% lower than $RS_{76}H$ and $RS_{66+10}H$, respectively.

The CAT of cotton under CI was significantly higher ($P< 0.05$; Figure 4) than that of cotton with the same row spacing configuration under LI, especially 11.9–33.9% higher in FB. $RS_{76}L$ had the lowest CAT activity in all fertility periods under the same irrigation amount. Under LI, the CAT of $RS_{66+10}H$ was 9.2–23.5% and 3.8–22.3% higher than that of $RS_{76}L$. Under CI, The CAT of $RS_{76}L$ was 10.1–30.6% and 3.9–22.2% lower than that of $RS_{66+10}H$ and $RS_{76}H$, respectively, throughout the critical reproductive period.

3.4 Canopy apparent photosynthesis

The irrigation amount was found to significantly affect cotton CAP ($P< 0.05$; Table 6). The CAP under CI was 5.2–16.7%(FF), 8.8–23.4%(FB), and 20.7–71.6% (BO) higher than that under LI. $RS_{66+10}H$ had the largest CAP from FF to BO under the same irrigation amount. Under LI, $RS_{66+10}H$ had 2.8–5.1% and 19.8–45.7% higher CAP than $RS_{76}L$ in FB and BO, respectively. However, the CAP of $RS_{76}L$ was not significantly different from that of $RS_{66+10}H$ under CI ($P > 0.05$).

3.5 Leaf area index

The LAI under CI was significantly greater than that under LI and was 9.5–22.2% greater at the FB (Figure 5). Under the same

10.3389/fpls.2023.1158329

TABLE 5 Daily water consumption intensity as affected by the combination of irrigation amount and row spacing configuration of cotton under mulch drip irrigation.

Irrigation amount	Planting pattern	Daily water consumption intensity (DWCI, mmd^{-1})											
		2016			2017			2018			2019		
		FF	FB	BO	FF	FB	BO	FF	FB	BO	FF	FB	BO
CI	$RS_{66+10}H$	5.7a	8.2a	3.1a	6.3a	8.1a	2.5a	6.7a	7.2a	3.1a	6.4a	7.4a	3.1a
	$RS_{76}H$	5.9a	8.3a	3.1a	6.3a	8.0a	2.6a	6.6a	7.4a	3.2a	6.5a	7.4a	3.2a
	$RS_{76}L$	5.2b	7.5b	2.7b	5.8c	7.4c	2.2b	6.1b	6.6b	2.9b	5.9b	6.9b	2.9ab
	$RS_{66+10}L$	5.3ab	7.8ab	2.9ab	6.0ab	7.7ab	2.3ab	6.3ab	6.8ab	3.0ab	6.0ab	7.1ab	3.0a
LI	$RS_{66+10}H$	4.9bc	7.5b	1.8b	4.9d	6.6d	1.3c	5.2c	5.0c	1.8c	5.0c	5.5c	2.0c
	$RS_{76}H$	5.0b	7.6b	1.7b	4.9d	6.5d	1.3c	5.3c	4.9c	1.7c	5.1c	5.6c	1.9c
	$RS_{76}L$	4.7c	7.1bc	1.5bc	4.6cd	6.2cd	1.2cd	4.8d	4.6d	1.6d	4.7cd	5.2cd	1.7cd
	$RS_{66+10}L$	4.7c	7.3bc	1.6bc	4.7cd	6.3cd	1.3c	4.9cd	4.7cd	1.7cd	5.0c	5.2cd	1.8cd
I (Irrigation amount)		**	**	**	**	**	**	**	**	**	**	**	**
R (row)		**	**	NS	**	**	NS	**	**	NS	**	NS	**
D (Density)		NS	NS	NS	**	**	**	**	**	*	**	**	**
I×R		**	NS	**	NS	**	**	**	**	*	**	**	**
I×D		**	NS	NS	**	**	**	NS	**	NS	**	NS	**
R×D		**	**	**	**	**	**	**	**	**	**	**	**

FF, FB, and BO mean full flowering, full boll and boll opening of cotton growth stage, respectively. CI means conventional irrigation; LI means limited irrigation; $RS_{66+10}H$ and $RS_{66+10}L$ mean high/low-density planting with 66 + 10 cm row spacing configuration, respectively; $RS_{76}H$ and $RS_{76}L$ mean high/low-density planting with 76 cm row spacing configuration, respectively. I means irrigation amount; R means row spacing configuration; D means plant density. Values are means ± SD (n=4).* Significant at $P \leq 0.05$; ** Significant at $P \leq 0.01$; NS, not significant.
Values followed by different lowercase letters are significantly different at the 0.05 probability level.

irrigation level, the trend variation of cotton LAI was $RS_{66+10}H$, $RS_{76}H > RS_{66+10}L > RS_{76}L$. The LAI of $RS_{66+10}H$ under LI was 5.1–25.8% and 7.8–32.3% higher than that of $RS_{76}L$ throughout the critical reproductive period. Under CI, $RS_{76}L$ was 6.0–17.1% and 7.3–16.8% lower than $RS_{66+10}H$ and $RS_{76}H$, respectively, throughout the critical reproductive period from four replicates.

3.6 Relationship between LAI, yield and water consumption

The relationship between LAI_{max} and seed yield was fitted to a quadratic function ($P< 0.01$; Figure 6), and the maximum cotton seed yield (7366 kg hm^{-2}) was obtained when LAI_{max} was approximately 5.5. However, based on the correlation between LAI and CAP, after LAI_{max} reached 5, further increases in LAI did not significantly increase CAP, which indicated that a higher LAI (peak >5) did not significantly increase cotton CAP and seed yields. There was a highly significant positive correlation between CAT, DWCI, ETc, and LAI ($P< 0.01$). This indicates that the LAI is a principal factor affecting cotton transpiration water consumption under mulch drip irrigation. Therefore, combine maintaining cotton high yields and water conservation, the optimal cotton LAI_{max} range should be between 5.0 and 5.5 (Figures 6C, F).

4 Discussion

4.1 Leaf area index, photosynthetic rate, and cotton yield

The first objective of this study was to evaluate the effect of irrigation amount and row spacing configuration on LAI, photosynthesis, dry matter accumulation rate, and yield. Adjusting the planting density and row spacing configuration is an important agronomic measure for achieving high and stable cotton yields (Zhang et al., 2004; Brodrick et al., 2010; Chen et al., 2019). An important condition for achieving high cotton yields in cotton production areas with short frost–free periods and limited light and heat resources is to achieve a high rate of dry matter accumulation per unit area. This study showed that cotton DDMA was higher under CI than under the same row spacing configuration with LI (Table 3), indicating that higher dry matter accumulation rates and higher yields were obtained under CI conditions during shorter reproductive periods (Figure 6A). A suitable LAI range is essential for the rapid growth of cotton dry matter accumulation (Srinivasan et al., 2017). Hu et al. (2021) concluded that a larger LAI caused canopy shading between leaves, resulting in lower CAP and lower cotton yield levels. This study showed a highly significant quadratic relationship between cotton LAI_{max} and seed yield (Figure 6C). It had the highest yield when LAI_{max} reached

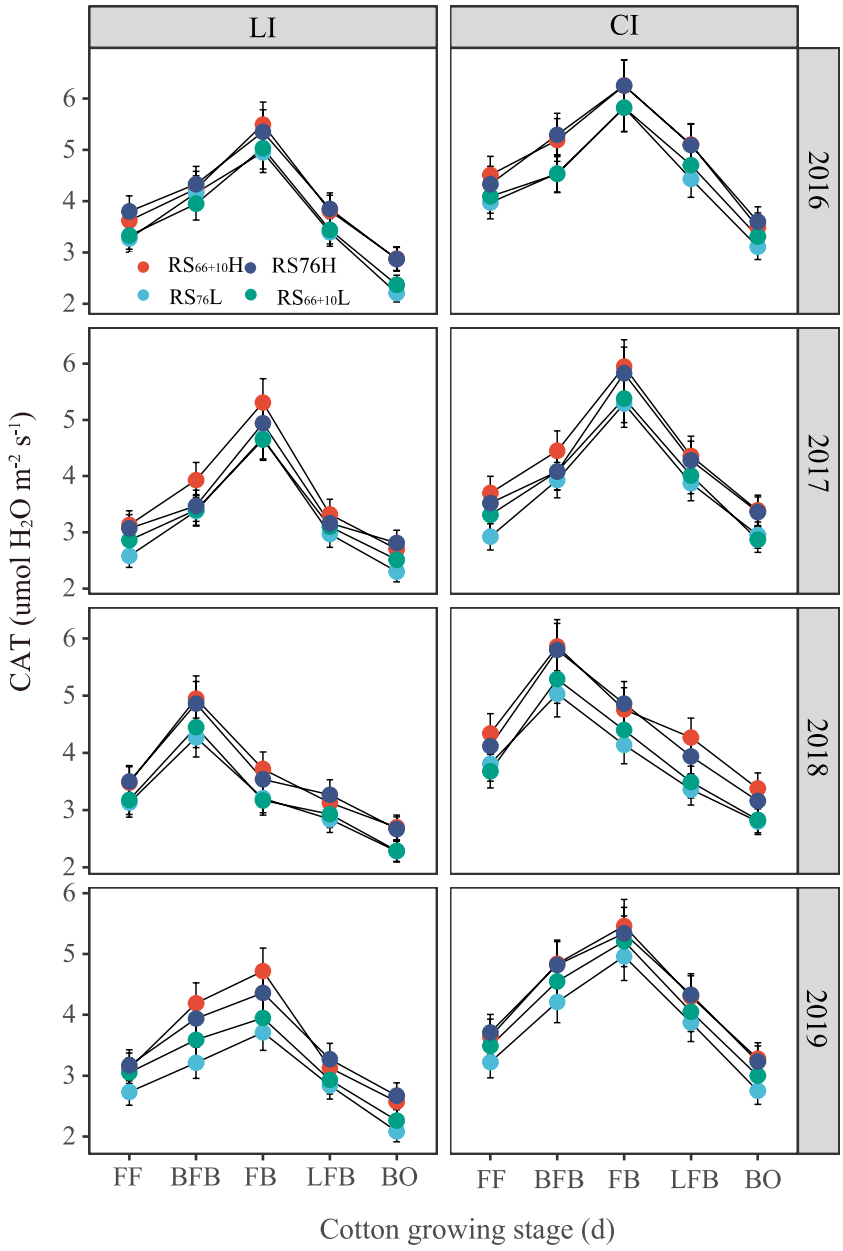


FIGURE 4
Canopy apparent transpiration rate (CAT) affected by the combination of irrigation amount and row spacing configuration. FF, BFB, FB, LFB and BO means full flowering, before full boll, full boll, later full boll and boll opening stage, respectively. Vertical bars represent the standard error. Mean values ± SE are from four replicates.

approximately 5.4 (Figure 5). Analysis of the relationship between LAI and CAP (Figure 6F) showed that CAP increased significantly when 0< LAI_{max}< 5. However, when LAI_{max} > 5, CAP did not increase significantly, which may be related to mutual shading between the groups. In combination with the CAP of treatments under CI, the cotton population maintained a relatively stable and high CAP when LAI_{max} between 5.0 and 6.0. This may be attributed to the adoption of densely tolerant cotton varieties in Xinjiang and the optimization of canopy structure through chemical regulation to shape compact plants (Wang et al., 2021a; Shi et al., 2022). However, $RS_{76}L$ under CI reached the same yield level as $RS_{66+10}H$, however its LAI_{max} was between 5.0 and 5.5, which was significantly smaller than $RS_{66+10}H$ and $RS_{76}H$. Combining LAI_{max}, CAP, and cotton yield, the appropriate LAI_{max} for achieving a high cotton yield in Xinjiang was between 5.0 and 5.5.

The cotton regions of the Yellow River Basin and Yangtze River Basin in China have a long cotton fertility period. The suitable planting density was 50000–60000 plants hm^{-2} and the largest cotton seed yield was 3700–4500 kg hm^{-2} in the Yellow River basin (Li et al., 2020); The optimum density was 19500–37500 plants hm^{-2} and the largest seed yield was 3800–4200 kg hm^{-2} in the Yangtze River basin (Lv et al., 2021). Based on the literature, our analysis of the relationship between cotton LAI and CAP in the Yangtze and Yellow River basins revealed a quadratic relationship

TABLE 6 Canopy apparent photosynthesis as affected by the combination of irrigation amount and row spacing configuration of cotton under mulch drip irrigation.

Irrigation amount	Planting pattern	Canopy apparent photosynthesis (CAP, μ mol $m^{-2}s^{-1}$)											
		2016			2017			2018			2019		
		FF	FB	BO	FF	FB	BO	FF	FB	BO	FF	FB	BO
CI	$RS_{66+10}H$	30.3ab	37.3a	15.4a	30.4a	36.3a	14.0a	31.8a	30.8a	18.6a	29.2a	35.7a	18.3a
	$RS_{76}H$	30.6ab	35.7ab	13.7ab	29.6a	34.5ab	13.5ab	30.5ab	28.9ab	17.3ab	28.1ab	36.4a	16.2b
	$RS_{76}L$	31.4a	36.0ab	12.6bc	29.3a	34.2ab	13.7a	31.1ab	28.4ab	17.6ab	29.0a	35.3ab	17.1ab
	$RS_{66+10}L$	32.9a	36.2ab	13.9ab	28.3ab	35.0ab	12.8b	32.5a	29.0ab	17.1ab	29.3a	34.5b	15.9b
LI	$RS_{66+10}H$	28.6c	33.1c	10.9c	26.9cd	30.2cd	11.6c	30.4b	24.8c	15.3c	28.5ab	32.4cd	14.0c
	$RS_{76}H$	27.6c	32.3c	11.1c	25.9d	31.7c	10.6c	29.0c	23.8cd	13.2d	27.5b	33.4c	13.2d
	$RS_{76}L$	28.3c	31.7d	9.1d	25.9d	29.5d	9.0cd	30.2b	24.6c	10.5e	27.2b	30.9e	11.2e
	$RS_{66+10}L$	28.2c	32.3c	8.1de	27.6c	30.9c	8.3d	29.7bc	23.5cd	11.0e	28.5ab	31.5d	10.2d
I (Irrigation amount)		**	**	**	**	**	**	**	**	**	**	**	**
R (row)		**	**	**	**	**	**	**	**	**	**	**	**
D (Density)		**	**	**	**	**	**	**	**	**	**	**	**
I×R		**	**	**	**	**	**	**	**	**	**	**	**
I×D		**	**	**	**	**	NS	**	**	**	**	**	**
R×D		**	**	**	**	**	**	**	**	**	**	**	**

FF, FB, and BO mean full flowering, full boll and boll opening of cotton growth stage, respectively. CI means conventional irrigation; LI means limited irrigation; $RS_{66+10}H$ and $RS_{66+10}L$ mean high/low-density planting with 66 + 10 cm row spacing configuration, respectively; $RS_{76}H$ and $RS_{76}L$ mean high/low-density planting with 76 cm row spacing configuration, respectively. I means irrigation amount; R means row spacing configuration; D means plant density. Values are means ± SD (n=4).* Significant at $P \leq 0.05$; ** Significant at $P \leq 0.01$; NS, not significant.
Values followed by different lowercase letters are significantly different at the 0.05 probability level.

(Figure S1) and the optimum LAI_{max} was between 3.5 and 4.0. Therefore, in cotton areas with a short reproductive period, higher cotton population photosynthetic capacity and higher yield could be achieved by using a combination of adequate irrigation with appropriate low-density row spacing, or with high density to improve LAI under limited drip irrigation.

4.2 Optimal planting pattern of cotton is regulated by the local water resource condition

The second objective of the study was to integrate yield and soil water consumption to optimize the row spacing configuration under different irrigation conditions. The analysis showed that LAI was linearly and positively correlated with CAT, DWCI, and ET_C (Figures 6D, E, B), and we concluded that LAI was a key factor affecting the soil evaporation and transpiration of cotton. The cotton yields under CI were significantly higher than those under LI, but the ET_C increased by 22.9–32.6% under CI compared to those observed under LI, mainly because of the higher LAI. Therefore, CI is recommended to achieve higher yields in areas with sufficient water. Under CI, $RS_{76}L$ reduced CAT and DWCI because of lower LAI and overall reduced cotton ET_C and significantly increased WUE (Table 4, $P< 0.05$) relative to $RS_{66+10}H$ and $RS_{76}H$ under adequate irrigation. Both CAT and DWCI were significantly lower in $RS_{76}L$ than in $RS_{66+10}H$ and $RS_{76}H$ under the same irrigation amount (Figure 4; Table 5).

Planting density is also an important factor affecting crop ET_C, and related studies have shown that an increased planting density of maize significantly increases ET_C (Guo et al., 2021). Deep soil water consumption is significantly elevated owing to the high planting density (Magaia et al., 2017; Meng et al., 2020). $RS_{76}L$ under CI significantly reduced SAWC in the 20–60 cm soil layer within a horizontal distance of 19–38 cm from the cotton row radius compared with $RS_{66+10}H$ and $RS_{76}H$ (Figure 3). High-density planting, such as $RS_{66+10}H$, did not result in significant drought stress relative to $RS_{76}L$ under adequate irrigation. However, the ET_C of $RS_{76}L$ under CI decreased by 51–60 mm, but the WUE increased by 5.6–8.3% compared to $RS_{66+10}H$, which indicated that $RS_{76}L$ could further reduce irrigation to improve water use efficiency under CI. In conjunction with the development of machine harvesting cotton in China, $RS_{76}L$ under adequate irrigation is more conducive to cotton defoliation than $RS_{66+10}H$ and $RS_{76}H$ because of the larger row spacing and lower LAI in late reproduction (Li et al., 2016; Hu et al., 2021), which reduced cotton seed inclusion and improved cotton quality after mechanical harvesting. Moreover, $RS_{76}L$ saved seed cost and cotton labor topping cost compared to $RS_{66+10}H$, owing to half of the seeding volume. In summary, the combination of low-density equal row spacing with CI could reduce soil water consumption in the 20–60 cm soil layer while maintaining high cotton yields and has the

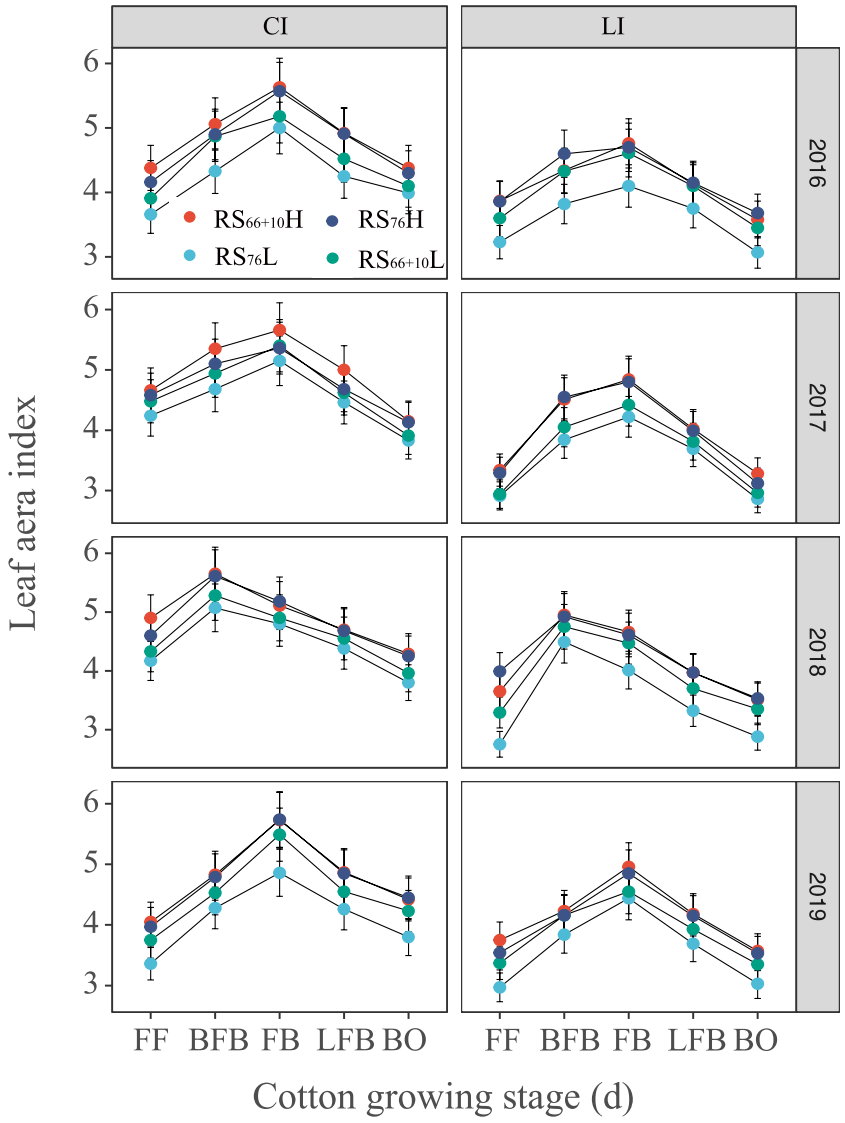


FIGURE 5
Leaf are index (LAI) of cotton in cotton (*Gossypium hirsutum*) affected by the combination of irrigation amount and row spacing configuration when evaluated in Xinjiang, China. FF, BFB, FB, LFB and BO means full flowering, before full boll, full boll, later full boll, and boll opening stage, respectively. Vertical bars represent the standard error. Mean values ± SE are from four replicates.

potential to further reduce irrigation. Therefore, $RS_{76}L$ under CI is also conducive to improving the machine–harvested quality of cotton and reducing management costs and is an optimum cotton planting pattern for mulch drip irrigation in arid areas.

Under LI, $RS_{66+10}H$ had the highest cotton yield, but the WUE did not differ from that of $RS_{76}L$ and $RS_{66+10}L$ ($P > 0.05$). High-density planting of cotton under LI can make full use of deep soil water by increasing root length and root surface area, inducing root growth in the deep and lateral soil layers, and promoting water uptake and transport for normal aboveground growth and development (Dong et al., 2010; Chen et al., 2018). Our results showed that $RS_{66+10}H$ significantly increased soil water consumption in the 20–60 cm soil layer at a horizontal distance of 19 –38 cm from the cotton row compared with $RS_{76}L$ under LI. Because of the use of mulch drip irrigation, soil water and roots are mainly distributed in the 0–60 cm soil layer range (Wang et al., 2014; Chen et al., 2022). The distribution of cotton roots under high-density planting coincided with the water supply in the 20–60 cm soil layer, which was fully utilized. Therefore, in cotton production areas where water resources are scarce, high-density planting can be used to tap the soil water production potential to achieve high and stable cotton yields.

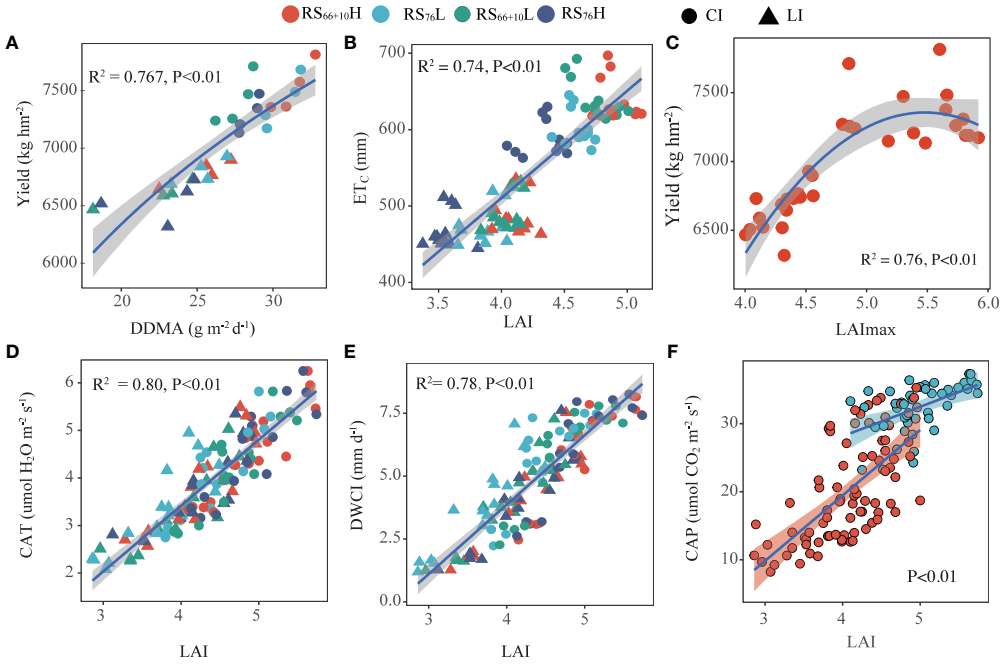


FIGURE 6
Correlation between leaf area index (LAI) and canopy apparent transpiration rate (CAT), daily water consumption intensity (DWCI), canopy apparent photosynthesis (CAP), crop evapotranspiration (ET_C), and Yield, and the correlation between daily dry matter accumulation (DDMA) and Yield. DDMA **(A)**, ET_C **(B)** and LAI **(E, F)** were the mean values from full flowering to the boll opening stage; LAImax **(C)** is the maxmum LAI value in critical growth period. CAP **(F)**, CAT **(D)**, and DWCI **(E)** were the mean values of each growth period (full flowering stage, full boll stage, and boll opening stage) corresponding to LAI. CI (conventional irrigation); LI (limited irrigation); RS_{66+10H} and RS_{66+10L} (high/low-density planting with 66+10cm row spacing configuration); $RS_{76}H$ and $RS_{76}L$ (high/low-density planting with 76 cm row spacing configuration).

5 Conclusion

This study proposed the most suitable planting pattern based on different irrigation conditions. This research showed that a maximum LAI (LAI_{max}) maintained between 5.0 and 5.5 was most conducive to high yield and higher WUE. Under sufficient water, optimize low density row spacing configuration ($RS_{76}L$) could reach the same yield level as high-density planting, whereas suitable LAI_{max} reduced CAT, DWCI, and soil water consumption of 20-60 cm soil layers. Under water restriction condition, high-density planting ($RS_{66\ +\ 10}H$) could fully exploit the soil water potential of 20-60 cm soil layers to improve cotton yields. Our results suggest that in northern of Xinjiang's moisture-rich areas, $RS_{76}L$ has the advantage of receiving high cotton yield while improving cotton benefits and further reducing the irrigation amount to improve WUE. However, moisture-limited areas are more suitable for high-density planting to increase yields.

Data availability statement

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

Author contributions

WQZ, BW and YW conceptualized the study. WQZ, WFZ and JT involved in methodology. WQZ and BW involved in formal analysis and writing— original draft. WQZ, XJ and HD investigated the study. WQZ and SX involved in writing—review and editing. WFZ supervised the study. All authors contributed to the article and approved the submitted version.

Funding

This experimental work was supported by the National Key Research and Development Program of China (2020YFD1001001).

Conflict of interest

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

The reviewer ZW declared a shared affiliation with the authors WQZ, BW, YW, SX, JT, WFZ to the handling editor at the time of review.

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1158329/full#supplementary-material

References

Anwar, M. R., Wang, B., Li Liu, D., and Waters, C. (2020). Late planting has great potential to mitigate the effects of future climate change on Australian rainfed cotton. *Sci. Total Environ.* 714, 136806.

Araus, J. L., Sanchez-Bragado, R., and Vicente, R. (2021). Improving crop yield and resilience through optimization of photosynthesis: panacea or pipe dream? *J. Exp. Bot.* 72, 3936–3955.

Bilal, A., Ahmad, A., Rasul, F., and Murtaza, G. (2019). Optimization of the sowing time for bt-cotton production in punjab, pakistan. *Pak. J. Biol. Agric. Sci.* 56, 95–100.

Brodrick, R., Bange, M. P., Milroy, S. P., and Hammer, G. L. (2010). Yield and maturity of ultra-narrow row cotton in high input production systems. *Agron. J.* 102 (3), 843–848.

Chen, W., Chen, F., Lai, S., Jin, M., Xu, S., Liu, Y., et al. (2022). Spatial distribution and dynamics of cotton fine root under film-mulched drip irrigation. *Ind. Crops Prod.* 179, 114693.

Chen, Z., Niu, Y., Zhao, R., Han, C., Han, H., and Luo, H. (2019). The combination of limited irrigation and high plant density optimizes canopy structure and improves the water use efficiency of cotton. *Agric. Water Manage.* 218, 139–148.

Chen, Z., Tao, X., Khan, A., Tan, D. K., and Luo, H. (2018). Biomass accumulation, photo-synthetic traits and root development of cotton as affected by irrigation and nitrogen-fertilization. *Front. Plant Sci.* 9, 173–187.

Cui, Z., Zhang, H., Chen, X., Zhang, C., Ma, W., Huang, C., et al. (2018). Pursuing sustainable productivity with millions of smallholder farmers. *Nature.* 555, 363–366.

Dai, J., and Dong, H. (2014). Intensive cotton farming technologies in China: achievements, challenges and countermeasures. *Field Crops Res.* 155, 99–110.

Di, N., Wang, Y., Clothier, B., Liu, Y., Jia, L., Xi, B., et al. (2019). Modeling soil evaporation and the response of the crop coefficient to leaf area index in mature populus tomentosa plantations growing under different soil water availabilities. *Agric. For. Meteorol.* 264, 125–137.

Dong, H., Kong, X., Luo, Z., Li, W., and Xin, C. (2010). Unequal salt distribution in the root zone increases growth and yield of cotton. *Eur. J. Agron.* 33, 285–292.

Dong, H., Li, W., Tang, W., Li, Z., Zhang, D., and Niu, Y. (2006). Yield, quality and leaf senescence of cotton grown at varying planting dates and plant densities in the yellow river valley of China. *Field Crops Res.* 98, 106–115.

Feng, L., Dai, J., Tian, L., Zhang, H., Li, W., and Dong, H. (2017). Review of the technology for high-yielding and efficient cotton cultivation in northwest inland cotton-growing region of China. *Field Crops Res.* 208, 18–26.

Forouzani, M., and Karami, E. (2011). Agricultural water poverty index and sustainability. *Agron. Sustain. Dev.* 31, 415–431.

Guo, Q., Huang, G., Guo, Y., Zhang, M., Zhou, Y., and Duan, L. (2021). Optimizing irrigation and planting density of spring maize under mulch drip irrigation system in the arid region of Northwest China. *Field Crops Res.* 266, 108141.

Heitholt, J. J. (1994). Canopy characteristics associated with deficient and excessive cotton plant population densities. *Crop Sci.* 34, 1291–1297.

Hernandez, M. D., Alfonso, C., Echarte, M. M., Cerrudo, A., and Echarte, L. (2021). Maize transpiration efficiency increases with n supply or higher plant densities. *Agric. Water Manage.* 250, 106816.

Hou, X., Fan, J., Zhang, F., Hu, W., Yan, F., Xiao, C., et al. (2022). Determining water use and crop coefficients of drip-irrigated cotton in south xinjiang of China under various irrigation amounts. *Ind. Crops Prod.* 176, 114376.

Hu, L., Pan, X., Wang, X., Hu, Q., Wang, X., Zhang, H., et al. (2021). Cotton photosynthetic productivity enhancement through uniform row-spacing with optimal plant density in xinjiang, China. *Crop Sci.* 61, 2745–2758.

Kerby, T. A., Cassman, K. G., and Keeley, M. (1990). Genotypes and plant densities for narrow-row cotton systems. II. leaf area and dry-matter partitioning. *Crop Sci.* 30, 649–653.

Khan, A., Najeeb, U., Wang, L., Tan, D. K. Y., Yang, G., Munsif, F., et al. (2017b). Planting density and sowing date strongly influence growth and lint yield of cotton crops. *Field Crops Res.* 209, 129–135.

Khan, A., Wang, L., Ali, S., Tung, S. A., Hafeez, A., and Yang, G. (2017a). Optimal planting density and sowing date can improve cotton yield by maintaining reproductive organ biomass and enhancing potassium uptake. *Field Crops Res.* 214, 164–174.

Kodur, S. (2017). Improving the prediction of soil evaporation for different soil types under dryland cropping. *Agric. Water Manag* 193, 131–141.

Li, J. F., Liang, F. B., Chen, H. C., Zhang, W. F., and Kang, P. (2016). Effect of row spacing configuration on agronomic traits and yield of cotton under machine-harvested pattern. *Xinjiang Agric. Sci.* 08), 1390–1396.

Li, H., Wan, H. L., Tian, L. W., Liu, L. T., Zhang, Y. J., Bai, Z. Y., et al. (2020). The effects of increased-density on canopy apparent photosynthesis, dry matter accumulation and distribution of cotton under late-sown condition. *Cotton Sci.* 32, 339–347.

Li, C., Wu, P. T., Li, X. L., Zhou, T. W., Sun, S. K., Wang, Y. B., et al. (2017). Spatial and temporal evolution of climatic factors and its impacts on potential evapotranspiration in loess plateau of northern shaanxi, China. *Sci. Total Environ.* 589, 165–172.

Li, H., Zhu, Y. F., and Cai, D. L. (2021). Study on the characteristics of water use in agriculture and major crops in xinjiang. *Agric. Technol.* 41 (21), 40–43.

Liao, Z., Zeng, H., Fan, J., Lai, Z., Zhang, C., Zhang, F., et al. (2022). Effects of plant density, nitrogen rate and supplemental irrigation on photosynthesis, root growth, seed yield and water-nitrogen use efficiency of soybean under ridge-furrow plastic mulching. *Agric. Water Manage.* 268, 107688.

Luo, H. H., Zhang, H. Z., Han, H. Y., Zhang, Y. L., and Zhang, W. F. (2014). Effects of water sto-rage in deeper soil layers on growth, yield, and water productivity of cotton (Gossypium hirsutum l.) in arid areas of northwestern China. *Irrig. Drain.* 63, 59–70.

Lv, X., Wang, Z., Ma, L., Cao, N., Meng, Y., and Zhou, Z. (2021). Crop residue incorporation combined with potassium fertilizer increased cotton canopy apparent photosynthesis and seed cotton yield in barley-cotton rotation system. *Arch. Agron. Soil Sci.* 67, 300–312.

Magaia, E., Famba, S., Wesström, I., Brito, R., and Joel, A. (2017). Modelling maize yield response to plant density and water and nitrogen supply in a semi-arid region. *Field Crop Res.* 205, 170–181.

Malone, S., Herbert, D. A., and Holshouser, D. L. (2002). Evaluation of the LAI-2000 plant canopy analyzer to estimate leaf area in manually defoliated soybean. *Agron. J.* 94, 1012–1019.

Meng, X., Lian, Y., Liu, Q., Zhang, P., Jia, Z., and Han, Q. (2020). Optimizing the planting density under the ridge and furrow rainwater harvesting system to improve crop water productivity for foxtail millet in semiarid areas. *Agric. Water Manage.* 238, 106220.

Neumann, K., Stehfest, E., Verburg, P. H., Siebert, S., Müller, C., and Veldkamp, T. (2011). Exploring global irrigation patterns: a multilevel modelling approach. *Agr. Syst.* 104, 703–713.

Paul, M., Rajib, A., Negahban-Azar, M., Shirmohammadi, A., and Srivastava, P. (2021). Improved agricultural water management in data-scarce semi-arid watersheds: value of integrating remotely sensed leaf area index in hydrological modeling. *Sci. Total Environ.* 791, 148177.

Qin, S., Li, S., Kang, S., Du, T., Tong, L., and Ding, R. (2016). Can the drip irrigation under film mulch reduce crop evapotranspiration and save water under the sufficient irrigation condition? *Agric. Water Manage.* 177, 128–137.

Rafiee, M., and Kalhor, M. (2016). Economic water use efficiency of corn (Zea mays l.) hybrids as affected by irrigation regimes: a case study in West Iran. *Arch. Agron. Soil Sci.* 62, 781–789.

Rahman, M. A., Moser, A., Gold, A., Rötzer, T., and Pauleit, S. (2018). Vertical air temperature gradients under the shade of two contrasting urban tree species during different types of summer days. *Sci. Total Environ.* 633, 100–111.

Reddy, V., Reddy, K., and Hodges, H. (1995). Carbon dioxide enrichment and temperature effects on cotton canopy photosynthesis transpiration, and water-use effciency. *Field Crops Res.* 41, 13–23.

Shi, F., Li, N. N., Khan, A., Lin, H. R., Tian, Y., Shi, X., et al. (2022). DPC can inhibit cotton apical dominance and increase seed yield by affecting apical part structure and hormone content. *Field Crop Res.* 282, 108509.

Srinivasan, V., Kumar, P., and Long, S. P. (2017). Decreasing, not increasing, leaf area will raise crop yields under global atmospheric change. *Glob. Change Biol.* 23, 1626–1635.

Sui, J., Wang, J., Gong, S., Xu, D., Zhang, Y., and Qin, Q. (2018). Assessment of maize yield-increasing potential and optimum n level under mulched drip irrigation in the northeast of China. *Field Crop Res.* 215, 132–139.

Tabashnik, B. E., and Carrière, Y. (2019). Global patterns of resistance to bt crops highlighting pink bollworm in the united states, China, and India. *J. Econ. Entomol.* 112, 2513–2523.

Tian, J., Zhang, X., Yang, Y., Xu, S., Zuo, W., Zhang, W., et al. (2017). How to reduce cotton fiber damage in the xinjiang China. *Ind. Crops Prod.* 109, 803–811.

Wang, Y., Chen, M., Liang, F., Tian, J., Zhang, Y., Jiang, C., et al. (2021a). Photosynthates competition within the boll–leaf system is alleviated with the improvement of photosynthetic performance during the succession of xinjiang cotton cultivars. *Ind. Crops Prod.* 160, 113121.

Wang, F., Xiao, J., Ming, B., Xie, R., Wang, K., Hou, P., et al. (2021b). Grain yields and evapotranspiration dynamics of drip-irrigated maize under high plant density across arid to semi-humid climates. *Agric. Water Manage.* 247, 106726.

Wang, Z., Yang, P., Zheng, X., He, X., Zhang, J., and Li, W. (2014). Soil salt dynamics in cotton fields with mulched drip irrigation under the existing irrigation system in xinjiang. *Agric. Machinery* 45, 149–159.

Wei, K., Zhang, J., Wang, Q., Guo, Y., and Mu, W. (2022). Irrigation with ionized brackish water affects cotton yield and water use efficiency. *Ind. Crops Prod.* 175, 114244.

Wu, Y., Huang, F. Y., Jia, Z. K., Ren, X. L., and Cai, T. (2017). Response of soil water, temperature, and maize (Zea may l.) production to different plastic film mulching patterns in semi-arid areas of northwest China. *Soil Tillage Res.* 166, 113–121.

Xie, T. T., Su, P. X., and Gao, S. (2010). Photosynthetic rate, transpiration rate, and water use efficiency of cotton canopy in oasis edge of linze. *J. Appl. Ecol.* 21, 1425–1431.

Yao, H., Zhang, Y., Yi, X., Zuo, W., Lei, Z., Sui, L., et al. (2017). Characters in light-response curves of canopy photosynthetic use efficiency of light and n in responses to plant density in field-grown cotton. *Field Crop Res.* 203, 192–200.

Zhang, W., Wang, Z., Yu, S., Li, S., Fang, J., and Tong, W. (2004). Effects of planting density on canopy photosynthesis, canopy structure and yield formation of high-yield cotton in xinjiang, China. *Chin. J. Plant Ecol.* 28, 164.

Zou, H., Fan, J., Zhang, F., Xiang, Y., Wu, L., and Yan, S. (2020). Optimization of drip irrigation and fertilization regimes for high grain yield, crop water productivity and economic benefits of spring maize in Northwest China. *Agric. Water Manage.* 230, 105986.

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 05 June 2023
DOI 10.3389/fpls.2023.1172354

OPEN ACCESS

EDITED BY
Meng Kou,
Xuzhou Institute of Agricultural Sciences in Jiangsu Xuhuai District, China

REVIEWED BY
Oswaldo Valdes-Lopez,
National Autonomous University of Mexico, Mexico
Qiuming Yao,
University of Nebraska-Lincoln, United States

*CORRESPONDENCE
Lixiang Wang
lxwang@sxau.edu.cn
Shujin Guo
sxndgsj@163.com

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

RECEIVED 23 February 2023
ACCEPTED 02 May 2023
PUBLISHED 05 June 2023

CITATION
Zheng J, Sun L, Wang D, He L, Du W, Guo S and Wang L (2023) Roles of a CCR4–NOT complex component GmNOT4-1 in regulating soybean nodulation. *Front. Plant Sci.* 14:1172354. doi: 10.3389/fpls.2023.1172354

Roles of a CCR4–NOT complex component GmNOT4-1 in regulating soybean nodulation

Jiangtao Zheng[1†], Lili Sun[1†], Dongmei Wang[1], Lin He[1], Weijun Du[1], Shujin Guo[1*] and Lixiang Wang[1,2*]

[1]College of Agronomy, Shanxi Agricultural University, Taigu, China, [2]State Key Laboratory of Crop Stress Adaptation Improvement, School of Life Sciences, Henan University, Kaifeng, China

Legume-rhizobial symbiotic nitrogen fixation is the most efficient nitrogen assimilation system in the ecosystem. In the special interaction between organ–root nodules, legumes supply rhizobial carbohydrates for their proliferation, while rhizobials provide host plants with absorbable nitrogen. Nodule initiation and formation require a complex molecular dialogue between legumes and rhizobia, which involves the accurate regulation of a series of legume genes. The CCR4–NOT complex is a conserved multi-subunit complex with functions regulating gene expression in many cellular processes. However, the functions of the CCR4–NOT complex in rhizobia–host interactions remain unclear. In this study, we identified seven members of the *NOT4* family in soybean and further classified them into three subgroups. Bioinformatic analysis showed that *NOT4s* shared relatively conserved motifs and gene structures in each subgroup, while there were significant differences between *NOT4s* in the different subgroups. Expression profile analysis indicated that *NOT4s* may be involved in nodulation in soybean, as most of them were induced by Rhizobium infection and highly expressed in nodules. We further selected *GmNOT4-1* to clarify the biological function of these genes in soybean nodulation. Interestingly, we found that either *GmNOT4-1* overexpression or down-regulation of *GmNOT4-1* by RNAi or CRISPR/Cas9 gene editing would suppress the number of nodules in soybean. Intriguingly, alterations in the expression of *GmNOT4-1* repressed the expression of genes in the Nod factor signaling pathway. This research provides new insight into the function of the *CCR4–NOT* family in legumes and reveals *GmNOT4-1* to be a potent gene for regulating symbiotic nodulation.

KEYWORDS

legume symbiotic nitrogen fixation, Ccr4-not complex, root nodule (symbiotic), CRISPR/Cas9, nod factor signalling

Background

Nitrogen is one of the essential macroelements for plant growth and development (Jia et al., 2017). Therefore, improving nitrogen utilization efficiency and appropriately applying nitrogen fertilizer are important guarantees for a high and stable yield of crops (Chun et al., 2005). In the ecosystem, given that leguminous plants have a high demand for nitrogen, they evolved an additional special root organ–root nodules—to fix atmospheric nitrogen and provide nitrogen to improve legume development (Ren et al., 2019).

Nodulation is a complex biological process involving direct interactions between rhizobial and legume signals (Howard, 1991). Nodulation is initiated by the flavonoids secreted by the legume roots, which were perceived by the surrounding compatible rhizobium strains and stimulated them to synthesize lipochitin oligosaccharides, called Nod factors (NF, Patra et al., 2017), which were sensed by Nod factor receptors (NFRs) (e.g., NF Perception [NFP] in *M. truncatula*, NF Receptors 1 and 5 [NFR1/5] in *L. japonicus*, and NFR1/5α in soybean) located on legume root hairs. This interaction then stimulated the consequent NF signaling pathway, which promoted root hair deformation, infection thread formation, outer cortical cell division, and root nodule primordia formation (Gao et al., 2002; Chou and Wei, 2010). Nodulation is a high-energy-consumption biological and host-dominant process; thus, legumes evolved an auto-regulation of nodulation (AON) mechanism; in brief, when the root nodules reach a certain number, the host will generate the CLE peptides as a signal molecular (Carroll et al., 2016), which would transmit to the legume shoot and be perceived by the NARK receptor. Furthermore, this recognition will generate shoot-derived molecular signals (e.g., cytokinin, miRNA2111), which transmit back to the roots and attenuate the nodulation process (Chou and Wei, 2010; Yuan et al., 2016).

Nodules are produced via a complex genetic program to allow rhizobial recognition and nodule formation, a series of transcription factors modulate the downstream responses to NF signaling, including NIN (L. Yuan et al., 2016; Wang et al., 2019), IPD3 (Interacting Protein of DMI3; Horváth, 2011), ERN1/ERN2/ERN3 (ERFs Required for Nodulation; Andriankaja et al., 2007; Middleton et al., 2007), NF-YA1 (Nuclear Factor-Y Subunit A1; Laloum et al., 2013), the NSP1 (Nodulation Signaling Pathway1) and NSP2 (Kaló et al., 2005; Smit et al., 2005; Oldroyd & Downie, 2008; Heckmann et al., 2011), NNC1 (Nodule number control1), and the DELLAs (Fonouni-Farde et al., 2016; Jin et al., 2016).

The ubiquitin–proteasome system is the most efficient and specific protein degradation mode for regulating plant growth and development. *ASTRAY and SINAT5* encoding RING-finger domains containing E3 ubiquitin ligase have been shown to function during legume nodulation (Fujita and Kawaguchi, 2002; Nishimura et al., 2002). The no-nodule alfalfa mutants *rh2* and *LjnsRING*, encoding RING-H2 domain-containing proteins (Shimomura et al., 2006), and LIN (in Medicago) and CERBERUS (in Lotus) shared 86% homology and both contained a U-box domain (Kiss et al., 2010). Plant U-box protein1 (PUB1) has the activity of an E3 ligase and interacts with LYK3/NFR1/DMI2/SYMRK to inhibit the infection of rhizobia and mycorrhizal fungi through its ubiquitination activity (Vernié et al., 2016).

The carbon catabolite repression 4–negative on TATA-less (CCR4–NOT) complex multi-subunit complex functions as a major regulator of gene expression homeostasis through ubiquitination (CCR4 and CAF1 [CCR4 associated factor 1]) and deadenylation (NOT4) in eukaryotes. In plants, core components of the CCR4–NOT complex were identified and revealed that Arabidopsis possesses AtCCR4–NOT complexes involved in mRNA recognition, AtCCR4-CAF1 has mRNA deadenylase activity to regulate environmental stresses (Liang et al., 2008; Walley et al., 2010; Suzuki et al., 2015), and NOT9B and CCR4–NOT can respond to far-red light and are involved in phyA-modulated gene expression. However, the identification of NOT4 in the complex and its function in legumes are yet to be unveiled.

In this study, genome-wide systematic characterization, including protein properties, chromosome distribution, phylogenetic relationship, protein motif, and gene expression pattern, was performed to study the soybean CCR4–NOT complex gene family. We selected *GmNOT4-1*, a member of the CCR4–NOT complex gene family that is highly expressed in nodules and significantly stimulated by Rhizobium infection, to determine its function in symbiotic nodulation using overexpression, RNAi, and CRISPR/Cas9. We found that in transgenic hairy roots harboring *GmNOT4-1*-overexpressing or in roots carrying *GmNOT4-1*-RNAi and CRISPR/Cas9-*GmNOT4-1*, the number of root nodules was significantly inhibited, and the marker genes for both the NF and AON signaling pathways were repressed. In conclusion, this study reports for the first time the function of the key eukaryotic gene expression regulatory complex CCR4–NOT in legume nodulation and identifies a distinguished nodulation regulator, *NOT4-1*, whose expression balance is relevant to soybean symbiosis.

Materials and methods

Plant materials and growth, hairy root transformation, and inoculation of soybean rhizobium

In this study, soybean [*G. max* (L.) *Merrill* cv. Williams 82] and *Agrobacterium rhizogenes* strain K599 were used for the hairy root transformation. The hairy root transformation procedure was previously described (Wang et al., 2014) with some modifications. The positive transformed composite plants were cultured in a low-nitrogen nutrient solution for 5 days for recovery; after that, the plants were transferred to vermiculite for inoculation with a suspension of *B. japonicum* strain USDA110 (30 ml, OD600 = 0.08). Nodule numbers were evaluated at 28 DAI (days after inoculation).

Identification of *NOT4* gene family members in soybean

Genome data, protein sequence, and genome annotation files of *Glycine max* were downloaded from the Phytozome database

(https://phytozome-next.jgi.doe.gov/). Protein families in the PANTHER database (http://www.pantherdb.org) were applied to download the Hidden Markov model (HMM) of the NOT transcription complex related family (PTHR12603). The soybean genome was searched using hmm search in HMMER (http://hmmer.org/) to identify NOT4 candidate genes (the screening threshold was 1.0e−10). The resulting candidate sequences were submitted to Intel Pro Scan (https://www.ebi.ac.uk/interpro/search/sequence-search) to check the PTHR12603 structure domain. Soybean *GmNOT4* family members were named according to their location distribution on soybean chromosomes.

Phylogenetic analysis and chromosome mapping of *GmNOT4s in* soybean

NOT4 protein sequences from *Medicago truncatula*, *Phaseolus vulgaris*, *Arabidopsis thaliana*, soybean [*G. max* (L.)], and rice (*Oryza sativa*) NOT4 were downloaded from the Phytozome database (https://phytozome-next.jgi.doe.gov/). The phylogenetic tree of the *GmNOT4* gene family members of the five species was constructed by MEGA-X software and calculated by the NJ (neighbor-joining) method. The parameters were set to self-expanding and repeated 1,000 times. Chromosome distribution was analyzed using TBtools.

Conserved motif and gene structure analysis

The MEME Suite online software (https://meme-suite.org/meme/) was used for conserved motif analysis, and the parameter was set to 10. The gene structure information of *GmNOT4s* was extracted from the soybean gene information GFF file, and the gene structures were visualized by TBtools.

Gene expression

RNAprep Plant Plus Trizol Kit was used to extract RNA from collected samples, and the first-strand cDNA was synthesized using the Super Mix Kit (Hifair II 1 strand cDNA Synthesis SuperMix, gDNA digester plus) (Yeasen Biotech Co. Ltd., Shanghai, China). qPCR was performed using SYBR Green JumpStart Taq ReadyMix (Sigma-Aldrich). *GmCYP2* was used as an internal control (Jian et al., 2008). The primers used in this study are shown in Table S1.

Plasmid construction

For the *GmNOT4-1* overexpression construct, the *GmNOT4-1* CDS fragment was inserted into the pCAMBIA1300-GFP vector through seamless cloning using the *Bam*H1 restriction site; for the *GmNOT4-1*-RNAi construct, the *GmNOT4-1* CDS fragment was ligated into the pDONOR207 entry vector, and then the target sequence was cloned into the pK7GWIW-GFP vector through the gateway LR reaction. For the *GmNOT4-1* CRISPR/Cas9 knock-out construct, *sgRNAs* were designed using *the* software Crispr-P (http://cbi.hzau.edu.cn/crispr/), and the top two reliable sgRNAs, CAAGGTGCGGTGAAGAGCA and TCGTCCTCTTCGCCTCTGC, were selected. Then, vector pCBC-DT1T2 was used as a template to clone the two CRISPR fragments, and the two obtained products were inserted into vector pKSE401-GFP.

Statistical analysis

GraphPad Prism 7 (GraphPad Software, La Jolla, CA, USA) was used to analyze the data in this study. A Student's *t*-test was performed to generate *P*-values. The statistical differences are marked as follows: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

Results and discussion

Identification and chromosomal distribution of soybean *NOT4* gene family

Seven *GmNOT4s* were identified from the soybean genome using BLAST and PANTHER searches and named *GmNOT4-1*–*GmNOT4-7* according to their positions on six chromosomes (Figure S1). The chromosomal distribution of *NOT4s* indicated that they were dispersed on chromosomes 5, 10, 12, 15, and 17. Except for *GmNOT4-4* and *GmNOT4-5*, which are on chromosome 13, each of the other five chromosomes contains one family member. The physical and chemical properties of *GmNOT4s* were analyzed, and it was found that the amino acid residues encoded by seven *GmNOT4s* ranged from 232 (GmNOT4-1) to 1,046 (GmNOT4-6), and the corresponding molecular weight ranged from 35,386.84 (GmNOT4-1) to 115,070.36 (GmNOT4-6) Da. The theoretical isoelectric points of seven *GmNOT4* family members ranged from 4.76 (GmNOT4-7) to 6.40 (GmNOT4-4), all of which belonged to weakly acidic proteins (Table S2). The total mean hydrophobic index was less than 0, indicating that they were hydrophilic proteins. To understand the functional characteristics and evolutionary relationship of the *NOT4* gene family, NOT4 protein sequences in soybean, *M. truncatula*, common bean, *A. thaliana*, and rice were retried to construct the phylogenetic tree. The results showed that the members of the *NOT4* gene family were divided into three subgroups (Groups I–III), in which groups I and III contained 10/11 members, and group II had three members. In the same subgroups, NOT4 proteins cluster together within species (Figure 1). Soybean *NOT4* family members were evenly distributed in three groups, with three members in Group I, two members in Group II, and two members in Group III.

Conserved motifs and gene structure analysis of *NOT4* family genes in soybean

To get a hint about the function of NOT4s, we first performed conserved motif analysis and gene structure analysis. There are 10

conserved motifs obtained from the GmNOT4 protein, and the number of motifs of GmNOT4s in different subgroups varied significantly, ranging from 2 to 10. Three members of Group I contain 10 motifs; there were eight motifs in GmNOT4-2 and seven for GmNOT4-4 in Group II and the members of Group III contain only two motifs. Among the 10 motifs, Motif1 was the most conserved one and was present in all GmNOT4 proteins (Figure 2). The variation in the types and amounts of conserved motifs in GmNOT4s reflects the functional diversity of these proteins. In addition, we have confirmed the conservation of motifs in NOT4 proteins in other plant species; we found these motifs are consistently present among Arabidopsis, rice, Medicago, and the common bean (Figure S2).

Diagram of *GmNOT4s* gene structures shows the exon number variation of GmNOT4s ranged from 3 to 13 (Figure 2). Combined with phylogenetic tree analysis, we found that genes with close genetic relationships mostly had similar gene structures. For example, all Group I GmNOT4s contained 12 exons, while Group III GmNOT4s contained three exons.

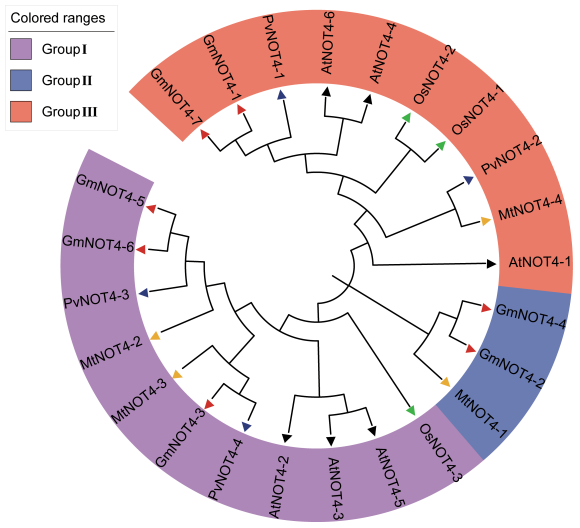


FIGURE 1
Phylogenetic tree of *NOT4* family members. NOT4 Protein sequences of *Medicago truncatula (Mt)*, *Phaseolus vulgaris (Pv)*, Arabidopsis thaliana (At), *Glycine max (Gm)* and *Oryza sativa (Os)* were divided into Group I to Group III and colored in purple, blue, and orange, respectively. Protein sequences were clustered using CLASTALW in MEGA 11.0. The phylogenetic tree was constructed by MEGA 11.0 using a bootstrap neighbor join method.

Cis-acting element distribution in *GmNOT4s* promoters

Most transcriptional factors function in gene expression regulation by binding specific *cis*-acting elements in gene promoters to modulate gene expression. To predict the functions of GmNOT4 genes, 2-kilobase pairs upstream of the translational initiation site were selected as the promoter sequence and submitted to PlantCARE online software to predict the *cis*-acting elements (Figure S3). A total of 63 *cis*-acting elements were identified in seven *GmNOT4* promoters, while 54 *cis*-acting elements were related to plant hormones, stress, growth, and development. Firstly, core *cis*-acting elements exist in almost all promoters and consist of AT-TATA-box, CAAT-box, TATA-box, and TATA. The second group contains 26 cis-acting elements related to plant growth and development; all *GmNOT4* promoters contain photoreactive elements, which coincide with the CCR4–NOT complex response to far-red light. However, only *GmNOT4-1* and *GmNot4-7* contain *cis*-regulatory elements for flavonoid biosynthesis, which play a vital role in rhizobia attraction and symbiosis construction. Plant hormones are important for legume nodulation; 13 kinds of plant hormone related *cis*-acting elements were harbored in *GmNOT4* promoters; all *GmNOT4* promoters contained abscisic acid and ethylene-related *cis*-acting elements; and auxin response elements

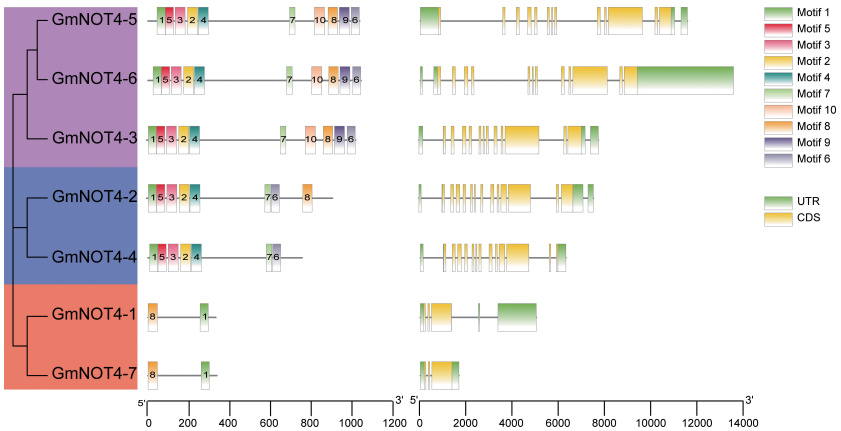


FIGURE 2
Gene structure and conserved motif distribution of *GmNOT4s*. MEME online software was used to identify conserved motifs. A total of 10 motifs were found in *GmNOT4s* genes and represented by different colored boxes and corresponding motif number, conserved amino acid sequence themes were shown below. (The left panel) shown the phylogenetic tree of group I–III genes in purple, blue, and orange colors; (The middle panel) shown the conserved motifs distribution in GmNOT4s; (The right panel) shown the gene structure of untranslated regions (UTRs) of the *GmNOT4s* gene are shown as green, yellow boxes and black lines.

(TGA-elements), gibberellin response elements (P-box, GARE-motif) existed in most *GmNOT4s* promoters. Low oxygen conditions were very important for nitrogen fixation in nodules, and we found that among the stress response *cis* elements, *GmNOT4s* harbored *the* most anaerobic induction and stress-related *cis* elements. The specific *cis*-acting element distributions in *GmNOT4* promoters suggest their different biological functions.

Gene expression pattern of *GmNOT4* genes

To investigate the potential function of *NOT4s* during soybean nodulation, we examined the temporal and spatial expression patterns of *GmNOT4s* by qRT-PCR.

We found that all but NOT4-7 were highly expressed in root nodules (Figures 3A–F). Further study has shown that the expression of *GmNOT4-1*, *GmNOT4-2*, *GmNOT4-3*, *GmNOT4-4*, *and GmNOT4-7* was dramatically stimulated by rhizobium infection at 1 DAI (Figures 3H–K, N). The expression of *GmNOT4-1 and GmNOT4-2* was induced by rhizobium USDA110 while *GmNOT4-3* was repressed at 3 DAI (Figure 3H–J). The expression of *GmNOT4-1*, *GmNOT4-5*, *and GmNOT4-6* was induced by rhizobium, while *GmNOT4-4* and *GmNOT4-7* were repressed at 6 DAI (Figures 3H, L, M). These results show that all *GmNOT4s* responded to rhizobium at least at one point. Combining the above tissue expression pattern to prove the function of *GmNOT4s in* legume root nodule symbiosis, we selected *GmNOT4-1 for* further study, whose expression was induced at all checkpoints within a short period of inoculation (1, 3, and 6 DAI).

GmNOT4-1 is an important regulator in regulating soybean nodulation

To genetically explore whether *GmNOT4-1* is involved in the regulation of soybean nodulation, overexpression and knockout/knockdown analyses of *GmNOT4-1* were performed using the hairy root transformation system. Firstly, we constructed *35S:GmNOT4-1 and* obtained overexpressing *GmNOT4-1* (*GmNOT4-1*-OX) roots by qPCR analysis (Figure 4A). The effects of *GmNOT4-1* overexpression on the early and late stages of nodulation were evaluated at 1, 6, and 28 days after inoculation (DAI) (Figure S4). Firstly, the expression of rhizobial infection-related genes, including *GmRPG*, *GmNPL*, *GmVPY*, *GmCYCLOPS*, and *GmSCARN*, was validated. As shown in Figure S5, all rhizobial infection-related genes we checked in this study were significantly inhibited in *GmNOT4-1* overexpressing roots. Then, we found the number of root hairs showing deformation was markedly decreased in *GmNOT4-1*-OX hairy roots at 6 DAI (Figure S6). Finally, the nodule numbers were quantified at 28 DAI. The average nodule number per *GmNOT4-1* overexpressed root was about 3.6, while empty vector (EV) control roots produced an average of 11.8 nodules per root, with an approximately 69% reduction by *GmNOT4-1* overexpression. These data suggest that *GmNOT4-1* plays a negative role in regulating soybean nodulation (Figures 4B, C).

To further validate the function of *GmNOT4-1* during nodulation, we constructed the *NOT4-1*-RNAi to analyze the soybean nodulation phenotype when the expression of *GmNOT4-1* was reduced (Figure S4). As shown in Figures 5B, C, transgenic roots harboring*GmNOT4-1*-RNAi produced fewer nodules (6.9 nodules per root) than empty vector (EV) control roots (19.5 nodules per root), so that the nodule number per *GmNOT4-1*-silenced root was reduced by approximately 64.6%. Further, the CRISPR/Cas9 system was also applied to knock out *GmNOT4-1* in soybean hairy roots (Figures S8, S9), and the gene editing was verified by sequencing (Table S3). The expression of rhizobial infection-related genes, including *GmRPG*, *GmNPL*, *GmVPY*, *GmCYCLOPS*, and *GmSCARN*, was significantly inhibited in *GmNOT4-1*-KO roots. The number of deformed root hairs was markedly decreased in *GmNOT4-1* edited knock-out hairy roots at 6 DAI compared to the vector control (Figure S6). Finally, we found that the *GmNOT4-1* edited roots produce functional nodules (Figure S7) with a significantly reduced nodule number, which showed the same result as the *GmNOT4-1*-RNAi roots (Figures 5B, C).

Combined with the overexpression results, it is suggested that homeostasis of the expression of *GmNOT4-1* is critical for the regulation of soybean nodulation.

Marker genes in the NF pathway were affected by *GmNOT4-1*

Nodule number was mainly modulated by both the NF signaling and AON signaling pathways. Given the phenotype that either overexpression or knockdown (out) of *GmNOT4-1* significantly decreased soybean nodule number, it is worth checking whether *GmNOT4-1* regulates nodulation through the above pathways. Thus, we examined the expression level of several nodulation marker genes in soybean, including NF signaling pathway genes *ENOD40*, *GmNINa*, *NSP1*, *HAP2-1*, and *HAP2-2*, and AON signaling pathway genes *GmRIC1/2*. In general, we found that either overexpression or silence of *GmNOT4-1* resulted in reduced expression of nodulation marker genes compared with the empty vector control roots (Figure 6).

Legumes plants can specifically interact with the phylogenetically diverse group of soil bacteria-rhizobia to form nodules. However, symbiotic nitrogen fixation is a highly energy-intensive biological process; thus, host legumes have evolved a root–shoot–root long-distance auto-regulation of nodulation (AON) system to refine the number of nodules (Patra et al., 2017; Ren et al., 2019). NODULE INCEPTION (NIN) induces the expression of CLE ROOT SIGNAL1 (*CLE-RS1*) and *CLE-RS2* to activate AON to inhibit excessive nodulation (Chun et al., 2005). NIN also modulates almost all nodulation processes, including nodule initiation, nodule organogenesis, and nitrogen fixation. Thus, NIN can act as a bifunctional transcription factor, fine-tuning legume nodulation. Interestingly, in this study, we observed that

10.3389/fpls.2023.1172354

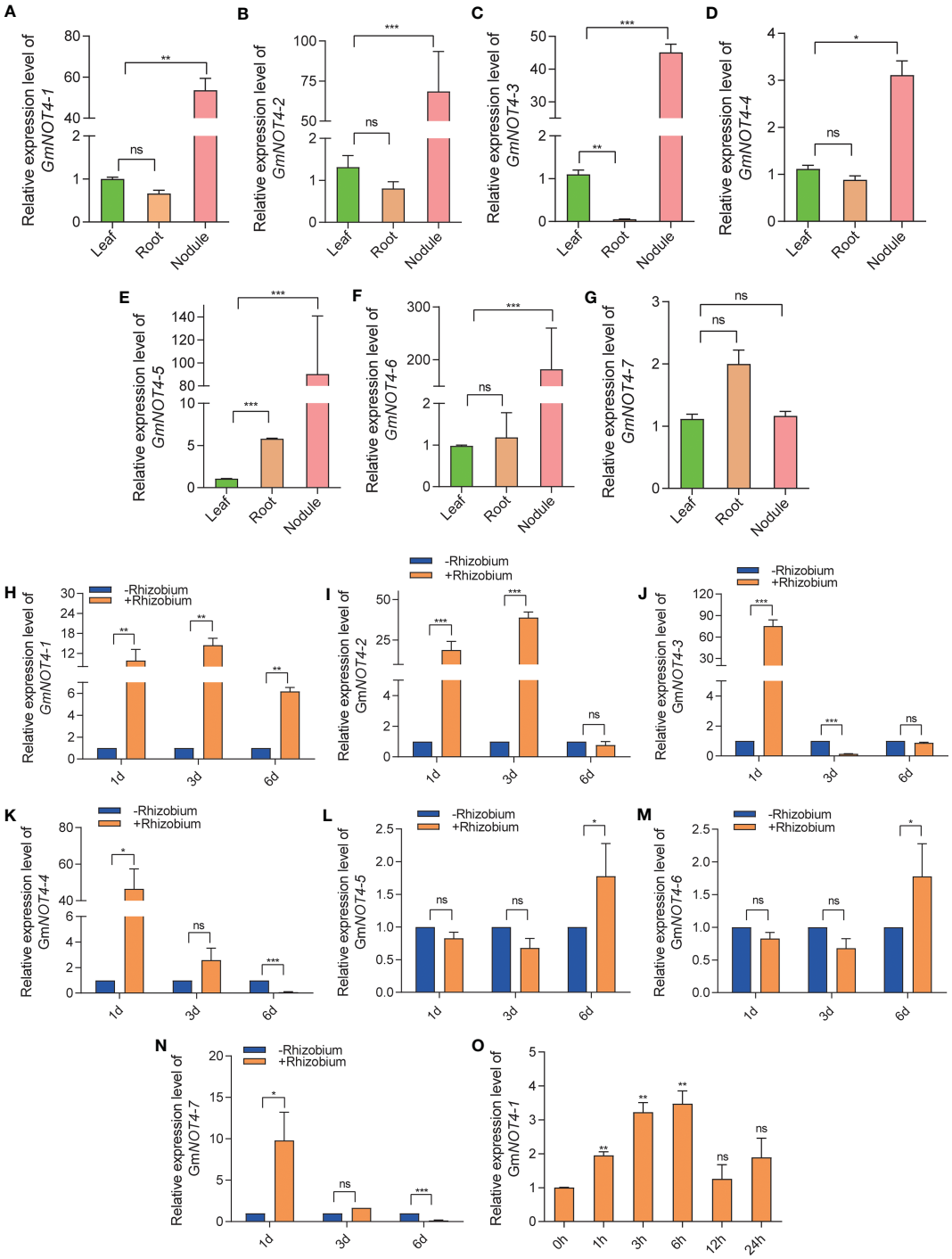


FIGURE 3
Expression pattern of soybean *GmNOT4s*. **(A–G)** show the relative expression level of *GmNOT4-1*, *GmNOT4-2*, *GmNOT4-3*, *GmNOT4-4*, *GmNOT4-5*, *GmNOT4-6*, *and GmNOT4-7* in soybean leaf, root, and nodule at 28 DAI, respectively. **(H–N)** show the relative expression level of *GmNOT4-1*, *GmNOT4-2*, *GmNOT4-3*, *GmNOT4-4*, *GmNOT4-5*, *GmNOT4-6*, *and GmNOT4-7* at 1, 3, and 6 DAI were validated by qPCR. **(O)** expression level of *GmNOT4-1* in soybean roots at 0, 1, 3, 6, 12, and 24 HAI (hours after inoculation). *GmCYP2* was used as an internal control. **(A–O)**, n = 12, Student's t-test; *P <0.05; **P <0.01; ***P <0.001, ns, no significance).

GmNOT4-1 also function as a biofunctional regulator in nodulation, both overexpression and CRISPR/Cas9 knock out of *GmNOT4-1* inhibited the number of nodules in soybean. Except for *GmNOT4-1*, we found all *GmNOT4s* responded to rhizobium at least at one time point (Figure 3). We do not exclude the possibility that different family members might have a role in a particular stage of root nodule symbiosis.

In previous studies, a series of transcriptional factors and microRNAs were identified (NIN, NSP1/2, miR172c, etc.) that modulate the expression of some key nodulation regulators

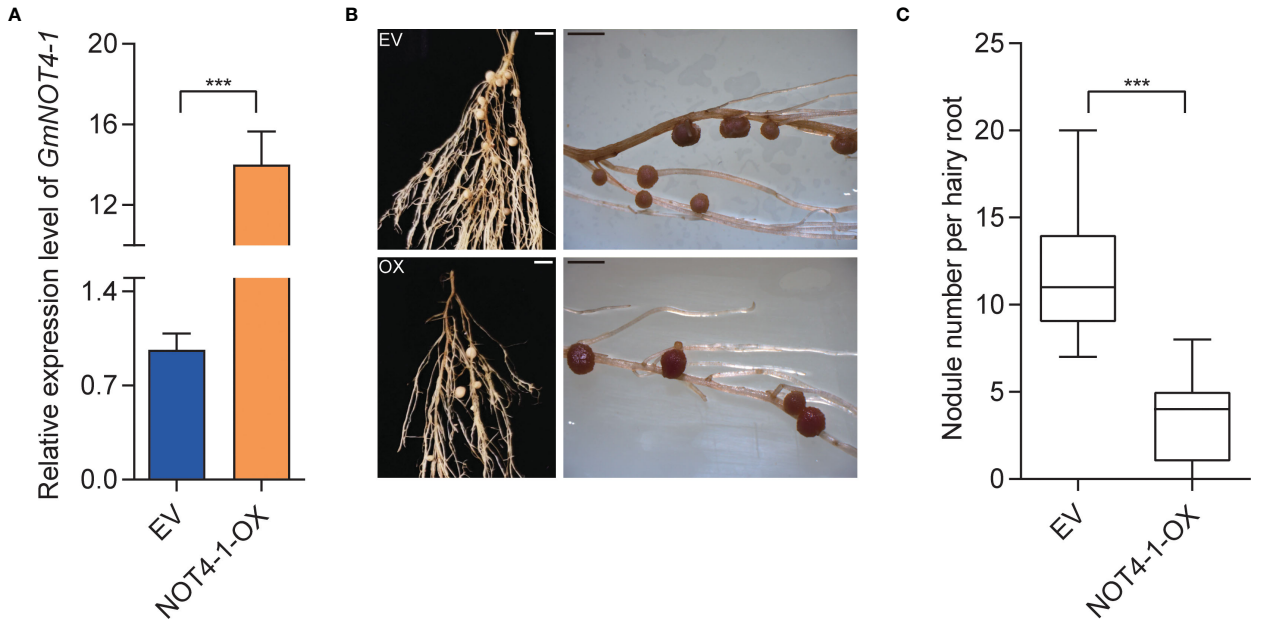


FIGURE 4
Phenotypic analysis of *GmNOT4-1* overexpression. **(A)** Expression level of transgenic hairy roots harboring empty vector and *35S:GmNOT4-1*. The expression levels were normalized against the housekeeping gene of soybean *GmCYP2*. Student's *t*-test was performed (***p <0.001, n = 15). **(B)** Nodule status of individual transgenic roots expressing EV1 and *35S:GmNOT4-1* at 28 DAI. Bar = 2 mm. **(C)**, Quantitative analysis of nodule number per hairy root carrying EV and *35S: GmNOT4-1* at 28 DAI. Values are the mean ± SD. A total of 20 hairy roots were collected for each biological replicate (n = 12, Student's *t*-test; ***p <0.001).

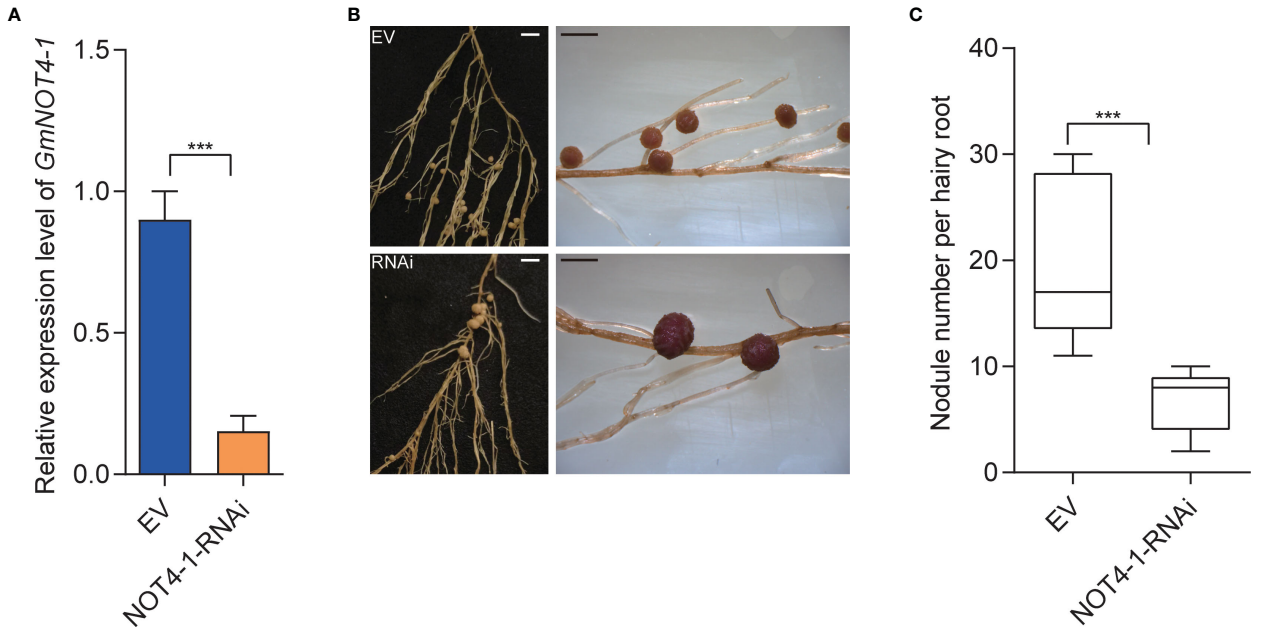


FIGURE 5
Knocking down *GmNOT4-1* inhibits nodulation. **(A)** qRT-PCR analysis of transgenic hairy roots harboring empty vector and *GmNOT4-1*-RNAi. The expression levels were normalized against the housekeeping gene of soybean *GmCYP2*. Student's *t*-test was performed (***p <0.001, n = 15). **(B)** Nodule status of individual transgenic roots expressing empty vector and *GmNOT4-1*-RNAi at 28 DAI. Bar = 2 mm. **(C)** Quantitative analysis of nodule number per hairy root carrying empty vector and *GmNOT4-1*-RNAi at 28 DAI. Values are the mean ± SD. A total of 20 hairy roots were collected for each biological replicate (n = 12, Student's *t*-test; ***p <0.001).

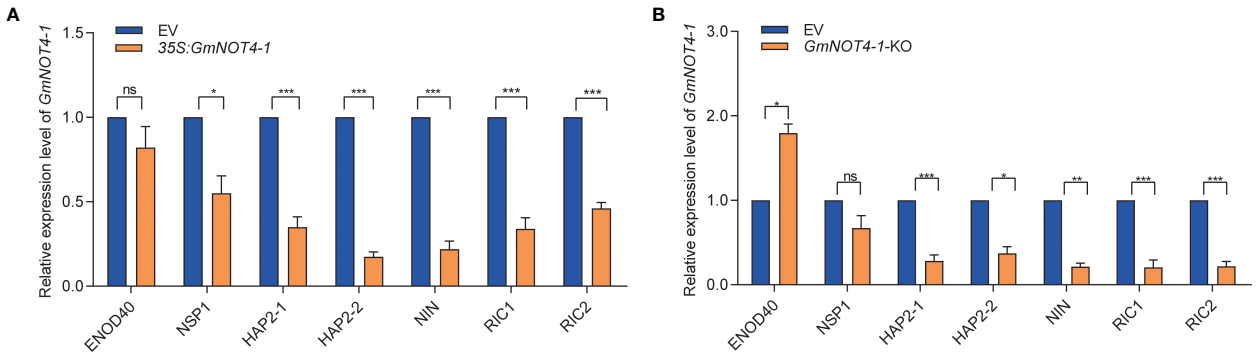


FIGURE 6
GmNOT4-1 expression alliterating inhibit the transcript levels of nodulation-related genes. **(A)** qRT-PCR analysis of *ENOD40*, *GmNINa*, *NSP1*, *HAP2-1*, *HAP2-2*, *GmRIC1*, *and GmRIC2* in roots transformed with empty vector and *GmNOT4-1* at 6 DAI (n = 6). **(B)** qRT-PCR analysis of *ENOD40*, *GmNINa*, *NSP1*, *HAP2-1*, *HAP2-2*, *GmRIC1*, *and GmRIC2* in roots transformed with empty vector and *GmNOT4-1* knock out at 6 DAI (n = 6). We set transcript level of the *ENOD40*, *GmNINa*, *NSP1*, *HAP2-1*, *HAP2-2*, *GmRIC1*, *and GmRIC2* at 6 DAI EV hairy roots as "1." The transcript amounts in each sample were normalized to those of *GmCYP2* (n = 6, Student's t-test; **p* <0.05, ***p* <0.01, and ****p* <0.001; ns, no significance).

(*ERN1*, *ENOD40*, *RICs*, etc.) and finally legume nodule number. In eukaryotic cells, except for transcriptional factors and miRNAs, the multiprotein complex CCR4–NOT also plays a vital role in regulating gene expression *via the* shortening of poly(A) tails of messenger RNA. In this study, we identified *GmNOT4-1*, one component of the soybean CCR4–NOT complex, whose expression homeostasis is important for soybean nodule initiation and subsequent nodule number (Figures 4, 5; S4-S6). We found that either overexpression or silencing of *GmNOT4-1* resulted in reduced expression of nodulation marker genes compared with empty vector control roots (Figure 6). *RIC1/2* was induced by nodule primordial formation to inhibit excessive nodulation; *GmNOT4-1* may function before nodule primordial formation; thus, the resulting *RIC1/2* induction and AON signaling were blocked by the overexpression or silencing of *GmNOT4-1*. Further study is needed to clarify the other components of the CCR4–NOT complex, including GmNOT4 family members in nodulation, and construct a direct link between the CCR4–NOT complex and nodulation signaling pathway genes. In addition, *GmNOT4-1* encodes a RING domain containing an E3 ligase; it may function by adjusting substrate protein levels. Identification of the target of GmNOT4-1 can further unveil its role in the process.

Conclusions

In this research, we identified seven members of the *NOT4* family in soybeans and found that *GmNOT4-1* was mainly expressed in soybean nodules. Interestingly, we observed that both overexpression and downregulation of *GmNOT4-1* inhibited the number of nodules in soybean. The CRISPR/Cas9 system was applied to validate this phenotype. Finally, we demonstrated that alterations in *GmNOT4-1* expression level repressed the expression of genes in the Nod factor signaling pathway. To our best knowledge, this is the first research to study the CCR4–NOT complex in legume nodulation. Like the founder transcription factor *NIN* (NODULE INCEPTION), essential for nodulation, the gene expression level of *NIN* was rigidly regulated for different processes of nodulation, including rhizobial infection, nodule organogenesis, and AON signaling. *GmNOT4-1* showed the same phenotype pattern as the *NIN* gene; overexpression and knockdown (out) both inhibit the proper nodulation. The further detailed phenotype and mechanism study will provide a better understanding of *GmNOT4-1* and its roles in nodulation.

This research will provide novel insight into the function of the *CCR4–NOT* family in legumes and reveal *GmNOT4-1* to be a valuable gene in regulating symbiotic nodulation.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

Author contributions

WLX, GSJ, and DWJ designed and conceived the study. ZJT cloned the gene and constructed the vector, grew the seedlings, and harvested them for quantitative detection. SLL completed the bioinformatics analysis. WDM and HL participated in the expression analysis. All authors contributed to the article and approved the submitted version.

Funding

This work is supported by the Natural Science Foundation of Henan Province (Grant No. 202300410056), the Science and Technology Innovation Young Talent Team of Shanxi Province (202204051001020), the Basic Research Program of Shanxi Province (20210302123365, 202103021224146), the National Natural Science Foundation of China (Grant No. 32241046),

the National Key Research and Development Program (Grant No. 2021YFD1600605-10), and the National Laboratory of Minor Crops Germplasm Innovation and Molecular Breeding (In preparation) (Grant Nos. 202105D121010-23 and 202204010910001-33).

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.2023.1172354/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Chromosomal distribution of *NOT4* gene family members *7 GmNOT4s* were identified from soybean genome and named *GmNOT4-1* to *GmNOT4-7* according to their distribution positions on chromosomes.

SUPPLEMENTARY FIGURE 2
Motif distribution of *GmNOT4s* NOT4 Protein sequences of Medicago truncatula (Mt), Phaseolus vulgaris (Pv), Arabidopsis thaliana (At), Glycine max (Gm) and Oryza sativa (Os) were subjected into MEME online software to identify conserved motifs. 10 motifs were found in NOT4s genes and represented by different colored boxes and corresponding motif number, conserved amino acid sequence themes were shown below. (The left panel) shown the phylogenetic tree of group I~III genes in purple, blue, and orange colors; (The middle panel) shown the conserved motifs distribution in NOT4s; (The right panel) shown the gene structure of untranslated regions (UTRs) of the NOT4s gene are shown as green, yellow boxes and black lines.

SUPPLEMENTARY FIGURE 3
Cis-acting elements of soybean *GmNOT4s* gene promoter Promoter sequences (2000 base pair upstream of start codon site) of *GmNOT4s* were submitted to PlantCARE database to identify *cis*-acting elements. Boxes with different colors indicate various *cis*-acting elements located in the promoter regions of *GmNOT4s*. Abbreviations: AT~TATA-box, CAAT-box, TATA-box, TATA: core *cis*-elements; Box 4, GT1-motif, ACE, G-box: optical response element; CAT-box, CCGTCC-box: associated with meristem expression; circadian: circadian rhythm control; GCN4-Motif: endosperm expression; HD-Zip1: cell differentiation in palisade tissue; AC-I: xylem synthesis; MBSI: flavonoid biosynthesis; TGA-element: Auxin responsive element; ABRE, AAGAA-motif, ABRE3a: abscisic acid reaction; TCA, TCA-element: salicylic acid responsive elements; P-box, GARE-motif: gibberellin response element; TGACG-Motif, CGTCA-motif: methyl jasmonate responsive elements; ERE: ethylene response element; ARE: anaerobic induction element; MYB, Myb-binding site, MYC, MBS: Drought induction elements; LTR: low temperature element; TC rich repeats, CCAAT-box, as-1, STRE: stress reaction elements; WRE3, WUN-motif, W box: traumatic stress elements; A-box, CARE, CCGTCC motif, AT-rich sequence: unknown function.

SUPPLEMENTARY FIGURE 4
No combined and off-target effects existed in GmNOT4-1 overexpression or knoct-out roots (A), qRT-PCR analysis of GmNOT4-2, GmNOT4-3, GmNOT4-4, GmNOT4-5, GmNOT4-6 and GmNOT4-7 in roots transformed with empty vector and GmNOT4-1 overexpression at 6 DAI (n = 6). (B), qRT-PCR analysis of GmNOT4-2, GmNOT4-3, GmNOT4-4, GmNOT4-5, GmNOT4-6 and GmNOT4-7 in roots transformed with empty vector and GmNOT4-1-KO at 6 DAI (n = 6). We set transcript level of the GmNOT4-2, GmNOT4-3, GmNOT4-4, GmNOT4-5, GmNOT4-6 and GmNOT4-7 at 6 DAI EV hairy roots as "1". The transcript amounts in each sample were normalized to those of GmCYP2 (n = 12, Student's t-test; "ns" =No significance).

SUPPLEMENTARY FIGURE 5
GmNOT4-1 regulate nodulation though affecting rhizobial infection (A), qRT-PCR analysis of GmNPL, GmRPG, GmVPY, GmCYCLOPS and GmSCARN in roots transformed with empty vector and GmNOT4-1 overexpression at 6 DAI (n = 6). (B), qRT-PCR analysis of GmNPL, GmRPG, GmVPY, GmCYCLOPS and GmSCARN in roots transformed with empty vector and GmNOT4-1-KO at 6 DAI (n = 6). The transcript amounts in each sample were normalized to those of GmCYP2 (n = 12, Student's t-test; *p < 0.05, **p < 0.01, and ***p < 0.001).

SUPPLEMENTARY FIGURE 6
GmNOT4-1 overexpression and knock out decreases the number of deformed root hairs **(A)**, Expression level of transgenic hairy roots harboring empty vector and 35S:GmNOT4-1. The expression levels were normalized against the housekeeping gene of soybean GmCYP2. Student's t-test was performed (***p < 0.001, n = 20). **(B)**, At 6 DAI, 2 cm root segments of hairy roots overexpressing GmNOT4-1 or expressing EV below the root-hypocotyl junction were cut and stained with 1% (w/v) methylene blue. Deformed root hairs were counted (n=20). (B) Root hair deformation in transgenic roots harboring EV and 35S:GmNOT4-1 vector. Bar=40 μm. **(C)**, Quantification of deformed root hairs in the transgenic lines (n=10 to 12). Values are averages ± SD from three independent experiments. Asterisks represent statistically significant differences. (n = 20, Student's t-test; ***p < 0.001). **(D)**, Expression level of transgenic hairy roots harboring empty vector and GmNOT4-1-KO. The expression levels were normalized against the housekeeping gene of soybean GmCYP2. Student's t-test was performed (***p < 0.001, n = 20). **(E)**, Root hair deformation in transgenic roots harboring EV and GmNOT4-1-KO. Bar=40 μm. **(F)**, Quantification of deformed root hairs in the transgenic root harboring EV and GmNOT4-1-KO (n=20). Values are averages ± SD from three independent experiments. Asterisks represent statistically significant differences. (n = 20, Student's t-test; ***p < 0.001).

SUPPLEMENTARY FIGURE 7
Nodules of GmNOT4-1 overexpression and RNAi roots are functional Nodule performance of mature nodule in roots expressing empty vector, overexpression, and GmNOT4-1-KO at 28 DAI. Bar =2 mm.

SUPPLEMENTARY FIGURE 8
Phenotypic analysis of GmNOT4-1 gene editing **(A)**, Relative expression level of GmNOT4-1 in GmNOT4-1 edited roots; The expression levels were normalized against the housekeeping gene of soybean GmCYP2. Student's t-test was performed (***p < 0.001, n = 15). **(B)**, Nodule performance of individual transgenic roots expressing empty vector and GmNOT4-1-KO at 28 DAI. Bar =2 mm. **(C)**, Quantitative data of nodule number per hairy root carrying empty vector and GmNOT4-1-KO at 28 DAI. Values are the mean ± SD. 20 hairy roots were collected for each biological replicate. (Student's t-test; ***p < 0.001).

SUPPLEMENTARY FIGURE 9
Alignment of GmNOT4-1 gRNAs with other GmNOT4 family members **(A)**, The front gRNA sequence aligned with the corresponding region of other GmNOT4 family members. **(B)**, The back gRNA sequence aligned with the corresponding region of other GmNOT4 family members, the software Daneman was applied for the sequence alignment.

References

Andriankaja, A., Boisson-Dernier, A., Frances, L., Sauviac, L., Jauneau, A., Barker, D. G., et al. (2007). AP2-ERF transcription factors mediate nod factor dependent Mt ENOD11 activation in root hairs *via* a novel cis-regulatory motif. *Plant Cell* 19 (9), 2866–2885. doi: 10.1105/tpc.107.052944

Carroll, R. K., Weiss, A., Broach, W. H., Wiemels, R. E., Mogen, A. B., Rice, K. C., et al. (2016). Genome-wide annotation, identification, and global transcriptomic analysis of regulatory or small RNA gene expression in staphylococcus aureus. *mBio* 7 (1), e01990–e01915. doi: 10.1128/mBio.01990-15

Chou, M. X., and Wei, X. Y. (2010). Review of research advancements on the molecular basis and regulation of symbiotic nodulation of legumes. *Chin. J. Plant Ecol.* 7), 876–888. doi: 10.3773/j.issn.1005-264x.2010.07.013

Chun, L., Mi, G. H., Li, G. S., Chen, F. G., and Zhang, F. S. (2005). Genetic analysis of maize root characteristics in response to low nitrogen stress. *Plant Soil* 276 (1-2), 369–382. doi: 10.1007/s11104-005-5876-2

Fonouni-Farde, C., Diet, A., and Frugier, F. (2016). Root development and endosymbioses: DELLAs lead the orchestra. *Trends Plant Sci.* 21 (11), 898–900. doi: 10.1016/j.tplants.2016.08.012

Fujita, H., and Kawaguchi, M. (2002). A lotus basic leucine zipper protein with a RING-finger motif negatively regulates the developmental program of nodulation. *Proc. Natl. Acad. Sci. United States America* 99 (No.23), 15206–15210. doi: 10.1073/pnas.222302699

Gao, L. F., Hu, Z. A., Wang, H. X., and Yang, W. L. (2002). Structure and function of Rhizobium Nod factors. *Chinese Bulletin of Life Sci.* (01), 17–19.

Heckmann, A. B., Sandal, N., Bek, A. S., Madsen, L. H., Jurkiewicz, A., Nielsen, M. W., et al. (2011). Cytokinin induction of root nodule primordia in lotus japonicus is regulated by a mechanism operating in the root cortex. *Mol. Plant-Microbe Interact.* 24 (No.11), 1385–1395. doi: 10.1094/mpmi-05-11-0142

Horváth, B., Yeun, L. H., Domonkos, A., Halász, G., Gobbato, E., Ayaydin, F., et al. (2011). Medicago truncatula IPD3 is a member of the common symbiotic signaling pathway required for rhizobial and mycorrhizal symbioses. *Molecular Plant-Microbe Interactions : MPMI* 24 (11), 1345. doi: 10.1094/MPMI-01-11-0015

Howard, J. (2010). A phylogenetic approach to examining symbiotic specificity and evolution in the legume-rhizobia nitrogen-fixing symbiosis. Diss. Arizona State University.

Jia, Z. R., Zhang, M. J., and Yang, W. D. (2017). Effects of nitrogen fertilizer on nitrogen uptake and utilization of intercropped maize. *Shanxi Agric. Sci.* 45 (12), 1960–1964. doi: 10.3969/j.issn.1002-2481.2017.12.16

Jian, B., Liu, B., Bi, Y., Hou, W., Wu, C., and Han, T. (2008). Validation of internal control for gene expression study in soybean by quantitative real-time PCR. *BMC Mol. Biol.* 9 (1), 1–14. doi: 10.1186/1471-2199-9-59

Jin, Y., Liu, H., Luo, D., Yu, N., Dong, W., Wang, C., et al. (2016). DELLA proteins are common components of symbiotic rhizobial and mycorrhizal signalling pathways. *Nat. Commun.* 7, 12433. doi: 10.1038/ncomms12433

Kaló, P., Gleason, C., Edwards, A., Marsh, J., Mitra, R. M., Hirsch, S., et al. (2005). Nodulation signaling in legumes requires NSP2, a member of the GRAS family of transcriptional regulators. *Sci. (New York N.Y.)* 308 (No.5729), 1786–1789. doi: 10.1126/science.1110951

Kiss, E., Olaíh, B. R., Kaloí, P. T., Morales, M., Heckmann, A. B., Borbola, A., et al. (2010). LIN, a novel type of U-box/WD40 protein, controls early infection by rhizobia in legumes. *Plant Physiol.* 151 (No.3), 1239–1249. doi: 10.1104/pp.109.143933

Laloum, T., De Mita, S., Gamas, P., Baudin, M., et al. (2013). CCAAT-box binding transcription factors in plants: y so many? *Trends Plant Sci.* 18 (No.3), 157–166. doi: 10.1016/j.tplants.2012.07.004

Liang, W., Li, C., Liu, F., Jiang, H., Li, S., Sun, J., et al. (2008). The arabidopsis homologs of CCR4-associated factor 1 show mRNA deadenylation activity and play a role in plant defence responses. *Cell Res.* 19, 307–316. doi: 10.1038/cr.2008.317

Middleton, E. T., Steel, S. A., and Doherty, S. M. (2007). The effect of prior bisphosphonate exposure on the treatment response to teriparatide in clinical practice. *Calcif Tissue Int.* 81 (5), 335–340. doi: 10.1007/s00223-007-9066-5

Nishimura, R., Ohmori, M., and Kawaguchi, M. (2002). The novel symbiotic phenotype of enhanced-nodulating mutant of lotus japonicus: astray mutant is an early nodulating mutant with wider nodulation zone. *Plant Cell Physiol.* 43 (No.8), 853–859. doi: 10.1093/pcp/pcf098

Oldroyd, G. E. D., and Downie, J. A. (2008). Coordinating nodule morphogenesis with rhizobial infection in legumes. *Annu. Rev. Plant Biol.* 59 (No.1), 519–546. doi: 10.1146/annurev.arplant.59.032607.092839

Patra, R. K., Pant, L. M., and Pradhan, K. (2017). Response of soybean to inoculation with rhizobial strains: effect on growth, yield, n uptake and soil n status. *World J. Agric. Sci.* 8 (1), 28–31. doi: 10.11942/j.issn1002-2767.2017.08.0028

Ren, B., Wang, X., Duan, J., and Ma, J. (2019). Rhizobial tRNA-derived small RNAs are signal molecules regulating plant nodulation. *Science* 365 (6456), 919–922. doi: 10.1126/science.aav8907

Shimomura, K., Nomura, M., Tajima, S., and Kouchi, H. (2006). LjnsRING, a novel RING finger protein, is required for symbiotic interactions between mesorhizobium loti and lotus japonicus. *Plant &Cell Physiol.* 47 (No.11), 1572–1581. doi: 10.1093/pcp/pcl022

Smit, P., Raedts, J., Portyanko, V., Debellé, F., Gough, C., Bisseling, T., et al. (2005). NSP1 of the GRAS protein family is essential for rhizobial nod factor-induced transcription. *Sci. (New York N.Y.)* 308 (No.5729), 1789–1791. doi: 10.1126/science.1111025

Suzuki, Y., Arae, T., Green, P. J., Yamaguchi, J., and Chiba, Y. (2015). AtCCR4a and AtCCR4b are involved in determining the Poly(A) length of granule-bound starch synthase 1 transcript and modulating sucrose and starch metabolism in arabidopsis thaliana. *Plant Cell Physiol.* 56 (No.5), 863–874. doi: 10.1093/pcp/pcv012

Vernié, T., Camut, S., Camps, C., Rembliere, C., Carvalho-Niebel, F., Mbengue, M., et al. (2016). PUB1 interacts with the receptor kinase DMI2 and negatively regulates rhizobial and arbuscular mycorrhizal symbioses through its ubiquitination activity in *Medicago truncatula*. *Plant Physiol.* 170 (No.4), 2312–2324. doi: 10.1104/pp.15.01694

Walley, J. W., Kelley, D. R., and Nestorova, G. (2010). Arabidopsis deadenylases AtCAF1a and AtCAF1b play overlapping and distinct roles in mediating environmental stress responses. *Plant Physiol.* 152 (No.2), 866–875. doi: 10.1104/pp.109.149005

Wang, L., Sun, Z., Su, C., Wang, Y., Yan, Q., Chen, J., et al. (2019). A GmNINa-miR172c-NNC1 regulatory network coordinates the nodulation and autoregulation of nodulation pathways in soybean. *Mol. Plant* 12 (9), 1211–1226. doi: 10.1016/j.molp.2019.06.002

Wang, Y., Wang, L., Zou, Y., Chen, L., Cai, Z., Zhang, S., et al. (2014). Soybean miR172c targets the repressive AP2 transcription factor NNC1 to activate ENOD40 expression and regulate nodule initiation. *PLANTCELL* 26 (No.12), 4782–4801. doi: 10.1105/tpc.114.131607

Yuan, S. L., Li, R., Wang, L., Chen, H. F., Zhang, C. J., Chen, L. M., et al. (2016). Cystatin genes in the genome of soybean (*Glycine max*). *Front. Plant Sci.* 7. doi: 10.11974/nyyjs.20171131011

frontiers | Frontiers in Plant Science

TYPE Opinion
PUBLISHED 20 June 2023
DOI 10.3389/fpls.2023.1216337

OPEN ACCESS

EDITED BY
Kaixuan Zhang,
Chinese Academy of Agricultural Sciences, China

REVIEWED BY
Federico Scossa,
Research Center for Genomics and Bioinformatics (CREA-GB), Italy

*CORRESPONDENCE
Dominik K. Großkinsky
dominik.grosskinsky@ait.ac.at
Claudia Jonak
claudia.jonak@ait.ac.at

RECEIVED 03 May 2023
ACCEPTED 08 June 2023
PUBLISHED 20 June 2023

CITATION
Großkinsky DK, Faure J-D, Gibon Y, Haslam RP, Usadel B, Zanetti F and Jonak C (2023) The potential of integrative phenomics to harness underutilized crops for improving stress resilience. *Front. Plant Sci.* 14:1216337. doi: 10.3389/fpls.2023.1216337

The potential of integrative phenomics to harness underutilized crops for improving stress resilience

Dominik K. Großkinsky [1*], Jean-Denis Faure [2], Yves Gibon [3,4], Richard P. Haslam [5], Björn Usadel [6,7], Federica Zanetti [8] and Claudia Jonak [1*]

[1]AIT Austrian Institute of Technology, Center for Health and Bioresources, Bioresources Unit, Tulln a. d. Donau, Austria, [2]Université Paris-Saclay, INRAE, AgroParisTech, Institut Jean-Pierre Bourgin, Versailles, France, [3]INRAE, Univ. Bordeaux, UMR BFP, Villenave d'Ornon, France, [4]Bordeaux Metabolome, INRAE, Univ. Bordeaux, Villenave d'Ornon, France, [5]Rothamsted Research, Harpenden, Hertfordshire, United Kingdom, [6]IBG-4 Bioinformatics, CEPLAS, Forschungszentrum, Jülich, Germany, [7]Biological Data Science, Heinrich Heine University, Universitätsstrasse 1, Düsseldorf, Germany, [8]Department of Agricultural and Food Sciences (DISTAL), Alma Mater Studiorum - Università di Bologna, Bologna, Italy

KEYWORDS

agriculture, breeding, climate change resilience, crop improvement, neglected crops, phenomics, sustainability

1 Introduction

The current agricultural and food system faces diverse and increasing challenges. These include feeding an ever-growing human population, expected to reach about 10 billion by 2050 combined with societal disruption, and the need to cope with the impact of climate change (FAO, 2022). Given that future environmental conditions will limit crop productivity (Zhao et al., 2017; Cooper et al., 2021) and the limited potential to continually increase the performance of staple crops by conventional breeding (Hickey et al., 2019), there is an urgent need to transform agricultural systems. Central to this transformation is the application of alternative, accelerated, and sustainable approaches for the improvement and development of underutilized crops (Hickey et al., 2019). Modern breeding strategies for major crops have widely integrated novel technologies, such as advanced phenotyping or genome-wide interactions, and even epigenomics within "beyond the gene" strategies (Crisp et al., 2022) to speed up crop/genotype selection (Hickey et al., 2019; Kumar et al., 2023). Deploying phenotyping at different scales has the potential to identify novel trait(s) components that can be targeted to accelerate crop improvement (Araus and Cairns, 2014; Großkinsky et al., 2015b; Zhao et al., 2019; Varshney et al., 2021). There is even greater potential for these technologies when used to improve underutilized crops and support the agricultural transformation, as underutilized crops typically lack a biased breeding/selection history, i.e., they often exhibit a high genetic diversity and potential, and are usually better adapted to

challenging environments (Kumar et al., 2021; Kumar et al., 2023). To illustrate the application of an integrative phenomics approach we discuss how combining multi-omics and advanced phenotyping is being applied to the underutilized oilseed crop *Camelina sativa* (camelina, gold-of-pleasure, false flax) to facilitate the generation of climate-smart crops for future agricultural systems.

2 Novel technologies and approaches for crop improvement

The time needed to select new varieties or species, but also to set up new agronomic practices (e.g., reduced input management) is currently too long in view of the rapidly changing climate[1]. While traditionally varieties have often been selected in a context of high input mechanized monoculture, it is now necessary to develop predictive strategies to formulate the genotypes and agronomic practices adapted to low input, sustainable and diversified cultures. That said, plant biologists have long been studying adaptation or acclimation to stresses associated with climate change (e.g., drought, heat, attacks from pests and pathogens) to unravel the involved mechanisms and identify targets to improve crop performance.

Reverse genetics usually starts from a gene and characterizes associated phenotypes. Here, the development of "omics" has revolutionized biology over the past 20 years and led to a dramatic increase in the number of candidate genes and networks (e.g., Kajala et al., 2021). Great progress was initially achieved in Arabidopsis but has increasingly been translated to major crops (Fernie and Tohge, 2017). Bottom-up systems biology goes one step further by assembling reductionist knowledge using models to explain phenotypes (Keurentjes et al., 2011). One consequence of this is that more complex mechanisms can be analysed, and more sophisticated improvement strategies proposed (Chen et al., 2021). This ensemble of approaches is being carried out in a growing number of species (Stevens et al., 2018; Hawkins et al., 2021). However, when going beyond model plants and major established crops, challenges remain. Whilst resolving homozygous genomes (van Rengs et al., 2021) and pangenomes (Bayer et al., 2020) of crops with a medium genome size is now achievable with the advent of novel sequencing technologies (Mascher et al., 2021), challenges remain with consistent gene annotation and ontology developments. Rapid progress is being made with more mature tools, facilitating structural (Holst et al., 2023) and functional genome annotation (Bolger et al., 2018; Schwacke et al., 2019), whilst plant genome data continues to be gathered in comprehensive databases (Vandepoele, 2017). However, several problems remain outstanding, including the accurate description of phenotypes together with necessary metadata (Papoutsoglou et al., 2020) and simple data management approaches (Pommier et al., 2019; Jacob et al., 2020). This is especially true for minor or even orphan crops where, without the necessary resources, these strategies have not yet reached the same level of maturity as for major field crops (Watt et al., 2020). To address this, remote sensing, deep phenotyping, and machine learning approaches are continually being developed for crops (Fernie et al., 2021; Wong et al., 2023).

Forward genetics starts from the phenotype and tries to identify the underlying mechanisms, genes and/or markers. Traditionally, crop improvement has been based on traits of interest, especially yield. The most promising lines are retained, which results in an enrichment of the germplasm with performance-promoting alleles. More recently, proxies such as plant height have started to be used to speed up the process (Madec et al., 2017). The aim being to find phenotype-based parameters that are easy and cheap to measure. Classical selection in the field using plots is very expensive and generally subject to seasons and environmental hazards. It is in this context that multiple phenotyping facilities have been established[2]. By enabling plant phenotyping under controlled conditions, their purpose is to generate field performance proxies at affordable cost (Reynolds et al., 2019), although the confirmation of agronomic trait evaluations carried out under controlled conditions in the field is challenging (Hohmann et al., 2016). Indeed, when transferring phenotypic information from controlled environments to agronomic field conditions, phenotypes are not always confirmed (Shi et al., 2017; Zhang et al., 2019; Metje-Sprink et al., 2020). For example, evaluations of edited plants that have targeted mutations for specific and well-characterized traits have shown limitations, as only 13 out of 22 publications confirmed yield effects of edited plants identified in controlled environments in subsequent field trials (Metje-Sprink et al., 2020). Therefore, field trials remain essential for the study of traits interacting with complex environments. Phenotyping under controlled conditions requires continued evaluation of growth scenarios to determine the best predictive variables. Over the past decade more and more sophisticated sensors and platforms have emerged. Whether in fields or greenhouses, they use an ever-increasing number of imaging techniques (Großkinsky et al., 2015a; Brichet et al., 2017; Yang et al., 2020; Dodig et al., 2021) as non-destructive tools that allow the generation of performance proxies throughout plant growth (e.g., Millet et al., 2019). Genomic selection (e.g., R2D2 Consortium et al., 2021), increasingly used by breeders, consists of predicting the genetic value of candidates for selection through the development of models linking phenotype and genotype. The advantage of this approach is that the phenotyping step can be done with a small subset of the germplasm and the search for potentially interesting genotypes (pre-breeding) can be done with young plants and therefore at low cost. Paradoxically, sampling of plant material is still largely manual and the development of robots for destructive-type observations is still in its infancy (Foix et al., 2018). This may be one of the reasons why "omics" are not yet fully considered as selection tools, in particular metabolomics, whose costs have fallen sharply (López-Ruiz et al., 2019). Evidence continues to emerge demonstrating that transcriptomics or metabolomics (Westhues et al., 2017) can carry more condensed information than the genome, and these approaches can be used to build top-down models capable of predicting traits, particularly in

1 https://climate.copernicus.eu/esotc/2022

2 www.plant-phenotyping.org

integrative approaches combining omics technologies with advanced phenotyping.

3 Underutilized crops and their agricultural potential

Despite a huge variety of cultivated plants, historical developments in global agriculture have resulted in a very narrow diversity of crops contributing to our current food system, with only few species accounting for over 90% of world food production, e.g., the staple cereals rice, wheat, and maize (Milla and Osborne, 2021). In contrast, the vast majority of more than 5000 known edible crops are only cultivated on a small scale, if at all, and are largely neglected in modern agricultural settings due to their limited competitiveness (Padulosi, 2017). These wild, cultivated, or semi-domesticated crops are not yet fully integrated into typical agricultural systems worldwide (Padulosi, 2017). These crops are typically underutilized in agriculture due to a lack of improvement and in-depth agronomic knowledge and/or experience, moreover, they are poorly represented in fundamental research, which is mostly focused on model plants or staple crops. Thus, there is often a lack of complementary knowledge, technologies, and resources from basic research to support use and integration of such crops. Underutilised crops (for examples see Table 1) have recently gained attention as many of them hold untapped potential for coping with the challenges of current and future agriculture (Padulosi, 2017; Kumar et al., 2023), e.g., increasing agro-biodiversity, improving the overall resilience of agricultural systems against climate change and sourcing alternative (high-quality) food ingredients. This is evident in research today e.g., EU-funded projects, aiming to promote the use of underutilized crops, such as INCREASE[3] aiming at improving the sustainable use of the pulses chickpea, common bean, lentil and lupin, PROTEIN2FOOD[4] addressing the development and sustainable use of quinoa, amaranth, buckwheat and pulses like lupins, and UNTWIST[5] aiming at the promotion of the sustainable use of camelina as a promising climate-resilient oilseed crop. All these projects focus on crop phenotypes and environmental plasticity.

Among underutilized crops with significant potential, camelina has recently attracted interest as a promising alternative to other oilseeds (e.g., rapeseed). From a scientific perspective, this is also due to its genetic similarity with the model plant Arabidopsis (Kagale et al., 2014). Camelina can be easily integrated into existing agricultural systems (Zanetti et al., 2021). It has a natural capacity to withstand harsh environmental conditions, and tolerance to pests and diseases represent additional features that make it a suitable crop for many environments worldwide (Berti et al., 2016; Zanetti et al., 2021), except for too wet conditions (Stasnik et al., 2022). This particularity also makes it a potential source of resistance mechanisms that would have disappeared or have become insufficient in species that have been selected above all for their yield in optimal conditions. So, the aim here is not to rediscover mechanisms that are already known. The emerging interest in camelina has spurred genomic and genetic investigations, including the description of a reference genome (Kagale et al., 2014). The development of Recombinant Inbred Lines (RILs) has further allowed the identification of important QTLs e.g., 1000-seed weight, seed number per pod, days to flowering or seed oil content (King et al., 2019; Li et al., 2021). Nevertheless, the throughput of these approaches is limited particularly at the phenotype level, as for other (underutilized) crops. Hence, further focused improvement of camelina is required and the use of novel approaches and technologies, specifically integrating the crop phenotype, is necessary and achievable as respective protocols for high-throughput phenotyping are established (Vello et al., 2022).

4 Integrative phenomics to unlock the potential of the oilseed crop camelina as a blueprint to improve stress tolerance

Several projects are emerging that take advantage of the availability and affordability of modern technologies (e.g., multi-omics, phenomics, network discovery) to investigate plant plasticity in novel crops (e.g., Girija et al., 2021; Marks et al., 2021). Understanding plant plasticity is key for the development of novel crops for resilient food production systems in increasingly variable climates. As an exemplary project, UNTWIST[5] targets camelina to develop a systems-based understanding of the key mechanisms underlying crop stress tolerance using a phenomics approach. The project is based on the premise that the unravelling of stress adaptation mechanisms in naturally resilient camelina (Zanetti et al., 2021), which has not yet undergone intensive breeding, will reveal successful stress adaptation strategies. Research is centred on agronomically relevant stress scenarios (linking experiments in controlled environments and field conditions) to improve our understanding of plant responses to challenging growth conditions. Therefore, diverse germplasm is analysed at the genomic, epigenetic, transcriptional, physiological, proteomic, and metabolomic level (incl. lipidome, redox status) complemented by agronomic parameters and advanced non-invasive multi-sensor phenotyping. Despite limited basic knowledge on camelina compared to well-studied plant species, this integrative multi-omics approach deploying phenotyping from cell to whole plant level (Dhont et al., 2013; Ghanem et al., 2015; Großkinsky et al., 2015b) allows the identification and dissection of mechanisms contributing to camelina performance under challenging environments representative of future conditions. Integration of the multi-layered data serves modelling approaches and supports potential marker development for crop improvement under adverse climate conditions. Furthermore, the results will underpin databases, protocols, and advice for downstream

3 www.pulsesincrease.eu

4 www.protein2food.eu

5 www.untwist.eu

TABLE 1 Examples of promising underutilized crops exploitable for agricultural use.

Underutilized crop	Origin	Alternative to	Purpose / Advantage	Suitable for
Amaranth	Global	Wheat/cereals	Alternative grain, protein source, gluten-free	Dry, hot tropical to temperate climates
Buckwheat	Asia	Cereals	Alternative grain, protein source, gluten-free	Cool, moist climates
Camelina	Europe	Rapeseed, soybean	Alternative oilseed, high oil quality, protein source, climate-resilient	Marginal land/poor soils, dry, warm to cold, not too moist climates
Carinata	Africa	Rapeseed, soybean	Alternative oilseed, aviation/biofuel production	Marginal land, dry climates
Cowpea	Africa, Asia	Other legumes	Alternative protein	Dry climates, sandy soils
Cucurbits	Americas	Oilseeds, (sweet) potato	High-quality seed oil, nutritional value of flesh	Warm, dry, not too moist climates
Linseed	Global	Oilseeds, cotton	High-quality oil, natural fibres	Wet, not too warm climates
Lupin	Africa, Andes, Europe	Other legumes	Alternative protein	Dry, not too warm climates, high sun exposure, acidic soils (some types)
Millets	Global	Cereals	High nutritional value, climate-resilient, gluten-free	Dry, hot climates, areas of high flood risk
Pennycress	Americas, Eurasia	Oilseeds	Alternative oilseed, protein source	Wet, cold climates
Quinoa	Andes	Wheat/cereals	Alternative grain, protein source, gluten-free	Dry, not too warm climates, mountainous regions, saline soils
Sorghum	Global	Cereals	Alternative grain, gluten-free	Hot, dry climates, saline soils
Teff	Africa	Wheat/cereals	Alternative grain, protein source, gluten-free	Dry, hot climates, poor soil, areas of high waterlogging risk

application. The new knowledge can be utilized for further improvements of camelina and other (underutilized) crops in combination with modern strategies in breeding and engineering crops (Jiang et al., 2017; Morineau et al., 2017; Lemmon et al., 2018; Fernie and Yan, 2020; Rönspies et al., 2021; Yu et al., 2021; Bellec et al., 2022; Han et al., 2022).

5 Discussion

Global agriculture urgently requires alternative strategies in crop development and improvement. So far, the great potential of underutilized crops to contribute to the agricultural transition in our food systems remains untapped. This is due to a lack of knowledge and appropriate technologies; shortfalls that are now being addressed. Recognizing that phenotypes are as or even more important than genotypic information has driven recent advances in crop phenotyping. These advances greatly contribute to the development of multi-omics systems-based phenomics approaches, which have shown their potential for accessing previously untapped resources and can be applied more directly to underutilized non-staple crops in the future, significantly reducing the dependency on information from staple crops and model plants as the only information source. This allows the effective improvement and utilization of such crops to support the re-configuration of agricultural and food systems in a sustainable manner. Most notably, phenomics and multi-omics approaches will enable agriculture to harness the potential of underutilized crops. Projects like UNTWIST, which apply these approaches, are needed to prove, and validate the potential of these novel tools. Ultimately, such projects can serve as a blueprint for improving underutilized crops and provide a platform for establishing more climate-resilient food production systems.

Author contributions

All authors jointly defined the outline of the manuscript. DKG drafted the manuscript with input from all authors. All authors contributed to the article and approved the submitted version.

Funding

This work was performed within the UNTWIST project (www.untwist.eu) funded by the European Commission Horizon 2020 Research & Innovation Programme (Grant Agreement No 862524). Rothamsted Research (UK) receives grant-aided support from the BBSRC (UK) and RH was partially supported by the BBSRC (UK) Institute Strategic Programme Green Engineering (BB/X010988/1). BU was partially supported by DFG grant 456082119.

Conflict of interest

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

Publisher's note

References

Araus, J. L., and Cairns, J. E. (2014). Field high-throughput phenotyping: the new crop breeding frontier. *Trends Plant Sci.* 19, 52–61. doi: 10.1016/j.tplants.2013.09.008

Bayer, P. E., Golicz, A. A., Scheben, A., Batley, J., and Edwards, D. (2020). Plant pan-genomes are the new reference. *Nat. Plants* 6, 914–920. doi: 10.1038/s41477-020-0733-0

Bellec, Y., Guyon-Debast, A., François, T., Gissot, L., Biot, E., Nogué, F., et al. (2022). New flowering and architecture traits mediated by multiplex CRISPR-Cas9 gene editing in hexaploid *Camelina sativa*. *Agronomy* 12, 1873. doi: 10.3390/agronomy12081873

Berti, M., Gesch, R. W., Eynck, C., Anderson, J., and Cermak, S. (2016). Camelina uses, genetics, genomics, production, and management. *Ind. Crops Products* 94, 690–710. doi: 10.1016/j.indcrop.2016.09.034

Bolger, M. E., Arsova, B., and Usadel, B. (2018). Plant genome and transcriptome annotations: from misconceptions to simple solutions. *Briefings Bioinf.* 19, 437–449. doi: 10.1093/bib/bbw135

Brichet, N., Fournier, C., Turc, O., Strauss, O., Artzet, S., Pradal, C., et al. (2017). A robot-assisted imaging pipeline for tracking the growths of maize ear and silks in a high-throughput phenotyping platform. *Plant Methods* 13, 96. doi: 10.1186/s13007-017-0246-7

Chen, J., Beauvoit, B., Génard, M., Colombié, S., Moing, A., Vercambre, G., et al. (2021). Modelling predicts tomatoes can be bigger and sweeter if biophysical factors and transmembrane transporters are fine-tuned during fruit development. *New Phytol.* 230, 1489–1502. doi: 10.1111/nph.17260

Cooper, M., Voss-Fels, K. P., Messina, C. D., Tang, T., and Hammer, G. L. (2021). Tackling G × e × m interactions to close on-farm yield-gaps: creating novel pathways for crop improvement by predicting contributions of genetics and management to crop productivity. *Theor. Appl. Genet.* 134, 1625–1644. doi: 10.1007/s00122-021-03812-3

Crisp, P. A., Bhatnagar-Mathur, P., Hundleby, P., Godwin, I. D., Waterhouse, P. M., and Hickey, L. T. (2022). Beyond the gene: epigenetic and cis-regulatory targets offer new breeding potential for the future. *Curr. Opin. Biotechnol.* 73, 88–94. doi: 10.1016/j.copbio.2021.07.008

Dhont, S., Wuyts, N., and Inzé, D. (2013). Cell to whole-plant phenotyping: the best is yet to come. *Trends Plant Sci.* 18, 428–439. doi: 10.1016/j.tplants.2013.04.008

Dodig, D., Božinović, S., Nikolić, A., Zorić, M., Vančetović, J., Ignjatović-Micić, D., et al. (2021). Dynamics of maize vegetative growth and drought adaptability using image-based phenotyping under controlled conditions. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.652116

FAO (2022). *The future of food and agriculture – drivers and triggers for transformation* Vol. 3 (Rome: The Future of Food and Agriculture). doi: 10.4060/cc0959en

Fernie, A. R., Alseekh, S., Liu, J., and Yan, J. (2021). Using precision phenotyping to inform *de novo* domestication. *Plant Physiol.* 186, 1397–1411. doi: 10.1093/plphys/kiab160

Fernie, A. R., and Tohge, T. (2017). The genetics of plant metabolism. *Annu. Rev. Genet.* 51, 287–310. doi: 10.1146/annurev-genet-120116-024640

Fernie, A. R., and Yan, J. (2020). Targeting key genes to tailor old and new crops for a greener agriculture. *Mol. Plant* 13, 354–356. doi: 10.1016/j.molp.2020.02.007

Foix, S., Alenyà, G., and Torras, C. (2018). Task-driven active sensing framework applied to leaf probing. *Comput. Electron. Agric.* 147, 166–175. doi: 10.1016/j.compag.2018.01.020

Ghanem, M. E., Marrou, H., and Sinclair, T. R. (2015). Physiological phenotyping of plants for crop improvement. *Trend Plant Sci.* 20, 139–144. doi: 10.1016/j.tplants.2014.11.006

Girija, A., Han, J., Corke, F., Brook, J., Doonan, J., Yadav, R., et al. (2021). Elucidating drought responsive networks in tef (*Eragrostis tef*) using phenomic and metabolomic approaches. *Physiol. Plant.* 174, e13597. doi: 10.1111/ppl.13597

Großkinsky, D. K., Pieruschka, R., Svensgaard, J., Rascher, U., Christensen, S., Schurr, U., et al. (2015a). Phenotyping in the fields: dissecting the genetics of quantitative traits and digital farming. *New Phytol.* 207, 950–952. doi: 10.1111/nph.13529

Großkinsky, D. K., Svensgaard, J., Christensen, S., and Roitsch, T. (2015b). Plant phenomics and the need for physiological phenotyping across scales to narrow the genotype-to-phenotype knowledge gap. *J. Exp. Bot.* 66, 5429–5440. doi: 10.1093/jxb/erv345

Han, L., Silvestre, S., Sayanova, O., Haslam, R. P., and Napier, J. A. (2022). Using field evaluation and systematic iteration to rationalize the accumulation of omega-3 long-chain polyunsaturated fatty acids in transgenic *Camelina sativa*. *Plant Biotechnol. J.* 20, 1833–1852. doi: 10.1111/pbi.13867

Hawkins, C., Ginzburg, D., Zhao, K., Dwyer, W., Xue, B., Xu, A., et al. (2021). Plant metabolic network 15: a resource of genome-wide metabolism databases for 126 plants and algae. *J. Integr. Plant Biol.* 63, 1888–1905. doi: 10.1111/jipb.13163

Hickey, L. T., Hafeez, A. N., Robinson, H., Jackson, S. A., Leal-Bertioli, S. C. M., Tester, M., et al. (2019). Breeding crops to feed 10 billion. *Nat. Biotechnol.* 37, 744–754. doi: 10.1038/s41587-019-0152-9

Hohmann, M., Stahl, A., Rudloff, J., Wittkop, B., and Snowdon, R. J. (2016). Not a load of rubbish: simulated field trials in large-scale containers. *Plant Cell Environ.* 39, 2064–2073. doi: 10.1111/pce.12737

Holst, F., Bolger, A., Günther, C., Maß, J., Triesch, S., Kindel, F., et al. (2023). Helixer–*de novo* prediction of primary eukaryotic gene models combining deep learning and a hidden Markov model. *bioRxiv*. doi: 10.1101/2023.02.06.527280

Jacob, D., David, R., Aubin, S., and Gibon, Y. (2020). Making experimental data tables in the life sciences more FAIR: a pragmatic approach. *GigaScience* 9, giaa144. doi: 10.1093/gigascience/giaa144

Jiang, W. Z., Henry, I. M., Lynagh, P. G., Comai, L., Cahoon, E. B., and Weeks, D. P. (2017). Significant enhancement of fatty acid composition in seeds of the allohexaploid, *Camelina sativa*, using CRISPR/Cas9 gene editing. *Plant Biotechnol. J.* 15, 648–657. doi: 10.1111/pbi.12663

Kagale, S., Koh, C., Nixon, J., Bolina, V., Clarke, W. E., Tuteja, R., et al. (2014). The emerging biofuel crop *Camelina sativa* retains a highly undifferentiated hexaploid genome structure. *Nat. Commun.* 5, 3706. doi: 10.1038/ncomms4706

Kajala, K., Gouran, M., Shaar-Moshe, L., Mason, G. A., Rodriguez-Medina, J., Kawa, D., et al. (2021). Innovation, conservation, and repurposing of gene function in root cell type development. *Cell* 184, 3333–33348. doi: 10.1016/j.cell.2021.04.024

Keurentjes, J. J. B., Angenent, G. C., Dicke, M., Martins Dos Santos, V. A. P., Molenaar, J., van der Putten, W. H., et al. (2011). Redefining plant systems biology: from cell to ecosystem. *Trends Plant Sci.* 16, 183–190. doi: 10.1016/j.tplants.2010.12.002

King, K., Li, H., Kang, J., and Lu, C. (2019). Mapping quantitative trait loci for seed traits in *Camelina sativa*. *Theor. Appl. Genet.* 132, 2567–2577. doi: 10.1007/s00122-019-03371-8

Kumar, A., Anju, T., Kumar, S., Chhapekar, S. S., Sreedharan, S., Singh, A., et al. (2021). Integrating omics and gene editing tools for rapid improvement of traditional food plants for diversified and sustainable food security. *Int. J. Mol. Sci.* 22, 8093. doi: 10.3390/ijms22158093

Kumar, B., Singh, A. K., Bahuguna, R. N., Pareek, A., and Singla-Pareek, S. L. (2023). Orphan crops: a genetic treasure trove for hunting stress tolerance genes. *Food Energy Secur.* 12, e436. doi: 10.1002/fes3.436

Lemmon, Z. H., Reem, N. T., Dalrymple, J., Soyk, S., Swartwood, K. E., Rodriguez-Leal, D., et al. (2018). Rapid improvement of domestication traits in an orphan crop by genome editing. *Nat. Plants* 4, 766–770. doi: 10.1038/s41477-018-0259-x

Li, H., Hu, X., Lovell, J. T., Grabowski, P. P., Mamidi, S., Chen, C., et al. (2021). Genetic dissection of natural variation in oilseed traits of camelina by whole-genome resequencing and QTL mapping. *Plant Genome* 14, e20110. doi: 10.1002/tpg2.20110

López-Ruiz, R., Romero-González, R., and Garrido Frenich, A. (2019). Ultrahigh-pressure liquid chromatography-mass spectrometry: an overview of the last decade. *TrAC Trends Analytical Chem.* 118, 170–181. doi: 10.1016/j.trac.2019.05.044

Madec, S., Baret, F., de Solan, B., Thomas, S., Dutartre, D., Jezequel, S., et al. (2017). High-throughput phenotyping of plant height: comparing unmanned aerial vehicles and ground LiDAR estimates. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.02002

Marks, M. D., Chopra, R., and Sedbrook, J. C. (2021). Technologies enabling rapid crop improvements for sustainable agriculture: example pennycress (*Thlaspi arvense* l.). *Emerging Topics Life Sci.* 5, 325–335. doi: 10.1042/ETLS20200330

Mascher, M., Wicker, T., Jenkins, J., Plott, C., Lux, T., Koh, C. S., et al. (2021). Long-read sequence assembly: a technical evaluation in barley. *Plant Cell* 33, 1888–1906. doi: 10.1093/plcell/koab077

Metje-Sprink, J., Sprink, T., and Hartung, F. (2020). Genome-edited plants in the field. *Curr. Opin. Biotechnol.* 61, 1–6. doi: 10.1016/j.copbio.2019.08.007

Milla, R., and Osborne, C. P. (2021). Crop origins explain variation in global agricultural relevance. *Nat. Plants* 7, 598–607. doi: 10.1038/s41477-021-00905-1

Millet, E. J., Kruijer, W., Couple-Ledru, A., Alvarez Prado, S., Cabrera-Bosquet, L., Lacube, S., et al. (2019). Genomic prediction of maize yield across European environmental conditions. *Nat. Genet.* 51, 952–956. doi: 10.1038/s41588-019-0414-y

Morineau, C., Bellec, Y., Tellier, F., Gissot, L., Kelemen, Z., Nogué, F., et al. (2017). Selective gene dosage by CRISPR-Cas9 genome editing in hexaploid *Camelina sativa*. *Plant Biotechnol. J.* 15, 729–739. doi: 10.1111/pbi.12671

Padulosi, S. (2017). Bring NUS back to the table! *GREAT Insights* 6, 21–22.

Papoutsoglou, E. A., Faria, D., Arend, D., Arnaud, E., Athanasiadis, I. N., Chaves, I., et al. (2020). Enabling reusability of plant phenomic datasets with MIAPPE 1.1. *New Phytol.* 227, 260–273. doi: 10.1111/nph.16544

Pommier, C., Michotey, C., Cornut, G., Roumet P Duchêne, E., Flores, R., Lebreton, A., et al. (2019). Applying FAIR principles to plant phenotypic data management in GnpIS. *Plant Phenomics* 2019, 1671403. doi: 10.34133/2019/1671403

R2D2 Consortium, Fugeray-Scarbel, A., Bastien, C., Dupont-Nivet, M., Lemarié, S. (2021). Why and how to switch to genomic selection: lessons from plant and animal breeding experience. *Front. Genet.* 12. doi: 10.3389/fgene.2021.629737

Reynolds, D., Baret, F., Welcker, C., Bostrom, A., Ball, J., Cellini, F., et al. (2019). What is cost-efficient phenotyping? Optimizing costs for different scenarios. *Plant Sci.* 282, 14–22. doi: 10.1016/j.plantsci.2018.06.015

Rönspies, M., Dorn, A., Schindele, P., and Puchta, H. (2021). CRISPR-cas-mediated chromosome engineering for crop improvement and synthetic biology. *Nat. Plants* 7, 566–573. doi: 10.1038/s41477-021-00910-4

Schwacke, R., Ponce-Soto, G. Y., Krause, K., Bolger, A. M., Arsova, B., Hallab, A., et al. (2019). MapMan4: a redefined protein classification and annotation framework applicable to multi-omics data analysis. *Mol. Plant* 12, 879–892. doi: 10.1016/j.molp.2019.01.003

Shi, J., Gao, H., Wang, H., Lafitte, H. R., Archibald, R. L., Yang, M., et al. (2017). ARGOS8 variants generated by CRISPR-Cas9 improve maize grain yield under field drought stress conditions. *Plant Biotechnol. J.* 15, 207–216. doi: 10.1111/pbi.12603

Stasnik, P., Großkinsky, D. K., and Jonak, C. (2022). Physiological and phenotypic characterization of diverse *Camelina sativa* lines in response to waterlogging. *Plant Physiol. Biochem.* 183, 120–127. doi: 10.1016/j.plaphy.2022.05.007

Stevens, R. G., Baldet, P., Bouchet, J.-P., Causse, M., Deborde, C., Deschodt, C., et al. (2018). A systems biology study in tomato fruit reveals correlations between the ascorbate pool and genes involved in ribosome biogenesis, translation, and the heat-shock response. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.00137

Vandepoele, K. (2017). "A guide to the PLAZA 3.0 plant comparative genomic database," in *Plant genomics databases. methods in molecular biology*, vol. 1533 . Ed. A. van Dijk (New York, NY: Humana Press). doi: 10.1007/978-1-4939-6658-5_10

van Rengs, W. M. J., Schmidt, M. H.-W., Effgen, S., Wang, Y., Zaidan, M. W. A. M., Huettel, B., et al. (2021). A gap-free tomato genome built from complementary PacBio and nanopore long DNA sequences reveals extensive linkage drag during breeding. *BioRxiv*. doi: 10.1101/2021.08.30.456472

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

Vello, E., Aguirre, J., Shao, Y., and Bureau, T. (2022). "Camelina sativa high-throughput phenotyping under normal and salt conditions using a plant phenomics platform," in *High-throughput plant phenotyping: methods and protocols. methods in molecular biology*. Eds. A. Lorence and K. Medina-Jimenez (New York, NY: Humana Press), 2539. doi: 10.1007/978-1-0716-2537-8_4

Watt, M., Fiorani, F., Usadel, B., Rascher, U., Muller, O., and Schurr, U. (2020). Phenotyping: new windows into the plant for breeders. *Annular Rev. Plant Biol.* 71, 689–712. doi: 10.1146/annurev-arplant-042916-041124

Westhues, M., Schrag, T. A., Heuer, C., Thaller, G., Utz, H. F., Schipprack, W., et al. (2017). Omics-based hybrid prediction in maize. *Theor. Appl. Genet.* 130, 1927–1939. doi: 10.1007/s00122-017-2934-0

Wong, C. Y., Gilbert, M. E., Pierce, M. A., Parker, T. A., Palkovic, A., Gepts, P., et al. (2023). Hyperspectral remote sensing for phenotyping the physiological drought response of common and tepary bean. *Plant Phenomics* 5, 21. doi: 10.34133/plantphenomics.0021

Yang, W., Feng, H., Zhang, X., Zhang, J., Doonan, J. H., Batchelor, W. D., et al. (2020). Crop phenomics and high-throughput phenotyping: past decades, current challenges, and future perspectives. *Mol. Plant* 13, 187–214. doi: 10.1016/j.molp.2020.01.008

Yu, H., Lin, T., Meng, X., Du, H., Zhang, J., Liu, G., et al. (2021). A route to *de novo* domestication of wild allotetraploid rice. *Cell* 184, 1156–1170. doi: 10.1016/j.cell.2021.01.013

Zanetti, F., Alberghini, B., Marjanović Jeromela, A., Grahovac, N., Rajković, D., Kiprovski, B., et al. (2021). Camelina, an ancient oilseed crop actively contributing to the rural renaissance in europe. a review. *Agron. Sustain. Dev.* 41, 2. doi: 10.1007/s13593-020-00663-y

Zhang, A., Liu, Y., Wang, F., Li, T., Chen, Z., Kong, D., et al. (2019). Enhanced rice salinity tolerance *via* CRISPR/Cas9-targeted mutagenesis of the OsRR22 gene. *Mol. Breed.* 39, 47. doi: 10.1007/s11032-019-0954-y

Zhao, C., Liu, B., Piao, S., Wang, X., Lobell, D. B., Huang, Y., et al. (2017). Temperature increase reduces global yields of major crops in four independent estimates. *Proc. Natl. Acad. Sci. U. S. A.* 114, 9326–9331. doi: 10.1073/pnas.1701762114

Zhao, C., Zhang, Y., Du, J., Guo, X., Wen, W., Gu, S., et al. (2019). Crop phenomics: current status and perspectives. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.00714

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 20 June 2023
DOI 10.3389/fpls.2023.1136643

OPEN ACCESS

EDITED BY
Meng Kou,
Xuzhou Institute of Agricultural Sciences in Jiangsu Xuhuai District, China

REVIEWED BY
Chen Junfeng,
Shanghai University of Traditional Chinese Medicine, China
Albert Ferrer,
University of Barcelona, Spain

*CORRESPONDENCE
Heqiang Lou
20170030@zafu.edu.cn
Jiasheng Wu
wujs@zafu.edu.cn

†PRESENT ADDRESS
Feicui Zhang
Institute of Maize and Featured Dryland Crops, Zhejiang Academy of Agricultural Sciences, Dongyang, China

‡These authors have contributed equally to this work

RECEIVED 03 January 2023
ACCEPTED 04 May 2023
PUBLISHED 20 June 2023

CITATION
Zhang F, Kong C, Ma Z, Chen W, Li Y, Lou H and Wu J (2023) Molecular characterization and transcriptional regulation analysis of the *Torreya grandis* squalene synthase gene involved in sitosterol biosynthesis and drought response.
Front. Plant Sci. 14:1136643.
doi: 10.3389/fpls.2023.1136643

Molecular characterization and transcriptional regulation analysis of the *Torreya grandis* squalene synthase gene involved in sitosterol biosynthesis and drought response

Feicui Zhang†‡, Congcong Kong†‡, Zhenmin Ma, Wenchao Chen, Yue Li, Heqiang Lou* and Jiasheng Wu*

State Key Laboratory of Subtropical Silviculture, Zhejiang A&F University, Hangzhou, China

The kernel of *Torreya grandis* cv. 'Merrillii' (Cephalotaxaceae) is a rare nut with a variety of bioactive compounds and a high economic value. β-sitosterol is not only the most abundant plant sterol but also has various biological effects, such as antimicrobial, anticancer, anti-inflammatory, lipid-lowering, antioxidant, and antidiabetic activities. In this study, a squalene synthase gene from *T. grandis*, *TgSQS*, was identified and functionally characterized. *TgSQS* encodes a deduced protein of 410 amino acids. Prokaryotic expression of the TgSQS protein could catalyze farnesyl diphosphate to produce squalene. Transgenic *Arabidopsis* plants overexpressing *TgSQS* showed a significant increase in the content of both squalene and β-sitosterol; moreover, their drought tolerance was also stronger than that of the wild type. Transcriptome data from *T. grandis* seedlings showed that the expression levels of sterol biosynthesis pathway-related genes, such as *HMGS*, *HMGR*, *MK*, *DXS*, *IPPI*, *FPPS*, *SQS*, and *DWF1*, increased significantly after drought treatment. We also demonstrated that TgWRKY3 directly bound to the *TgSQS* promoter region and regulated its expression through a yeast one-hybrid experiment and a dual luciferase experiment. Together, these findings demonstrate that *TgSQS* has a positive role in β-sitosterol biosynthesis and in protecting against drought stress, emphasizing its importance as a metabolic engineering tool for the simultaneous improvement of β-sitosterol biosynthesis and drought tolerance.

KEYWORDS

***Torreya grandis*, transcriptional regulation, squalene synthase, sitosterol biosynthesis, drought**

Introduction

Sterols are isoprenoid-derived molecules found in bacteria, fungi, insects, mammals, and plants (Yuan et al, 2010). In plants, more than 250 kinds of sterols and sterol conjugates have been identified, including free sterols, sterol esters, sterol glucosides, and acylated sterol glucosides. The most abundant plant sterols are sitosterol, followed by stigmasterol and campesterol (Goldstein and Brown, 1990; Schluttenhofer and Yuan, 2015). Sitosterol is not only a component of the plant cell membrane, which plays an important role in seed germination and organ development; it can also participate in the plant's response to low temperatures, wounds, salt, and drought stress (Elkeilsh et al, 2019). To initiate plant sterol biosynthesis, isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) are generated from either the cytosolic mevalonate (MVA) pathway or the plastidial methylerythritol phosphate (MEP) pathway (Guan et al, 1998; Jiang et al, 2017). Then, the "head-to-tail" condensation of IPP and DMAPP forms C15 farnesyl diphosphate (FPP). Subsequently, two molecules of FPP are condensed head-to-head by squalene synthase (SQS) to form squalene (Nguyen et al, 2013). Finally, sterols, such as sitosterol, stigmasterol, and campesterol, are produced through a series of reactions (Singh et al., 2017). Because FPP is also the precursor of other non-sterol isoprenoids, such as ubiquinones and sesquiterpenoids, regulation of SQS has been considered important for sterol biosynthesis (Nie et al, 2019).

Some studies have shown that *SQS* expression significantly increases in roots, stems, and leaves when *Panax ginseng* and *Portulaca oleracea* are treated with methyl jasmonate (Kumar et al., 2015). In addition, several transcription factors have been reported to regulate *SQS* expression. For example, two bHLH family transcription factors, SREBP1a and SREBP2 (sterol regulatory element binding proteins), are involved in cholesterol synthesis by regulating the expression of *SQS* in human cells (He et al., 2016). TPO1 negatively regulates *Saccharomyces cerevisiae SQS* (*ERG9*) expression, while YER064C and SLK19 positively regulate its expression (Tai et al., 2023; Kennedy and Bard, 2001). WRKY1 increases the content of triterpenes, such as withaferin, by positively regulating *SQS* expression in *Withania somnifera* (Sun et al., 2012).

Torreya grandis cv. 'Merrillii' is an excellent variety, with a cultivation history of more than one thousand years. The kernel of *T. grandis* cv. 'Merrillii' is a precious nut that is rich in unsaturated fatty acids and β-sitosterol, with insecticidal, anti-inflammatory, and antioxidant effects (Huang et al., 2007; Divi and Krishna, 2009). Several studies have indicated that β-sitosterol can protect against breast cancer (Awad et al., 2007; Ju et al., 2004), prostate cancer (Awad et al., 2001), colon cancer (Awad et al., 1997; Clough and Bent, 1998), and gastric cancer (Zhao et al., 2017) by inhibiting cell proliferation and inducing apoptosis. In addition, β-sitosterol can inhibit the increase in blood lipids caused by a high-fat diet (Singh et al., 2015; Salehi et al., 2020). With the given benefits, improving the β-sitosterol content in *T. grandis* can increase the kernel's nutritional values. In this study, we cloned the *SQS* of *T. grandis* for the first time and proved that it encodes an active squalene synthase. The β-sitosterol content increased significantly in mature seeds of *TgSQS*-overexpressing homozygous lines of *Arabidopsis thaliana*; moreover, their drought tolerance was also stronger than that of the wild type. We also demonstrated that TgWRKY3 directly bound to the *TgSQS* promoter region and regulated its expression through a yeast one-hybrid experiment and a dual luciferase experiment.

Materials and methods

Plant materials, growth conditions, and drought treatments

For transcriptome analyses, one-year-old *T. grandis* seedlings were transplanted to pots with 300 g of a soil mixture and grown in a shady plastic greenhouse. After two weeks, the surviving seedlings were irrigated for 12 h, and then irrigation was withdrawn to start drought treatment.

For drought resistance tests, plants were grown in a potting soil mixture in growth chambers at 22°C with 16 h/day illumination. The relative humidity was approximately 70% (± 5%). After growing for one week, the plants were irrigated with 1 L of water per tray for 6 h, and then drought treatment was imposed by withdrawing irrigation for half of the plants until most died. The other half were grown under a standard irrigation regime as a control. After 20 days of drought treatment, representative pots were placed to take photos.

For DAB and NBT staining, *Arabidopsis* seeds were surface sterilized and cold treated at 4°C for three days. Then, the seeds were plated on an MS medium containing 3% (w/v) sucrose and 0.8% (w/v) agar and grown at 22°C with a 16-h daily light period. The 10-day-old seedlings were transferred to the MS (Murashige and Skoog) medium with 10% PEG (with controls) for three days before staining. For DAB staining, the samples were immersed in a DAB solution (SL1805, Coolaber) for 4 h and then in 95% ethanol for decoloring. For NBT staining, the samples were first immersed in NBT staining solution (S19048, Yuanye) until a dark blue color appeared (approximately 1 h) and then in 95% ethanol for decoloring. Photographs were taken using a stereo microscope (SZX16, Olympus).

Phylogenetic analysis

Phylogenetic analysis generated a bootstrap neighbor-joining evolutionary tree using MEGA 7.0 with 1,000 bootstrap replicates. Amino acid sequence alignment was conducted using DNAMAN software.

Gene cloning, construction, and transformation in A. thaliana

To obtain *TgSQS* overexpression in transgenic plants, the coding sequence of *TgSQS* was PCR-amplified with the primers

listed in Table S1 and introduced into the Super1300 vector driven by a Super promoter (multiple CaMV35S are connected in series). The construct was transformed into *Agrobacterium* strain GV3101 for transfection of *Arabidopsis* using the floral-dip method (Devarenne et al, 2002). Seeds were screened on an MS agar medium containing 50 mg/L of hygromycin. The selected T3-generation transgenic plants with 100% resistance to hygromycin were considered homozygous lines and harvested for further analysis.

Subcellular localization

The coding sequence of TgSQS was cloned into a modified pCAMBIA1300, which had GFP fused to the N-terminal, and introduced into *Agrobacterium tumefaciens* strain GV3101. The transient gene expression analysis of TgSQS and an endoplasmic reticulum marker in *Nicotiana benthamiana* was performed according to a previous study (Nakashima et al., 1995). Agrobacteria were grown overnight in LB (Luria-Bertani) media and brought to an OD600 of 0.8 in the injection solution. After three days of injection, green fluorescent protein (GFP) and red fluorescent protein (RFP) fluorescence was observed *via* confocal laser scanning microscopy (LSM510, Carl Zeiss).

Prokaryotic protein expression and enzyme activity analysis

The coding sequence of *TgSQS* was introduced into the pET-32a vector (provided by Zuying Zhang) and transformed into competent Rosetta (DE3) to produce a recombinant protein containing a Trx-His-tag at the N-terminus. The cloning sites were BamHI and SacI. The transformed bacteria were grown to an OD600 of 0.6 at 37°C in an LB medium, and then 0.4 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to induce TgSQS expression. After culturing for 20 h at 16°C, the cells were collected and lysed by sonication. The crude enzyme solution was subjected to Trx-His-Tag purification using a Ni Sepharose 6 Fast Flow gravity column to obtain purified protein. Imidazole was removed from the purified protein using centrifugal ultrafiltration (Amicon Ultra 30 kDa, Millipore, Burlington, MA, USA). After BCA protein quantification, 10-μl samples were used to run a 10% SDS-PAGE gel.

The enzyme reaction of the recombinant protein TgSQS was conducted as follows: 50 μg of purified protein was incubated at 30°C with 20 μg of FPP triammonium salt (Sigma–Aldrich, St. Louis, MO, USA), 50 mM Tris–HCl (pH7.5), 25 mM $MgCl_2$, 1 mM DTT, 2% glycine, and 3 mM NADPH in a total reaction volume of 500 μl. After 8 h, the reaction product was extracted three times with 500 μl of n-hexane, concentrated to 100 μl, and analyzed using GC–MS. GC–MS detection was performed using Trace GC Ultra-ISQ (Thermo Scientific) with a DB-5MS column (30 m × 0.25 mm × 0.25 μm; Agilent). The GC temperature program was as follows: 80°C, raised to 300°C at a rate of 15°C/min, and held at this temperature for 18 min. The carrier gas was ultrahigh-purity helium at a flow rate of 1.0 ml/min. The GC interface temperature was 290°C, and the sample injection volume used was 1 μl.

Squalene analysis

Dry *A. thaliana* seeds were accurately weighed to 0.039 g and ground with liquid nitrogen. Ultrasonically extracted samples were extracted for 20 min three times, and the solvent was removed under low temperature and pressure and dissolved in 1 ml of n-hexane. A total of 1.5 ml of a 2 mol/L KOH-ethanol solution was added to 0.4 ml of the extraction solution. Samples were saponified ultrasonically for 10 min in a 60°C water bath and placed in a 60°C oven for 60 min after vortexing. They were shaken for 1 min, and 1 ml of water and 1 ml of n-hexane were added after cooling. Samples were extracted for 2 min, and then the supernatants were removed. We added 0.25 g of anhydrous sodium sulfate to purify the supernatants, dried the supernatants with liquid nitrogen, and then dissolved the precipitates in 0.1 ml of n-hexane to start GC–MS analysis. The GC–MS detection was performed using a Trace GC Ultra-ISQ (Thermo Scientific) with a DB-5MS column (30 m × 0.25 mm × 0.25 μm: Agilent). The GC temperature program was as follows: 80°C, raised to 300°C at a rate of 15°C/min, and held at this temperature for 18 min. The carrier gas was ultrahigh-purity helium at a flow rate of 1.0 ml/min. The GC interface temperature was 290° C, and the sample injection volume used was 1 μl.

RNA extraction and transcriptome analysis

Total RNA was isolated from the leaves of *T. grandis* seedlings treated with drought stress using the RNAprep Pure Plant Kit (DP441, Tiangen); mRNA was purified from total RNA using poly-T oligo-attached magnetic beads. First-strand cDNA was synthesized using a random hexamer primer and M-MuLV reverse transcriptase (RNase H-). The library fragments were purified with the AMPure XP system (Beckman Coulter, Beverly, USA) to select cDNA fragments that were preferentially 370–420 bp in length. Library quality was assessed using a Qubit2.0 Fluorometer, Agilent Bioanalyzer 2100 system, and qRT-PCR. The clustering of the index-coded samples was performed on a cBot Cluster Generation System using the TruSeq PE Cluster Kit v3-cBot-HS (Illumia) according to the manufacturer's instructions. After cluster generation, the library preparations were sequenced on an Illumina Novaseq platform, and 150-bp paired-end reads were generated.

Raw reads of the fastq format were first processed using in-house perl scripts. At the same time, the Q20, Q30, and GC content of the clean data were calculated. Differential expression analysis of two conditions/groups (two biological replicates per condition) was performed using the DESeq2 R package (1.20.0). The obtained P-values were adjusted using the Benjamini–Hochberg method, which is designed to control the false discovery rate. Genes with an adjusted P-value <0.05 found by DESeq2 were assigned as differentially expressed. Gene Ontology (GO) enrichment analysis

and KEGG pathways of differentially expressed genes were implemented using the cluster Profiler R package, in which gene length bias was corrected.

Quantitative real-time PCR (qRT-PCR)

qRT-PCR was performed with a C1000 Touch™ Thermal Cycler system (Bio-Rad) and the ChamQ SYBR qPCR Master Mix kit (Vazyme). The relative expression level was calculated according to the $2^{-\triangle\triangle Ct}$ method. The actin gene was used as a reference gene. Ct represents the PCR cycle number at which the amount of target reaches a fixed threshold. The corresponding primers are listed in Table S2.

Sterol extraction and analysis

We accurately weighed 0.039 g of dry *A. thaliana* seeds and 0.5 g of fresh leaves (from *T. grandis* seedlings treated with drought stress) and ground them in liquid nitrogen. Samples were extracted with ultrasound for 20 min three times, and the solvent was removed under low temperature and pressure and dissolved in 1 ml of n-hexane. A volume of 1.5 ml of a 2 mol/L KOH-ethanol solution was added to a 0.4 ml extraction solution. Samples were ultrasonically saponified for 10 min in a 60°C water bath and placed in a 60°C oven for 60 min after vortexing. They were shaken for 1 min, and 1 ml of water and 1 ml of n-hexane were added after cooling. Samples were extracted for 2 min, and the supernatants were removed. We added 0.25 g of anhydrous sodium sulfate to purify the supernatants, dried the supernatants with liquid nitrogen, and dissolved the precipitates with 0.1 ml of n-hexane for silylation to form trimethylsilyl ether derivatives of sterols, which were used for GC–MS.

Samples (1 μl) were injected into a GCMS-QP201PLUS apparatus (Shimadzu), which consisted of an automated sampler injection system, a split/splitless injector, and a DB-5MS column (30 m × 0.25 mm × 0.25 μm: Agilent). The GC temperature program was as follows: 80°C, raised to 290°C at a rate of 20°C/min, and held at this temperature for 18 min. The carrier gas was ultrahigh-purity helium at a flow rate of 1.0 ml/min and a 1:10 split injection ratio. Authentic campesterol, stigmasterol, β-sitosterol, and cycloartenol were purchased from Sigma-Aldrich. For quantification, a standard curve for individual authentic standards was generated.

Vector construction and dual luciferase assay

The coding sequence of *TgWRKYs* was obtained from RNA sequencing and cloned into the pCAMBIA 1300-GFP vector to express the effector. The 1870-bp promoter fragments of *TgSQS* were sub-cloned into the pGreenII 0800-LUC vector as a reporter. The corresponding primers are listed in Table S1. Individual effector vectors and recombinant reporter vectors were transferred to *A. tumefaciens* GV3101. The reporter was mixed with each kind of effector or the empty pCAMBIA 1300-GFP vector (control) at a 1:1 (v:v) ratio and then injected into tobacco leaves as described previously (Nakashima et al., 1995). A dual-luciferase assay was carried out using the Dual-Lumi™ Luciferase Reporter Gene Assay Kit (Beyotime, RG089). The firefly luciferase (FLUC) and Renilla luciferase (RLUC) values were measured using a Promega GloMax20/20 Luminescence Detector. The relative LUC activity was calculated as the ratio between FLUC and RLUC activities, and three biological replicates were applied.

Yeast one-hybrid assay

Yeast one-hybrid (Y1H) assays were carried out with the Matchmaker Gold Yeast One-Hybrid System Kit (Clontech) according to the manufacturer's protocol (PT4087-1, Clontech, USA). The corresponding primers are listed in Table S1. Based on the distribution of predicted W-box binding sites in the *TgSQS* promoter region, three short fragments of the *TgSQS* promoter (the adenine residue of the translational start codon ATG) were assigned position +1, and the numbers flanking the sequences of the *TgSQS* promoter fragments were counted based on this number. The positions of three 200-bp sequences containing a W-box were as follows: W1, 1775 to 1575 bp; W2, 1275 to 1075 bp; and W3, 375–175 bp). They were subcloned into the pAbAi vector to obtain *pAbAi-W1/W2/W3*. Figure S2 provides the sequence of the *TgSQS* promoter. The coding sequence of *TgWRKY3/6/7* was inserted into the pGADT7 vector to construct the prey-AD vector. Then, the linearized *pAbAi-W1/W2/W3* vector was transformed into Y1HGold. After determining the minimal inhibitory concentration of Aureobasidin A (AbA) for positive transformants, the prey-AD vector was transformed into the bait yeast strain. Successfully transformed yeast strains were grown on the corresponding SD medium (SD/−Leu, SD/−Leu + AbA) for 3–5 days to take photos.

Results

Identification of TgSQS

Since the genome of *T. grandis* has not yet been sequenced, we searched the transcriptome data (NCBI accession Nos. SRX9417001–SRX9417009) and found a unigene annotated as squalene synthase. This sequence was used as a query in the blastx of NCBI, and the results showed that this unigene had high identity (>80%) with squalene synthase from *Taxus cuspidate*, *Ginkgo biloba* L., and *Pinus massoniana*, indicating that this unigene likely encodes squalene synthase in *T. grandis*; thus, it was named *TgSQS*. The 1230-bp coding length of *TgSQS* was determined using sequence alignment in blastx. Phylogenetic tree analysis of TgSQS with SQSs from angiosperms, gymnosperms, algae, and animals showed that TgSQS was most closely related to gymnosperms, such as *T. cuspidate*, and furthest from animals (Figure 1A). Multiple alignments of TgSQS with squalene synthase

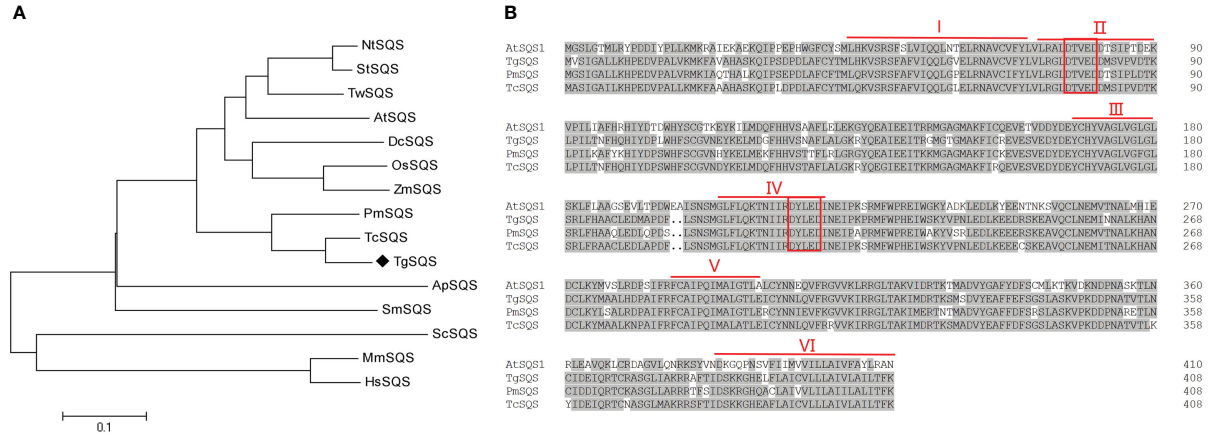


FIGURE 1
Analysis of the TgSQS amino acid sequence. **(A)** Construction of the phylogenetic tree of squalene synthase from different species using the neighbor-joining method. The accession numbers of the genes used were: NtSQS, *Nicotiana tabacum*, AAB08578.1; StSQS, *Solanum tuberosum*, BAA82093.1; AtSQS1, *Arabidopsis thaliana*, AEE86403.1; OsSQS, *Oryzasativa Japonica* Group, BAA22557.1; DcSQS, *Dendrobium catenatum*, AGI56082.1; ZmSQS, *Zea mays*, BAA22558.1; PmSQS, *Pinus massoniana*, AHI96421.1; TcSQS, *Taxus cuspidate*, ABI14439.1; ApSQS, *Auxenochlorella protothecoides*, KFM22694.1; HsSQS, *Homo sapiens*, AAB33404.1; ScSQS, *Saccharomyces cerevisiae*, AAA34597.1; TwSQS, *Tripterygium wilfordii*, AMR60779.1; SmSQS, *Selaginella moellendorffii*, XP024532546.1; MmSQS, *Mus musculus*, AP_034321.2. **(B)** Alignment of the full-length amino acid sequences of TgSQS and its homologs in *Arabidopsis thaliana*, *Taxus cuspidate*, and *Pinus massoniana*. The six conserved regions (I, II, III, IV, V, and VI) of squalene synthase are marked by red lines, and two aspartate rich domains (DXXXD) are marked by red squares.

from *A. thaliana*, *T. cuspidate*, and *P. massoniana* showed that there were six highly conserved regions, I–VI, in its amino acid sequence, among which II and IV contained aspartic acid (DXXXD) active sites (Figure 1B), which were reported to bind to isopentene phosphate groups (Rushton et al., 2010), and VI was reported to perform the function of membrane targeting and anchoring (Ishihama et al., 2011). The DeepTMHMM-based predictions of transmembrane domains revealed that TgSQS had one C-terminal transmembrane domain (amino acid residues 389–407), which was highly like those predicted for *A. thaliana* (Figure S1B, C) and *Taraxacum koksaghyz* SQS (Valitova et al, 2016). We further investigated the subcellular localization of TgSQS through transient expression in *N. benthamiana* leaves. As can be seen in Figure S1A, the TgSQS protein fused with GFP at the N-terminal colocalized with the endoplasmic reticulum marker, which was consistent with the literature (Valitova et al., 2016). These results suggest that *TgSQS* encodes functional squalene synthase.

Functional characterization of TgSQS *in vitro* and *in vivo*

To understand the catalytic function of TgSQS, the recombinant expression plasmid *pET-32a-TgSQS* was constructed and transformed into *E. coli* Rosetta (DE3) to produce a recombinant protein containing a Trx-His-tag at the N-terminus. The *pET-32a* empty vector served as a control. After induced expression, extraction, and purification, the purified recombinant proteins were analyzed using SDS-PAGE. TgSQS was successfully extracted from the supernatant of lysates with a molecular weight of around 70 kDa (Figure 2A). The enzyme activity of TgSQS was assayed using the purified protein with the cofactors NADPH and Mg^{2+}, while Trx-His-tag was used as a control. The catalytic products were analyzed using GC–MS; the peak at 13.02 min in the profile of TgSQS reaction products (Figure 2B) corresponded to that observed in the standard squalene sample (Figure 2C), but no such peak was detected in the control (Figure 2D). Moreover, the characteristics of peaks in the mass spectrum of TgSQS (Figure 2E) were consistent with those of standard squalene (Figure 2F). These results proved that TgSQS has the catalytic activity of squalene synthase *in vitro*.

To further verify the function of *TgSQS*, the overexpression plasmid *super1300-TgSQS* was transferred into *A. thaliana*, and three constitutive overexpression lines with significantly higher expression levels of *TgSQS* were obtained (Figure 3A). The content of squalene in mature seeds was detected, and the results showed that the content of squalene in the three overexpression lines was significantly higher than that of the wild type (Figure 3B), indicating that TgSQS also has the catalytic activity of squalene synthase *in vivo*.

Overexpression of TgSQS increased β-sitosterol content in Arabidopsis seeds

In plants, the dominant sterols comprise β-sitosterol, campesterol, and stigmasterol, while β-sitosterol is the most abundant. It has been evidenced in many *in vitro* and *in vivo* studies that β-sitosterol possesses various biological actions, such as immunomodulatory, antimicrobial, anticancer, anti-inflammatory, lipid-lowering, antioxidant, and anti-diabetic activities (Babu and Jayaraman, 2020). Therefore, increasing the β-sitosterol content in *T. grandis* seeds is very important to improve their nutritional value. As it is difficult to achieve genetic transformation in *T. grandis*, to

10.3389/fpls.2023.1136643

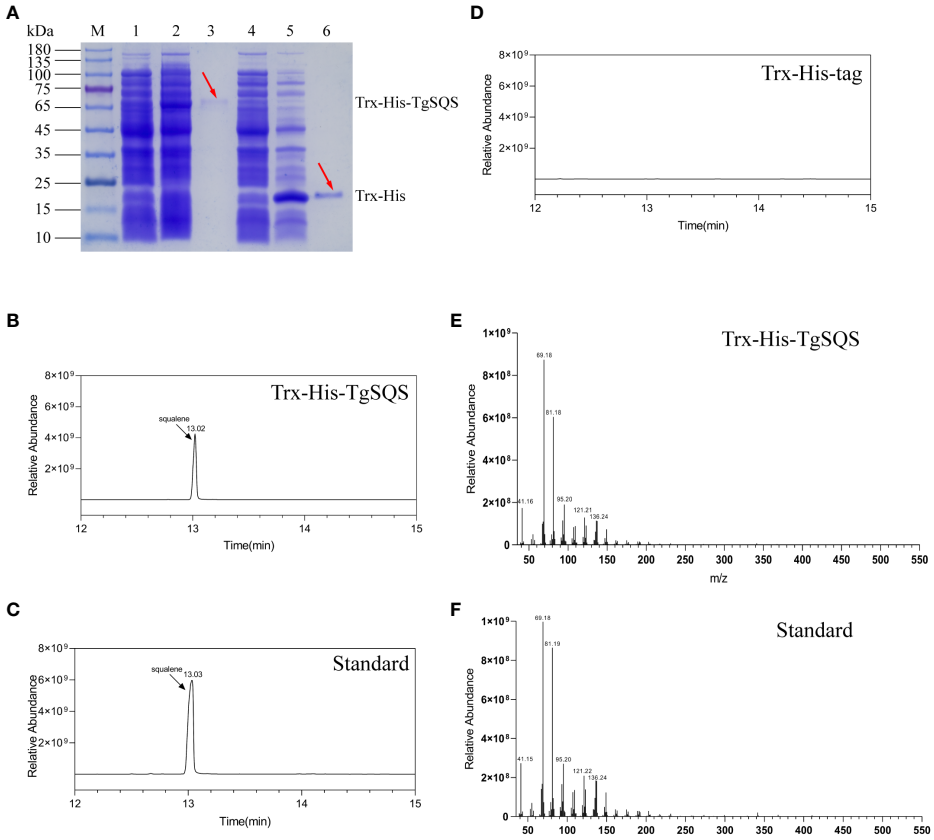


FIGURE 2
Identification of recombinant TgSQS protein and gas chromatography-mass spectroscopy (GC–MS) analysis of squalene in the catalytic products of TgSQS. **(A)** 10% SDS-PAGE detection of recombinant TgSQS protein expressed in ***(E)*** *coli* Rosetta (DE3). M, protein marker; Line 1, induced *pET-32a-TgSQS* bacteria; Line 2, supernatant of induced *pET-32a-TgSQS* bacteria after ultrasound; Line 3, purified *pET-32a-TgSQS* protein; Line 4, induced *pET-32a* bacteria; Line 5, supernatant of induced *pET-32a* bacteria after ultrasound; Line 6, purified Trx-His-Tag; **(B)** GC–MS detection of the reaction products of Trx-His-TgSQS; **(C)** GC–MS detection of the standard squalene; **(D)** GC–MS detection of the reaction products of Trx-His-Tag (control); **(E)** MS analysis of the catalytic products of Trx-His-TgSQS; **(F)** MS analysis of standard squalene.

verify the role of *TgSQS* in the sterol biosynthetic pathway, we used the super promoter to express *TgSQS* in *A. thaliana* (Figure 3A) and tested the β-sitosterol content in mature seeds, as can be seen from Figure 3C, which indicates that the β-sitosterol content in the mature seeds of three *TgSQS*-overexpressing lines increased by 15%, 15%, and 30% compared with the wild type, indicating that constitutively increasing *TgSQS* expression in *Arabidopsis* significantly increased β-sitosterol content in mature seeds. At the same time, the campesterol and stigmasterol content were also detected in these materials (Figure 3D); the results showed that the campesterol content in three *TgSQS*-overexpressing lines were also significantly higher than that of the wild type. However, the stigmasterol content showed no difference between the three *TgSQS*-overexpressing lines and the wild type.

TgWRKY3 bound the W-box elements in the TgSQS promoter

Due to the important role of *TgSQS* in sterol biosynthesis, we studied its transcriptional regulation mechanism. Through cloning and analysis of its promoter (1,860 bp before ATG), three W-box cis-acting elements were found in the entire promoter region. Since the W-box element was the DNA binding site of WRKY family transcription factors, WRKY TFs were regarded as candidates to regulate *TgSQS* expression. In total, 23 *TgWRKYs* with full-length coding sequences were found by searching the transcriptome data of *T. grandis* (NCBI accession Nos. SRX9417001–SRX9417009). All of them were constructed in the overexpression vector *pCAMBIA1300-GFP*, and the dual luciferase assay was carried out with the empty vector as the control. TgWRKY3, TgWRKY6, and TgWRKY7 significantly increased the expression level of the reporter gene (Figures 4B, C), indicating that they may regulate the expression of *TgSQS*.

To verify whether TgWRKY3, TgWRKY6, and TgWRKY7 could directly bind to the promoter of *TgSQS*, a yeast one-hybrid experiment was carried out. In the Y1H assay, the pAbAi vector carrying W1, W2, and W3 (200-bp sequences containing W-box from the *TgSQS* promoter, Figures 4D, S2) and the pGADT7-TgWRKY3/6/7 recombinant vector served as the reporter and the effector, respectively. Only Y1HGold yeast co-transformed with pGADT7-TgWRKY3 and pAbAi-W1 grew normally on the SD/−Leu medium supplied with 200 ng/ml AbA, whereas the yeast carrying the pGADT7-TgWRKY6/7 and pGADT7 empty

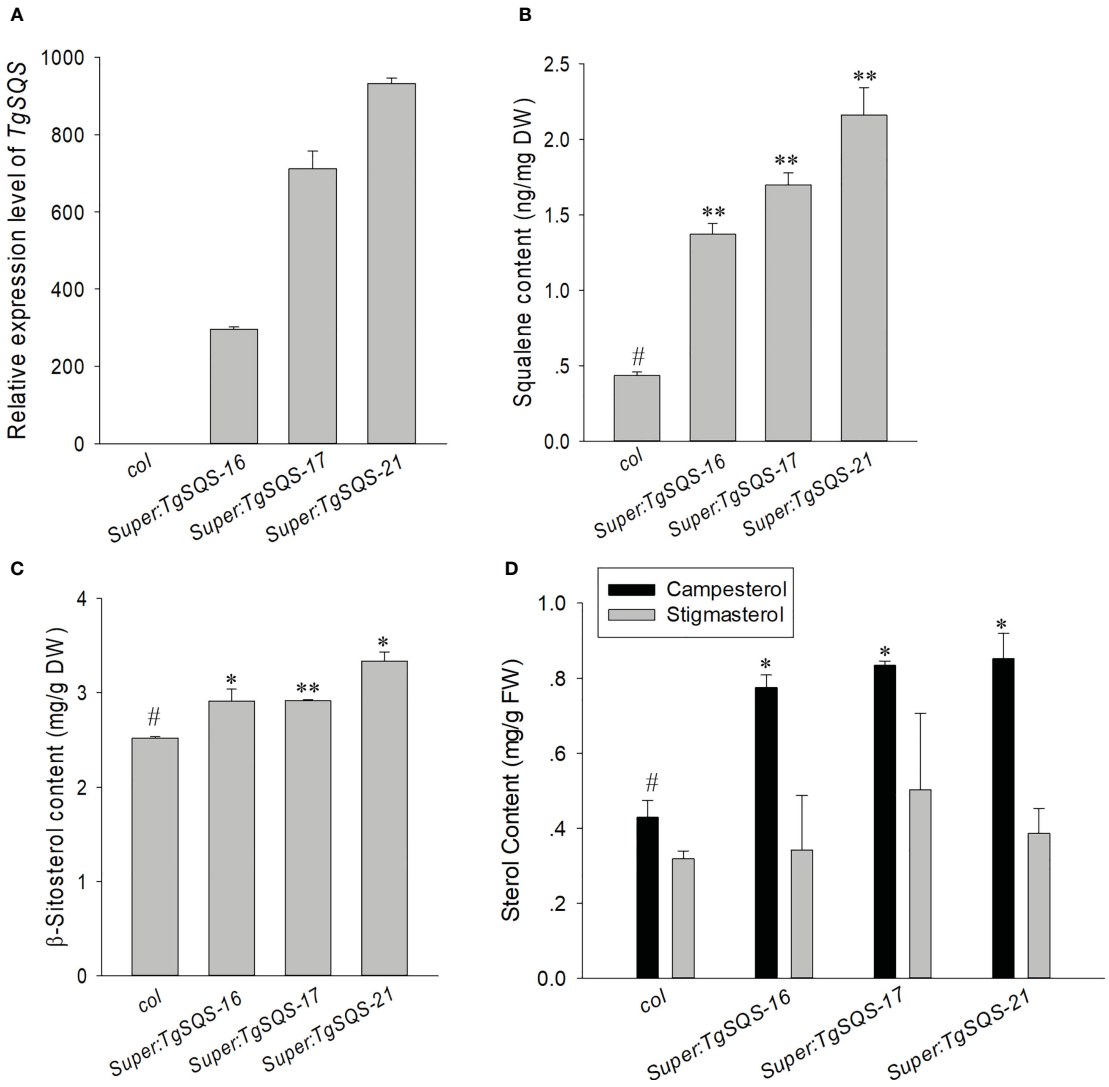


FIGURE 3
Increased squalene and β-sitosterol content in the seeds of *TgSQS*-overexpressing *Arabidopsis thaliana* lines. **(A)** Transcript level of *TgSQS*, **(B)** seed squalene content, **(C)** seed β-sitosterol content, and **(D)** stigmasterol and campesterol content in *col* and three *TgSQS*-overexpressing *Arabidopsis thaliana* lines. Data are mean ± SD (n = 3); asterisks indicate significant differences relative to the control by a two-tailed Student's *t*-test. #, control *p <0.05; **p <0.001.

vector did not (Figure 4E), indicating that TgWRKY3 directly bound the W-box-containing region in the *TgSQS* promoter.

Sterol biosynthesis-related genes in *T. grandis* can respond to drought stress

Preliminary experimental results showed that *T. grandis* seedlings had strong drought tolerance. Therefore, we carried out drought treatment on *T. grandis* seedlings and sampled the leaves at the D40 stage (drought treatment for 40 days when soil moisture was zero but seedlings stayed green) and D60 stage (drought treatment for 60 days when seedlings became wilted) to perform transcriptome sequencing with normally watered seedlings (CK, sampled at the D60 stage) as a control. PCA showed low variability among biological repeats, which indicated that there was a high correlation between repetitions (Figure 5A). The nine libraries produced over 6G of clean bases, with Q30 percentages (percentage of sequences with sequencing error rates <0.1%) ranging from 94% to 95% (Table S3). All bases were assembled into 66,798 unigenes with a mean length of 1,241 bp (Table S4). The GO, NT, NR, SwissProt, PFAM, KO, and KOG databases were used to annotate the predicted protein sequence, and 33,318 unigenes were annotated by at least one database (Table S5). These data showed that RNA-seq was of high quality and could be used for further analysis.

A total of 2,013 DEGs, including 1,505 up and 508 downregulated genes, were identified in the D40 stage compared with CK (Figure 5B). In the D60 stage, compared with CK, 9,156 genes were upregulated and 8,880 genes were downregulated (Figure 5C), indicating that there were many more DEGs in the latter. Moreover, the DEGs of D40vsCK and D60vsCK did not completely overlap. As shown in Figure S6, a total of 1,336 DEGs overlapped, of

10.3389/fpls.2023.1136643

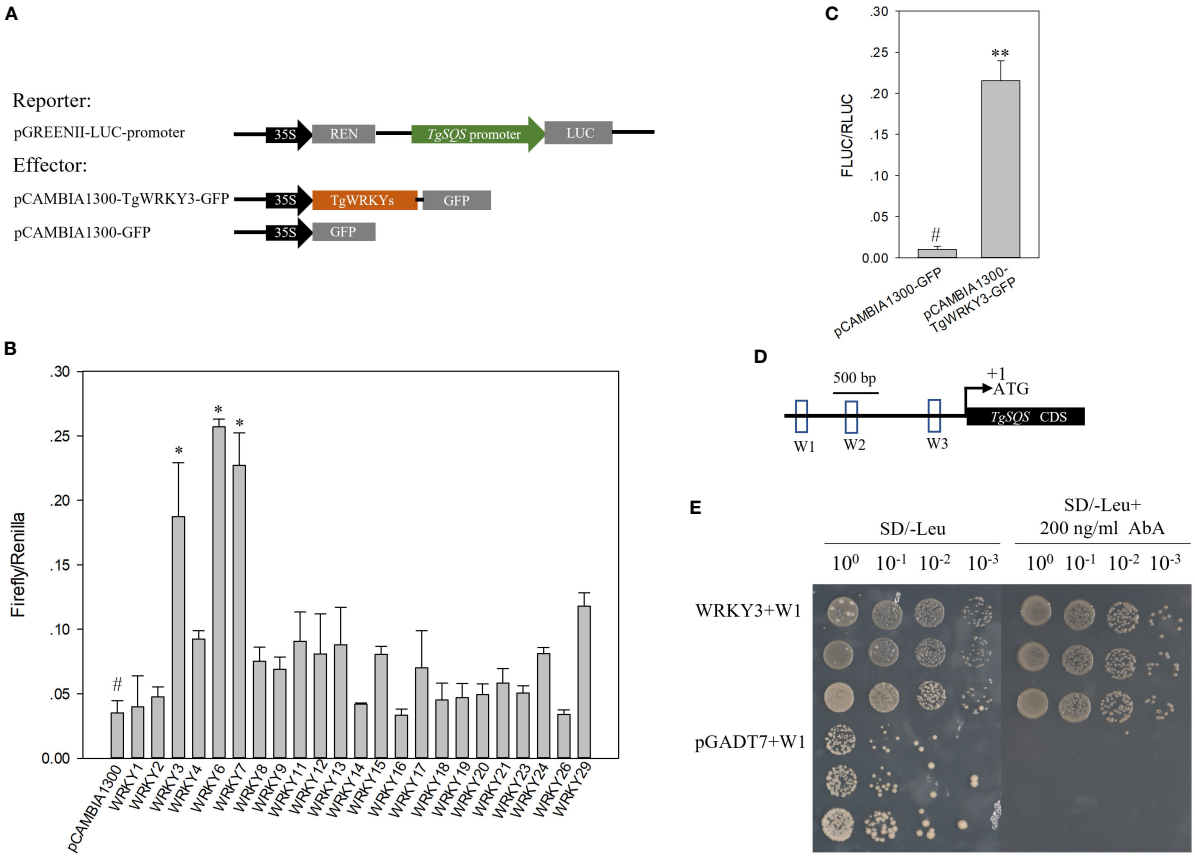


FIGURE 4
TgWRKY3 directly bound to the promoter of *TgSQS* and positively regulated *TgSQS* expression in tobacco leaves. **(A)** Schematic diagrams of vectors used for the dual luciferase assay; **(B)** Effects of TgWRKYs on the activities of the *TgSQS* promoter. Each column represents the mean ± SD of three biological replicates. Asterisks indicate significant differences relative to the control as determined by a two-tailed Student's *t*-test. #, control; *p <0.05. **(C)** Effects of TgWRKY3 on the activities of the *TgSQS* promoter. Data are the mean ± SD (n = 4); asterisks indicate significant differences relative to the control by a two-tailed Student's *t*-test. **p < 0.001. **(D)** Cis-element analysis in the sequence of the *TgSQS* promoter (the bait); W1, W2, and W3 represent 200-bp nucleotide sequences containing the W-box. The nucleotide sequences of W1, W2, and W3 are shown in Figure S2. **(E)** Y1H verification of the interaction between TgWRKY3 and the *TgSQS* promoter. The pAbAi vector carrying W1, W2, and W3 (200-bp sequences containing the W-box from the *TgSQS* promoter) and the pGADT7-TgWRKY3 recombinant vector were used to co-transfect Y1HGold receptive cells. Y1HGold cells co-transfected with pAbAi-W2/W3 and pGADT7-TgWRKY3 did not grow normally on the SD/–Leu medium supplied with 200 ng/ml AbA; thus, the pictures were not shown. From left to right, the dilutions of the bacterial solution are 1, 0.1, 0.01, and 0.001. Three monoclonal yeast strains are repeated.

which 461 were jointly upregulated and 266 were jointly downregulated. Brassinosteroid biosynthesis and terpenoid backbone biosynthesis were found in the KEGG path enrichment scatter plot at the D60 stage compared with CK (Figures S4, S5), indicating that genes related to terpenoid and sterol biosynthesis were involved in the response of *T. grandis* to drought stress, considering that campesterol is a precursor of brassinosteroids. Consequently, we combined the sterol biosynthesis pathway with the *T. grandis* transcriptome data in this study to draw a simplified map and mark the expression changes of related genes in CK, D40, and D60 stages in the form of a heat map (Figure 5D). The expression levels of *HMGS*, *HMGR*, *MK*, *DXS*, *IPPI*, *FPPS*, *SQS*, and *DWF1* increased significantly at the D60 stage (Table S6). At the same time, the sterol content in the leaves of the CK, D40, and D60 stages was detected, and the results showed that the content of sitosterol, campesterol, and stigmasterol in the D60 stage was significantly higher than in the CK, while there was no significant difference between the D40 stage and the CK (Figure 6). This result was consistent with the fact that the differential genes of the D40 stage in the transcriptome were far less than those in the D60 stage, which might be because the soil moisture at the D40 stage had just dropped to zero and the *T. grandis* seedlings were still in the early stages of drought stress. Together, these results indicate that the sterol biosynthesis pathway is involved in the response of *T. grandis* to drought stress.

TgSQS responds to drought stress by affecting ROS accumulation in A. thaliana

To verify the function of *TgSQS* under drought stress, three homozygous overexpressing *A. thaliana* lines and the wild type were used for drought treatment. After 20 days of drought, the wild-type leaves began to wilt, while the leaves of *TgSQS*-overexpressing lines stretched normally, indicating that they were more drought-tolerant than the wild type (Figure 7C). To deeply understand the drought tolerance mechanism, the accumulation of H_2O_2 and

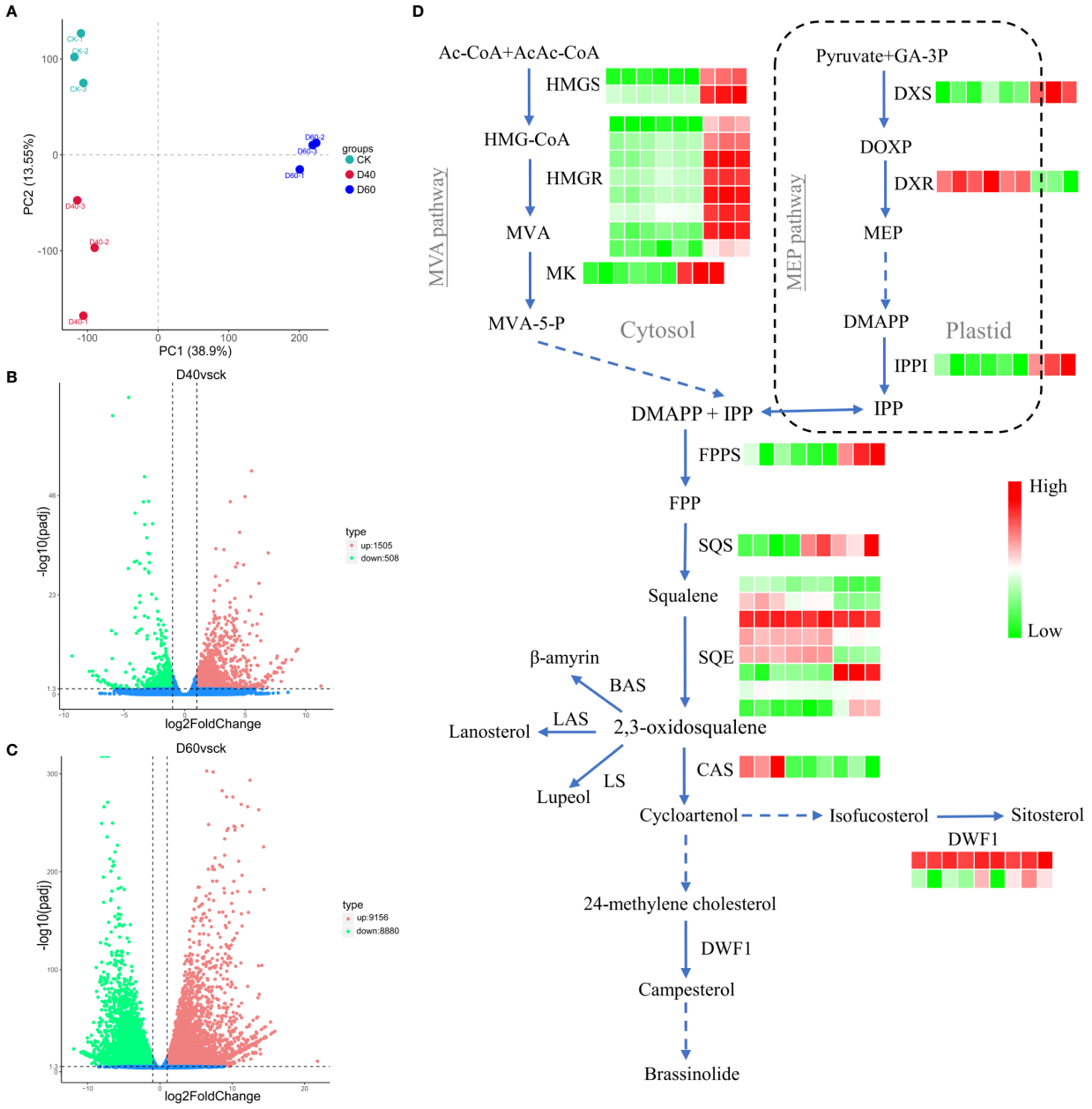


FIGURE 5
Sterol biosynthesis-related genes in *Torreya grandis* responded to drought stress. **(A)** Principal component analysis (PCA) of the genes. **(B)** Volcano plots for DEGs from CKvsD40; **(C)** Volcano plots for DEGs from CKvsD60. **(D)** Simplified view of isoprenoid biosynthesis in plants. Solid blue arrows indicate single-step reactions, dashed arrows denote several steps, and the double arrow between the cytosolic and plastid compartments indicates metabolic crosstalk between them. In the heat map, the first three squares from left to right represent CK samples, the three in the middle represent D40 samples, and the three on the right side represent the three D60 samples. Ac-CoA, acetyl CoA; AcAc-CoA, acetoacetyl CoA; BAS, b-amyrin synthase; BR6OX2, brassinosteroid-6-oxidase 2; CAS, cycloartenol synthase; DMAPP, dimethylallyl diphosphate; DWF1, D24 sterol reductase; DOXP, 1-deoxy-D-xylulose 5-phosphate; DXR, 1-deoxy-Dxylulose 5-phosphate reductoisomerase; DXS, 1-deoxy-D-xylulose-5-phosphate synthase; FPP, farnesyl diphosphate; FPPS, FPP synthase; GA-3P, glyceraldehyde 3-phosphate; HMG-CoA, 3-hydroxy-3-methylglutaryl-CoA; HMGR, HMG-CoA reductase; HMGS, HMG-CoA synthase; IPP, isopentenyl diphosphate; IPPI, isopentenyl diphosphate isomerase; LAS, lanosterol synthase; LS, lupeol synthase; MEP, 2-C-methyl-D-erythritol 4-phospahte; MVA, mevalonate; MVA-5-P, 5-phosphomevalonate; MK, mevalonate kinase; SQE, squalene monooxygenase/epoxidase; SQS, squalene synthase.

superoxide in the overexpressing lines and the wild type after simulated drought stress was observed by DAB and NBT staining. In this study, 10-day-old seedlings were treated with 10% PEG for 72 h, followed by 3,3-diaminobenzidine (DAB) staining to determine the presence of H_2O_2 and nitro blue tetrazolium (NBT) staining to show the presence of the superoxide anion. As shown in Figures 7A, B, under control conditions, wild-type and transgenic plants showed similar basal levels of H_2O_2 and superoxide, but under 10% PEG treatment, TgSQS-OE leaves and roots showed less H_2O_2 and superoxide accumulation than wild-type leaves and roots. These results suggest that *TgSQS* improves drought tolerance by reducing ROS accumulation.

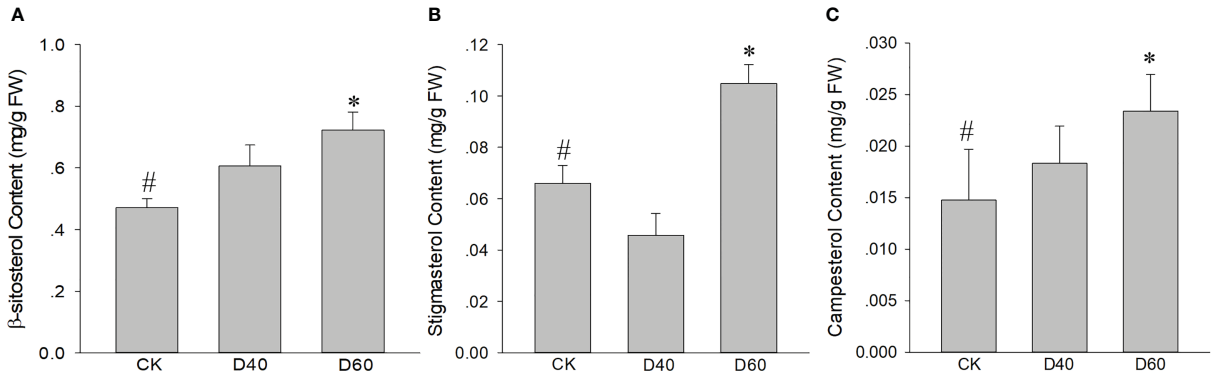


FIGURE 6
Increased **(A)** β-sitosterol content, **(B)** stigmasterol content, and **(C)** campesterol content in drought-treated *Torreya grandis* seedlings. Data are the mean ± SD (n = 3); asterisks indicate significant differences relative to the control by a two-tailed Student's *t*-test. #, control; *p <0.05.

Discussion

In this study, we found that a unigene in the transcriptome data of *T. grandis* might be the gene encoding functional squalene synthase, and we verified its function *in vitro* and *in vivo* (Figures 1–3). Since *AtSQS* was cloned in 1995 (Nguyen et al., 2013), the *SQS* of many species, such as *Tripterygium wilfordii* (Zhang et al., 2016), soybean (Nguyen et al., 2013), *P. ginseng*

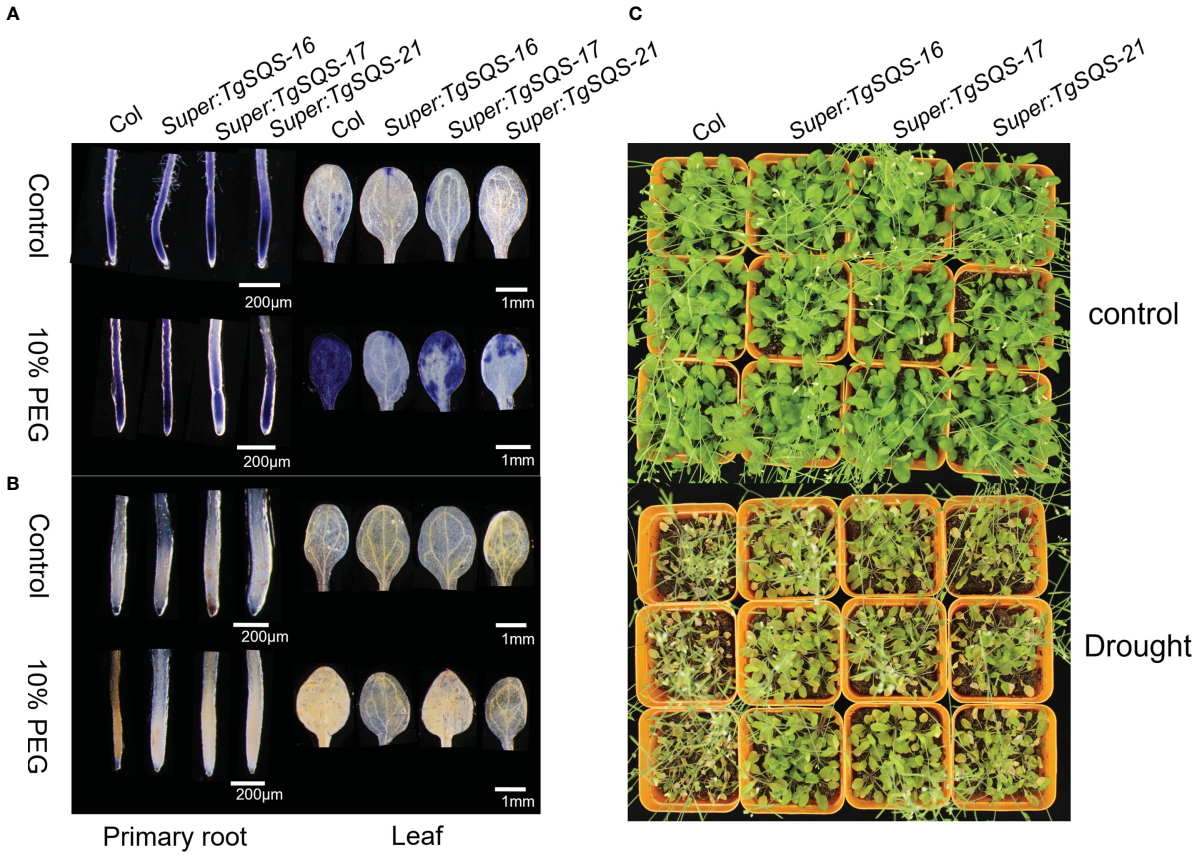


FIGURE 7
TgSQS responded to drought stress by affecting ROS accumulation in *Arabidopsis thaliana*. **(A)** Nitroblue tetrazolium (NBT) staining in the primary root tip and leaf of *TgSQS-OE* and *col* 13-day-old seedlings after treatment with 10% PEG for three days. The color strength shows the O^{2-} concentration in the root tips or leaves. **(B)** 3,3-Diaminobenzidine (DAB) staining in the primary root tip and leaf of *TgSQS-OE* and *col* 13-day-old seedlings after treatment with 10% PEG for three days. The color strength shows the H_2O_2 concentration in the root tips or leaves. **(C)** The phenotypes of *TgSQS-OE* and *Arabidopsis col* seedlings under drought stress treatment. Three-week-old *col* and *TgSQS-OE* seedlings were subjected to drought stress without water for 20 days.

(Kim et al., 2011), *Solanum nigrum* (Susan et al., 1991), *Siraitia grosvenorii* (Zhou et al., 2012), birch (Zhao et al., 2009), and persimmon (Zhou et al., 2012), have been cloned. Based on these reports, the number of *SQS* varies among different species, which may be a way for plants to meet changes in different developmental and environmental situations. Several literature reports have shown that SQS is a key enzyme for sitosterol synthesis, and changes in both SQS activity and gene expression levels can affect the phytosterol content. For example, treatment of tobacco suspension cells with SQS inhibitors and fungal elicitors could reduce their SQS activity, thereby significantly reducing their sterol content (Ding et al., 2020; Wollam and Antebi, 2011). The content of β-sitosterol significantly increased after *PgSQS* overexpression in *P. ginseng* and *Acanthopanax senticosus* (Li et al., 2019; Sikder et al., 2014). After reducing the expression of *SQS* by virus-induced gene silencing in *W. somnifera* leaves, a significant decrease in sterol content was detected (Singh et al., 2017). In this paper, we demonstrated that *TgSQS* overexpression in *Arabidopsis* significantly increased the content of β-sitosterol and campesterol in mature seeds (Figure 3), which was consistent with the literature (Nguyen et al., 2013), suggesting that we could increase the content of β-sitosterol and campesterol in *T. grandis* nuts to improve its nutritional value by increasing *TgSQS* expression.

Until now, only WsWRKY1 had been known to regulate *WsSQS* and *WsSQE* in *W. somnifera*, according to reports on transcriptional regulation of *SQS* (Sun et al., 2012). In the present study, we demonstrated that TgWRKY3 directly bound to the promoter region of *TgSQS* and upregulated its expression (Figure 4). WRKY domain-containing genes comprise one of the largest TF families in plants and are characterized by a highly conserved WRKYGQK motif at the N-terminal end, together with a

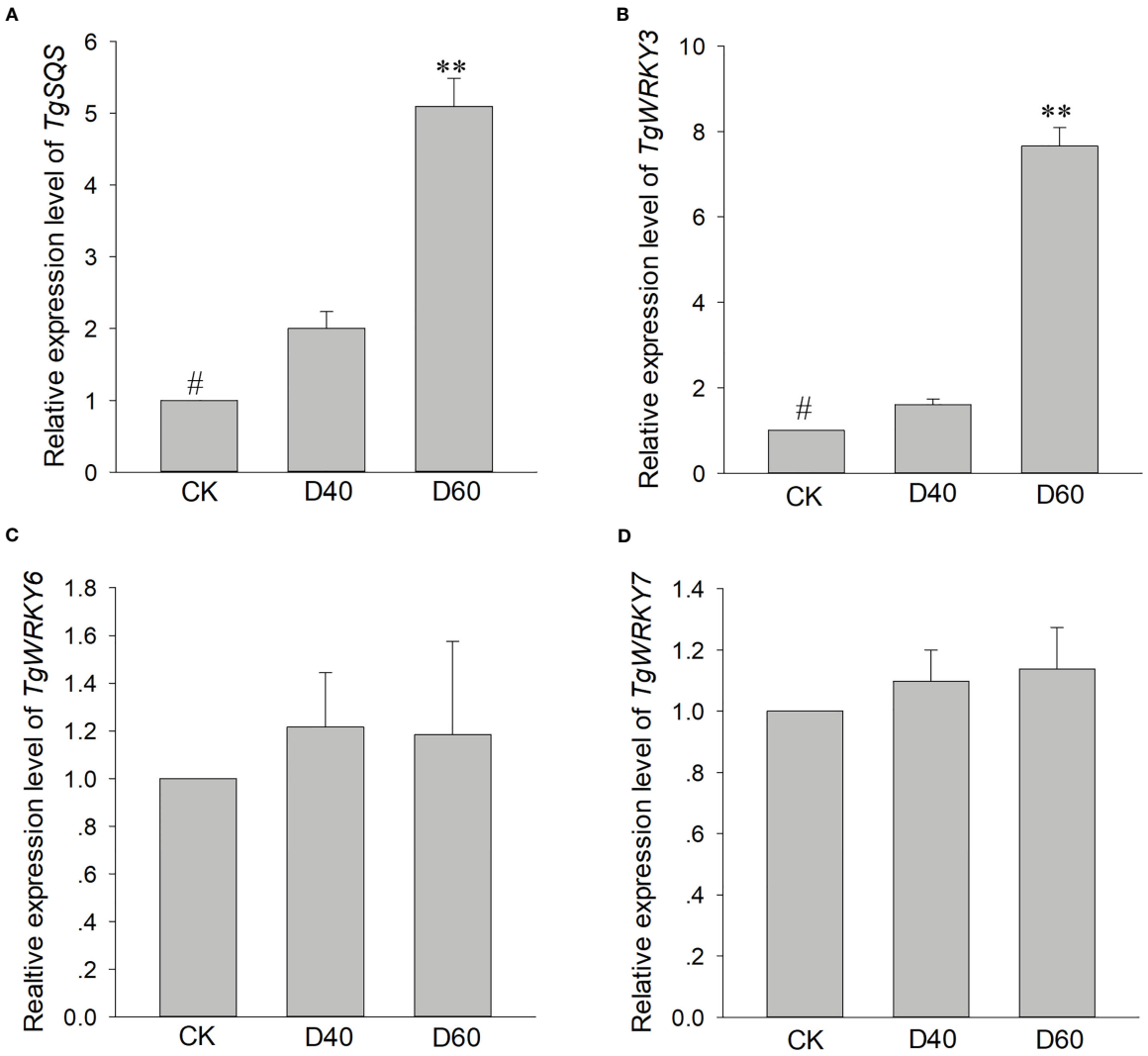


FIGURE 8
qRT-PCR analysis of gene expression patterns of *TgSQS* **(A)**, *TgWRKY3* **(B)**, *TgWRKY6* **(C)**, and *TgWRKY7* **(D)** among drought treatment stages. Data are the mean ± SD (n = 3); asterisks indicate significant differences relative to the control by a two-tailed Student's *t*-test. #, control; **p <0.0001.

novel zinc-finger-like motif (Ferrer et al., 2017). In *A. thaliana*, the WRKY family contains 72 members, which can be divided into three categories, and the second category can be further divided into five subclasses (a–e) (Rushtone et al., 2010). The phylogenetic tree analysis of TgWRKY3, WsWRKY1, and WRKY TFs in *A. thaliana* showed that TgWRKY3 has a close relationship with AtWRKY21, AtWRKY74, and AtWRKY39 in Group II-d (Figure S3) and that WsWRKY1 has a close relationship with AtWRKY41 and AtWRKY53 in Group III, indicating that we found another WRKY that regulates *SQS* expression. Multiple sequence alignment results showed that TgWRKY3 has a WRKYGQK conserved domain and the characteristic zinc finger domain C2H2, indicating that *TgWRKY3* indeed belongs to the WRKY TFs.

As one of the largest transcription factor families in plants, WRKY TFs have been found to play a role in plant development and tolerance to a variety of abiotic stressors, including wounding, drought, salt, heat, and cold pressure (Choi et al., 2003; Ju et al., 2004; Wentzinger et al., 2002; Chen et al., 2012). Recent research has also shown that WRKY TFs can participate in regulating the biosynthesis of multiple metabolites, such as sesquiterpenes, alkaloids, and terpenes (Seo et al, 2005). For example, MrWRKY1 interacts with the promoter of *MrFPS* to regulate the biosynthesis of α-Bisabolol in *Matricaria recutita L.* (Unland et al., 2018), and NbWRKY8 can regulate capsidiol biosynthesis by binding the promoter of NbHMGR2 in *N. benthamiana* (Jadaun et al., 2017). In our study, TgWRKY3, TgWRKY6, and TgWRKY7 both significantly increased the expression level of the reporter gene, but only TgWRKY3 directly bound to the promoter of TgSQS (Figure 4), indicating that TgWRKY6 and TgWRKY7 regulate *TgSQS* in a manner that interacts with other transcription factors. However, the expression level of *TgWRKY3* in *T. torreya* seedlings after drought treatment significantly increased, while *TgWRKY6* and *TgWRKY7* did not (Figures 8B–D), suggesting that TgWRKY3 is involved in both the synthesis of β-sitosterol and the plant's response to drought stress, while TgWRKY6 and TgWRKY7 might also regulate *TgSQS* and be involved in the response to different stressors.

In this study, we found that *TgSQS*-overexpressing lines showed stronger drought tolerance and less ROS accumulation than the wild type (Figure 7), indicating that *TgSQS* probably responded to drought stress by reducing ROS accumulation. However, the specific mechanism is still unclear. Our results also showed that the content of phytosterols (β-sitosterol, campesterol, and stigmasterol) in *T. grandis* seedlings after drought treatment was significantly higher than that in the control; furthermore, the β-sitosterol content was at least one order of magnitude higher than that of campesterol and stigmasterol (Figure 6). Studies have shown that an increase in β-sitosterol content can be observed in rice under drought conditions (Lee et al., 2004; Kumar et al., 2015). Additionally, drought tolerance and total antioxidant capacity were significantly improved with 100 μM sitosterol treatment in *T. aestivum* and 120 μM sitosterol treatment in *Trifolium repens* (Eulgem et al., 2000; Lou et al., 2019). These results imply that β-sitosterol is involved in the response to drought stress. β-sitosterol has been reported as a component of the cell membrane (Vriet et al., 2012); thus, the response of β-sitosterol to drought is probably realized by changing the membrane fluidity and permeability. However, the effect of *TgSQS* on ROS accumulation is perhaps related to brassinosteroids (BRs). As a plant hormone, BRs control several traits of agronomic importance, such as seed germination, plant architecture, seed yield, and tolerance to various abiotic and biotic stressors (Du et al, 2022; Wani et al., 2021). Several studies have shown that plants treated with exogenous BRs are drought tolerant, and in tomato, it has been clearly demonstrated that a rise in the level of BR biosynthesis is the key to enhancing the tolerance capacity (Nolan et al., 2020; Pandit et al., 2000). In another study, BR treatment rescued the heavy aggregation of ROS in the case of drought stress (Zhang et al., 2018). Therefore, *TgSQS* might affect ROS accumulation by regulating BR biosynthesis.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/ , SAMN32775921 https://www.ncbi.nlm.nih.gov/, SAMN32775922 https://www.ncbi.nlm.nih.gov/, SAMN32775923 https://www.ncbi.nlm.nih.gov/, SAMN32775924 https://www.ncbi.nlm.nih.gov/ , SAMN32775925 https://www.ncbi.nlm.nih.gov/, SAMN32775926 https://www.ncbi.nlm.nih.gov/, SAMN32775927 https://www.ncbi.nlm.nih.gov/, SAMN32775928 https://www.ncbi.nlm.nih.gov/ , SAMN32775929.

Author contributions

FZ conceived and guided the experiments, analyzed the data, and wrote the draft. CK performed the experiment and analyzed the data. ZM performed the experiment. WC performed the experiment. YL performed the experiment. HL conceived and guided the experiments. JW is responsible for resources and supervision. All authors contributed to the article and approved the submitted version.

Funding

This work was funded by the National Natural Science Foundation of China (31901339).

Acknowledgments

We are grateful to Dr. Zuying Zhang for providing pET-32a vector. We thank LetPub (www.letpub.com) for its linguistic assistance during the preparation 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.2023.1136643/full#supplementary-material

References

Awad, A. B., Chinnam, M., Fink, C. S., and Bradford, P. G. (2007). Beta-sitosterol activates fas signaling in human breast cancer cells. *Phytomedicine* 14, 747–754. doi: 10.1016/j.phymed.2007.01.003

Awad, A. B., Fink, C. S., and Williams, H. (2001). *In vitro* and *in vivo* (SCID mice) effect of phytosterols on the growth and dissemination of human prostate cancer PC-3 cells. *Eur. J. Cancer Prev.* 10, 507–513. doi: 10.1097/00008469-200112000-00005

Awad, A. B., Hernandez, A. Y. T., Fink, C. S., and Mendel, S. L. (1997). Effect of dietary phytosterols on cell proliferation and protein kinase cactivity in rat colonic mucosa. *Nutr. Cancer* 27, 210–215. doi: 10.1080/01635589709514527

Babu, S., and Jayaraman, S. (2020). An update on β-sitosterol: a potential herbal nutraceutical for diabetic management. *BioMed. Pharmacother.* 131, 110702. doi: 10.1016/j.biopha.2020.110702

Chen, L., Song, Y., Li, S., Zhang, L., Zou, C., and Yu, D. (2012). The role of WRKY transcription factors in plant abiotic stresses. *Biochim. Biophys. Acta* 1819, 120–128. doi: 10.1016/j.bbagrm.2011.09.002

Choi, Y. H., Kong, K. R., Kim, Y. A., Jung, K. O., Kil, J. H., Rhee, S. H., et al. (2003). Induction of bax and activation of caspases during beta-sitosterol-mediated apoptosis in human colon cancer cells. *Int. J. Oncol.* 23, 1657–1662. doi: 10.3892/ijo.23.6.1657

Clough, S. J., and Bent, A. F. (1998). Floral dip: a simplified method for agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16, 735–743. doi: 10.1046/j.1365-313x.1998.00343.x

Devarenne, T. P., Ghosh, A., and Chappell, J. (2002). Regulation of squalene synthase, a key enzyme of sterol biosynthesis, in tobacco. *Plant Physiol.* 129, 1095–1106. doi: 10.1104/pp.001438

Ding, M., Lou, H., Chen, W., Zhou, Y., Zhang, Z., Xiao, M., et al. (2020). Comparative transcriptome analysis of the genes involved in lipid biosynthesis pathway and regulation of oil body formation in *Torreya grandis* kernels. *Ind. Crop Prod.* 145, 112051. doi: 10.1016/j.indcrop.2019.112051

Divi, U. K., and Krishna, P. (2009). Brassinosteroid: a biotechnological target for enhancing crop yield and stress tolerance. *New Biotechnol.* 26, 131–136. doi: 10.1016/j.nbt.2009.07.006

Du, Y., Fu, X., Chu, Y., Wu, P., Liu, Y., Ma, L., et al. (2022). Biosynthesis and the roles of plant sterols in development and stress responses. *Int. J. Mol. Sci.* 23 (4), 2332. doi: 10.3390/ijms23042332

Elkeilsh, A., Awad, Y. M., Soliman, M. H., Abu-Elsaoud, A., Abdelhamid, M. T., and El-Metwally, I. M. (2019). Exogenous application of β-sitosterol mediated growth and yield improvement in water-stressed wheat (*Triticum aestivum*) involves up-regulated antioxidant system. *J. Plant Res.* 132, 881–901. doi: 10.1007/s10265-019-01143-5

Eulgem, T., Rushton, P. J., Robatzek, S., and Somssich, I. E. (2000). The WRKY superfamily of plant transcription factors. *Trends Plant Sci.* 5, 199–206. doi: 10.1016/s1360-1385(00)01600-9

Ferrer, A., Altabella, T., Arró, M., and Boronat, A. (2017). Emerging roles for conjugated sterols in plants. *Prog. Lipid Res.* 67, 27–37. doi: 10.1016/j.plipres.2017.06.002

Goldstein, J. L., and Brown, M. S. (1990). Regulation of the mevalonate pathway. *Nature* 343, 425–430. doi: 10.1038/343425a0

Guan, G. M., Dai, P. H., and Shechter, I. (1998). Differential transcriptional regulation of the human squalene synthase gene by sterol regulatory element-binding proteins (SREBP) 1a and 2 and involvement of 5′DNA sequence elements in the regulation. *J. Biol. Chem.* 273, 12526–12535. doi: 10.1074/jbc.273.20.12526

He, Z. Y., Zhu, H. D., Li, W. L., Zeng, M. M., Wu, S. F., Chen, S. W., et al. (2016). Chemical components of cold pressed kernel oils from different *Torreya grandis* cultivars. *Food Chem.* 209, 196–202. doi: 10.1016/j.foodchem.2016.04.053

Huang, Z. S., Jiang, K. J., Pi, Y., Hou, R., Liao, Z. H., Cao, Y., et al. (2007). Molecular cloning and characterization of the yew gene encoding squalene synthase from *Taxus cuspidata*. *J. Biochem. Mol. Biol.* 40, 625–635. doi: 10.5483/bmbrep.2007.40.5.625

Ishihama, N., Yamada, R., Yoshioka, M., Katou, S., and Yoshioka, H. (2011). Phosphorylation of the *Nicotiana benthamiana* WRKY8 transcription factor by MAPK functions in the defense response. *Plant Cell* 23, 1153–1170. doi: 10.1105/tpc.110.081794

Jadaun, J. S., Sangwan, N. S., Narnoliya, L. K., Singh, N., Bansal, S., Mishra, B., et al. (2017). Over-expression of DXS gene enhances terpenoidal secondary metabolite accumulation in rose-scented geranium and *Withania somnifera*: active involvement of plastid isoprenogenic pathway in their biosynthesis. *Physiol. Plant* 159, 381–400. doi: 10.1111/ppl.12507

Jiang, J. J., Ma, S. H., Ye, N. H., Jiang, M., Cao, J. S., and Zhang, J. H. (2017). WRKY transcription factors in plant responses to stresses. *J. Integr. Plant Biol.* 59 (2), 86–101. doi: 10.1111/jipb.12513

Ju, Y. H., Clausen, L. M., Allred, K. F., Almada, A. L., and Helferich, W. G. (2004). β-sitosterol, β -sitosterol glucoside, and a mixture of β -sitosterol and β-sitosterol glucoside modulate the growth of estrogen-responsive breast cancer cells *in vitro* and in ovariectomized athymic mice. *J. Nutr.* 134, 1145–1151. doi: 10.1093/jn/134.5.1145

Kennedy, M. A., and Bard, M. (2001). Positive and negative regulation of squalene synthase (*erg9*), an ergosterol biosynthetic gene, in *Saccharomyces cerevisiae*. *Biochim. Biophys. Acta* 1517, 177–189. doi: 10.1016/s0167-4781(00)00246-3

Kim, T. D., Han, J. Y., Huh, G. H., and Choi, Y. E. (2011). Expression and functional characterization of three squalene synthase genes associated with saponin biosynthesis in *Panax ginseng*. *Plant Cell Physiol.* 52, 125–137. doi: 10.1093/pcp/pcq179

Kumar, M. S., Ali, K., Dahuja, A., and Tyagi, A. (2015). Role of phytosterols in drought stress tolerance in rice. *Plant Physiol. Biochem.* 96, 83–89. doi: 10.1016/j.plaphy.2015.07.014

Lee, M. H., Jeong, J. H., Seo, J. W., Shin, C. G., Kim, Y. S., In, J. G., et al. (2004). Enhanced triterpene and phytosterol biosynthesis in *Panax ginseng* overexpressing squalene synthase gene. *Plant Cell Physiol.* 45 (8), 976–984. doi: 10.1093/pcp/pch126

Li, Z., Cheng, B., Yong, B., Liu, T., Peng, Y., Zhang, X., et al. (2019). Metabolomics and physiological analyses reveal β-sitosterol as an important plant growth regulator inducing tolerance to water stress in white clover. *Planta* 250, 2033–2046. doi: 10.1007/s00425-019-03277-1

Lou, H., Ding, M., Wu, J., Zhang, F., Chen, W., Yang, Y., et al. (2019). Full-length transcriptome analysis of the genes involved in tocopherol biosynthesis in *Torreya grandis*. *J. Agric. Food Chem.* 67, 1877–1888. doi: 10.1021/acs.jafc.8b06138

Nakashima, T., Inoue, T., Oka, A., Nishino, T., Osumi, T., and Hata, S. (1995). Cloning, expression, and characterization of cDNAs encoding *Arabidopsis thaliana* squalene synthase. *Proc. Natl. Acad. Sci. U.S.A.* 92 (6), 2328–2332. doi: 10.1073/pnas.92.6.2328

Nguyen, H. T., Neelakadan, A. K., Quach, T. N., Valliyodan, B., Kumar, R., Zhang, Z., et al. (2013). Molecular characterization of glycine max squalene synthase genes in seed phytosterol biosynthesis. *Plant Physiol. Biochem.* 73, 23–32. doi: 10.1016/j.plaphy.2013.07.018

Nie, S., Huang, S., Wang, S., Mao, Y., Liu, J., Ma, R., et al. (2019). Enhanced brassinosteroid signaling intensity *via* SlBRI1 overexpression negatively regulates drought resistance in a manner opposite of that *via* exogenous BR application in tomato. *Plant Physiol. Biochem.* 138, 36–47. doi: 10.1016/j.plaphy.2019.02.014

Nolan, T. M., Vukasinović, N., Liu, D., Russinova, E., and Yin, Y. (2020). Brassinosteroids: multidimensional regulators of plant growth, development, and stress responses. *Plant Cell* 32, 298–318. doi: 10.1105/tpc.19.00335

Pandit, J., Danley, D. E., Schulte, G. K., Mazzalupo, S., Pauly, T. A., Hayward, C. M., et al. (2000). Crystal structure of human squalene synthase, a key enzyme in cholesterol biosynthesis. *J. Biol. Chem.* 275 (39), 30610–30617. doi: 10.1074/jbc.M004132200

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

Salehi, B., Quispe, C., Sharifi-Rad, J., Cruz-Martins, N., Nigam, M., Mishra, A. P., et al. (2020). Phytosterols: from preclinical evidence to potential clinical applications. *Front. Pharmacol.* 11. doi: 10.3389/fphar.2020.599959

Schluttenhofer, C., and Yuan, L. (2015). Regulation of specialized metabolism by WRKY transcription factors. *Plant Physiol.* 167 (2), 295–306. doi: 10.1104/pp.114.251769

Seo, J. W., Jeong, J. H., Shin, C. G., Lo, S. C., Han, S. S., Yu, K. W., et al. (2005). Overexpression of squalene synthase in *Eleutherococcus senticosus* increases phytosterol and triterpene accumulation. *Phytochemistry* 66, 869–877. doi: 10.1016/j.phytochem.2005.02.016

Sikder, K., Das, N., Kesh, S. B., and Dey, S. (2014). Quercetin and beta-sitosterol prevent high fat diet induced dyslipidemia and hepatotoxicity in Swiss albino mice. *Indian J. Exp. Biol.* 52, 60–66. doi: 10.3109/03630269.2013.855936

Singh, A. K., Dwivedi, V., Rai, A., Pal, S., Reddy, S. G. E., Rao, D. K. V., et al. (2015). Virus-induced gene silencing of *Withania somnifera* squalene synthase negatively regulates sterol and defence-related genes resulting in reduced withanolides and biotic stress tolerance. *Plant Biotechnol. J.* 13, 1287–1299. doi: 10.1111/pbi.12347

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

Sun, Y., Zhao, Y., Wang, L., Lou, H. X., and Cheng, A. X. (2012). Cloning and expression analysis of squalene synthase, a key enzyme involved in antifungal steroidal glycoalkaloids biosynthesis from *Solanum nigrum*. *Drug Discovery Ther.* 6 (5), 242–248. doi: 10.5582/ddt.2012.v6.5.242

Susan, M. J., Yim, H. T., Tobe, M. F., and Gordon, W. R. (1991). Molecular cloning and characterization of the yeast gene for squalene synthetase. *Proc. Natl. Acad. Sci. U.S.A.* 88, 6038–6042. doi: 10.1073/pnas.88.14.6038

Tai, Y., Wang, H., Yao, P., Sun, J., Guo, C., Jin, Y., et al. (2023). Biosynthesis of α-bisabolol by farnesyl diphosphate synthase and α-bisabolol synthase and their related transcription factors in matricaria recutita l. *Int. J. Mol. Sci.* 24 (2), 1730. doi: 10.3390/ijms24021730

Unland, K., Pütter, K. M., Vorwerk, K., Deenen, N. V., Twyman, R. M., Prüfer, D., et al. (2018). Functional characterization of squalene synthase and squalene epoxidase in *Taraxacum koksaghyz*. *Plant Direct* 2 (6), e00063. doi: 10.1002/pld3.63

Valitova, J. N., Sulkarnayeva, A. G., and Minibayeva, F. V. (2016). Plant sterols: diversity, biosynthesis, and physiological functions. *Biochem. (Moscow)* 81 (8), 819–834. doi: 10.1134/S0006297916080046

Vriet, C., Russinova, E., and Reuzeau, C. (2012). Boosting crop yields with plant steroids. *Plant Cell* 24, 842–857. doi: 10.1105/tpc.111.094912

Wani, S. H., Anand, S., Singh, B., Bohra, A., and Joshi, R. (2021). WRKY transcription factors and plant defense responses: latest discoveries and future prospects. *Plant Cell Rep.* 40 (7), 1071–1085. doi: 10.1007/s00299-021-02691-8

Wentzinger, L. F., Bach, T. J., and Hartmann, M. A. (2002). Inhibition of squalene synthase and squalene epoxidase in tobacco cells triggers an up-regulation of 3-hydroxy-3-methylglutaryl coenzyme a reductase. *Plant Physiol.* 130, 334–346. doi: 10.1104/pp.004655

Wollam, J., and Antebi, A. (2011). Sterol regulation of metabolism, homeostasis, and development. *Annu. Rev. Biochem.* 80, 885–916. doi: 10.1146/annurev-biochem-081308-165917

Yuan, G. F., Jia, C. G., Li, Z., Sun, B., Zhang, L. P., Liu, N., et al. (2010). Effect of brassinosteroids on drought resistance and abscisic acid concentration in tomato under water stress. *Sci. Hortic.* 126, 103–108. doi: 10.1016/j.scienta.2010.06.014

Zhang, B., Liu, Y., Chen, M., Feng, J., Ma, Z., Zhang, X., et al. (2018). Cloning, expression analysis and functional characterization of squalene synthase (SQS) from *Tripterygium wilfordii*. *Molecules* 23 (2), 269. doi: 10.3390/molecules23020269

Zhang, M., Wang, S., Yin, J., Li, C., Zhan, Y., Xiao, J., et al. (2016). Molecular cloning and promoter analysis of squalene synthase and squalene epoxidase genes from *Betula platyphylla*. *Protoplasma* 253 (5), 1347–1363. doi: 10.1007/s00709-015-0893-3

Zhao, Y., Chang, S. K., Qu, G., Li, T., and Cui, H. (2009). Beta-sitosterol inhibits cell growth and induces apoptosis in SGC-7901 human stomach cancer cells. *J. Agr. Food Chem.* 57, 5211–5218. doi: 10.1021/jf803878n

Zhao, H., Tang, Q., Mo, C., Bai, L., Tu, D., and Ma, X. (2017). Cloning and characterization of squalene synthase and cycloartenol synthase from *Siraitia grosvenorii*. *Acta Pharm. Sin. B* 7, 215–222. doi: 10.1016/j.apsb.2016.06.012

Zhou, C., Zhao, D., Sheng, Y., Liang, G., and Tao, J. (2012). Molecular cloning and expression of squalene synthase and 2,3-oxidosqualene cyclase genes in persimmon (*Diospyros kaki l.*) fruits. *Mol. Biol. Rep.* 39, 1125–1132. doi: 10.1007/s11033-011-0841-z

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 21 December 2023
DOI 10.3389/fpls.2023.1255805

OPEN ACCESS

EDITED BY
Meng Kou,
Xuzhou Institute of Agricultural Sciences in Jiangsu Xuhuai District, China

REVIEWED BY
Lei Hua,
Peking University, China
Chao Wang,
University of California, Berkeley, United States

*CORRESPONDENCE
Zhimin Ma
sp13785181259@163.com
Lanfu Liu
llf6712@163.com
Xinliang Liu
liuxinliang0558@sina.com

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

RECEIVED 09 July 2023
ACCEPTED 06 November 2023
PUBLISHED 21 December 2023

CITATION
Hu Y, Dai Z, Huang J, Han M, Wang Z, Jiao W, Gao Z, Liu X, Liu L and Ma Z (2023) Genome-wide identification and expression analysis of the glutamate receptor gene family in sweet potato and its two diploid relatives.
Front. Plant Sci. 14:1255805.
doi: 10.3389/fpls.2023.1255805

Genome-wide identification and expression analysis of the glutamate receptor gene family in sweet potato and its two diploid relatives

Yaya Hu[1†], Zhuoru Dai[2†], Jinan Huang[1†], Meikun Han[1], Zhiwei Wang[3], Weijing Jiao[1], Zhiyuan Gao[1], Xinliang Liu[4*], Lanfu Liu[1*] and Zhimin Ma[1*]

[1]Hebei Key Laboratory of Crop Genetics and Breeding, Institute of Cereal and Oil Crops, Hebei Academy of Agriculture and Forestry Sciences, Shijiazhuang, Hebei, China, [2]Key Laboratory of Sweet Potato Biology and Biotechnology, Ministry of Agriculture and Rural Affairs, College of Agronomy & Biotechnology, China Agricultural University, Beijing, China, [3]Department of Agriculture Forestry and Biological Engineering, Baoding Vocational and Technical College, Baoding, Hebei, China, [4]School of Life Sciences, Jiangsu Normal University, Xuzhou, Jiangsu, China

Plant glutamate receptor (GLR) homologs are crucial calcium channels that play an important role in plant development, signal transduction, and response to biotic and abiotic stresses. However, the *GLR* gene family has not yet been thoroughly and systematically studied in sweet potato. In this study, a total of 37 *GLR* genes were identified in the cultivated hexaploid sweet potato (*Ipomoea batatas*), and 32 *GLR* genes were discovered in each of the two diploid relatives (*Ipomoea trifida* and *Ipomoea triloba*) for the first time. Based on their evolutionary relationships to those of *Arabidopsis*, these *GLRs* were split into five subgroups. We then conducted comprehensive analysis to explore their physiological properties, protein interaction networks, promoter *cis*-elements, chromosomal placement, gene structure, and expression patterns. The results indicate that the homologous *GLRs* of the cultivated hexaploid sweet potato and its two relatives are different. These variations are reflected in their functions related to plant growth, hormonal crosstalk, development of tuberous roots, resistance to root rot, and responses to abiotic stress factors, all of which are governed by specific individual *GLR* genes. This study offers a comprehensive analysis of *GLR* genes in sweet potato and its two diploid relatives. It also provides a theoretical basis for future research into their regulatory mechanisms, significantly influencing the field of molecular breeding in sweet potatoes.

KEYWORDS

glutamate receptor, tissue-specific expression, root rot stress, abiotic stress, sweet potato, *Ipomoea trifida*, *Ipomoea triloba*

1 Introduction

Glutamate, a ubiquitous amino acid, participates in various important chemical reactions in animals, plants, and microorganisms and plays an indispensable function in protein metabolism and signal transduction processes (Forde and Lea, 2007). Glutamate receptors (GLRs), including the ionotropic and metabotropic GLRs, were first discovered in animals. The ionotropic GLR of animal is a ligand-gated non-selective cation channel that regulates excitatory nerve signal transmission and guides neuronal development (Mayer, 2006; Zhu and Gouaux, 2017). Due to the existence of conserved signature domains, plant GLR structures are similar to those for animal ionotropic GLRs (Wudick et al., 2018a). Plant GLRs composed of receptor domains and four transmembrane helical domains are typical membrane proteins belonging to a class of amino acid-gated Ca^{2+} channels. When GLRs are activated by their corresponding ligands, they can mediate the transmembrane influx of Ca^{2+}, thus activating Ca^{2+} signaling to regulate plant responses to stress and simultaneously impact their overall development and growth (Sobolevsky et al., 2009; Traynelis et al., 2010; Kong et al., 2015; Tian et al., 2020; Luan and Wang, 2021; Ahmed et al., 2023).

Plant GLRs have been studied for more than 25 years, during which significant progress has been made in understanding their structural and functional characteristics. At present, 20, 13, 24, 34, 29, 35, 36, 16, 43, 16, and 21 *GLRs*, have been identified in *Arabidopsis thaliana* (Chiu et al., 2002), *Solanum lycopersicum* (Aouini et al., 2012), *Oryza sativa* (Singh et al., 2014), *Pyrus bretschneideri* (Chen et al., 2016), *Medicago truncatula* (Philippe et al., 2019), *Glycine max* (Zeng et al., 2020), *Gossypium hirsutum* (Liu et al., 2021), *Zea mays* (Zhou et al., 2021), *Saccharum* (Zhang et al., 2022), *Brassica rapa* (Yang et al., 2022), and *Erigeron breviscapus* (Yan et al., 2022), respectively. Furthermore, genetic research has demonstrated that *GLRs* play a crucial role in regulating various plant developmental processes and responding to environmental stresses such as salt, drought, heat, wounding, and pathogen attacks. These developmental processes encompass seed germination (Kong et al., 2015), root development (Vincill et al., 2013; Singh et al., 2016), hypocotyl elongation and pollen tube growth (Michard et al., 2011; Wudick et al., 2018b), xylem growth (Chen et al., 2016), Ca^{2+} distribution (Kim et al., 2001), stomatal closure (Cho et al., 2009), nitrogen and carbon metabolism (Kang et al., 2004), and abscisic acid (ABA) synthesis. Tea transcriptome data showed that the homologs for *GLR2.7* and *GLR2.8* manifest an upregulation in their expression levels under salt stress (Wan et al., 2018). Following coldstress, *AtGLR1.2* and *AtGLR1.3* activate core binding factors/dehydration-responsive element binding protein 1 signaling pathway through endogenous jasmonic acid (JA) accumulation, which contributes towards enhancing cold tolerance (Hu et al., 2013; Zheng et al., 2018). Moreover, *OsGLR3.4* is involved in brassinosteroid-mediated damage response from root to stem in rice. (Yu et al., 2022). GLRs can also improve the regeneration of plants after wounding and are crucial for establishing a balance between plant defense and regeneration following injury (Bian et al., 2022; Hernández-Coronado et al., 2022). As Ca^{2+} channels, plant GLRs can participate in disease-resistance responses by regulating Ca^{2+} signals. Kang et al. cloned an *RsGLuR* gene located on the plasma membrane from small radish, and found that overexpression of *RsGLuR* in *Arabidopsis* can improve resistance to the pathogen *Botrytis cinema* by triggering JA biosynthesis (Kang et al., 2006). However, there is no substantial volume of research on *GLRs* in sweet potato, in terms of their regulatory mechanisms and biological functions.

Sweet potato, *Ipomoea batatas* (L.) Lam. (2n = 6x = 90), is characterized by drought resistance, high and stable yields, strong adaptability, and rich nutrition, and is the best food recommended by the World Health Organization (Ma et al., 2012). It is both a food and cash crop with a wide range of uses. Specifically, it can be used as fresh food (Xie et al., 2018), starch processing (Zhou et al., 2020), food processing (Xiang et al., 2020), leafy vegetables (Su et al., 2018), and ornamental purposes (Meng and Lai, 2019). Accordingly, sweet potato has become an essential feed, food, and industrial raw material, and it is widely cultivated in over 100 regions and countries across the globe (Food and Agriculture Organization of the United Nations, 2021). The sweet potato genome is large and complex owing to its hexaploidy and high heterozygosity. In recent years, the assembly and reporting of the genomes for hexaploid sweet potato (Taizhong6) and its two diploid relatives (*Ipomoea trifida*, NCNSP0306, 2n=2x=30 and *Ipomoea triloba*, NCNSP0323, 2n=2x=30) have made it possible to analyze and identify the important gene family in sweet potato across the whole-genome level (Yang et al., 2017; Wu et al., 2018).

The cultivated hexaploid sweet potato and two of its diploid relatives (*I. trifida* and *I. triloba*) were used to screen and identify the *GLR* gene members in this study. The *GLR* genes were then analyzed for phylogenetic relationships, protein physicochemical properties, chromosomal localizations, gene structures, promoter *cis*-elements, protein interaction network, and expression patterns. The findings provide a basis for further understanding of the biological functions of *GLRs* and the future molecular breeding of sweet potatoes.

2 Materials and methods

2.1 Identification of *GLRs*

The whole-genome sequences for *I. batatas*, *I. trifida*, and *I. triloba* were obtained through *Ipomoea* Genome Hub (https://ipomoea-genome.org/) (viewed: 6 March 2023) and Sweetpotato Genomics Resource (http://sweetpotato.plantbiology.msu.edu/) (viewed: 6 March 2023). Three different screening techniques (Dai et al., 2022) were used at once to ensure that all members of *GLR* family were accurately identified.

2.2 Prediction for GLR protein properties

The molecular weight, hydrophilicity, instability index, and theoretical isoelectric point for GLRs were computed through the Expert Protein Analysis System (ExPASy, https://www.expasy.org/)

(viewed: 12 March 2023). Protein Subcellular Localization Prediction (PSORT, https://wolfpsort.hgc.jp/) (viewed: 12 March 2023) was employed to predict the subcellular localizations for GLRs.

2.3 Chromosomal distribution for *GLRs*

IbGLRs, *ItfGLRs*, and *ItbGLRs* were separately mapped to the respective chromosomes for *I. batatas*, *I. trifida*, and *I. triloba* by *Ipomoea* Genome Hub (https://ipomoea-genome.org/) (viewed: 14 March 2023) and Sweetpotato Genomics Resource (http://sweetpotato.plantbiology.msu.edu/) (viewed: 14 March 2023). The Toolkit for Biologists integrating various biological data handling tools (TBtools) program (South China Agricultural University, Guangzhou, China) was used for visualization.

2.4 Phylogenetic analysis for *GLRs*

MEGA ClustalW 7.0 was used to phylogenetically analyze the amino acid sequences for GLRs of *Arabidopsis*, *I. batatas*, *I. trifida*, and *I. triloba* (Kumar et al., 2016). Bootstrapping was carried out with 1000 replicates, and Interactive Tree of Life (iTOL) software (http://itol.embl.de/) (viewed: 8 June 2023) was used to construct a phylogenetic tree.

2.5 Domain identification and conserved motif analysis for GLRs

To perform an in-depth study of GLRs' conserved motifs, the Multiple Expectation Maximizations for Motif Elicitation (MEME) program (https://meme-suite.org/meme/) (viewed: 15 March 2023) was used. The maximum number of motifs that could be used has been set at five.

2.6 Exon-intron structures and promoter analysis for *GLRs*

A database of Plant Cis-Acting Regulatory Element (PlantCARE, (http://bioinformatics.psb.ugent.be/webtools/plantcare/html) (viewed: 16 March 2023) predicted *cis*-elements within 1500 bp promoter region for *GLRs* (Lescot et al., 2002). Gene Structure Display Server (GSDS) 2.0 developed by Peking University, Beijing, China (http://gsds.gao-lab.org/) (viewed: 16 March 2023) and the TBtools software developed by South China Agricultural University, Guangzhou, China was used to obtain and visualize the exon-intron structures for *GLRs*, respectively.

2.7 GLR protein interaction networking

Depending upon *Arabidopsis* orthologous proteins, the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING, https://www.string-db.org/) (viewed: 18 March 2023) predicted GLR protein interaction network. The Cytoscape version 3.2 was used to construct the network map (Kohl et al., 2011).

2.8 Data analysis

Transcriptome analysis and real-time quantitative PCR (qRT-PCR) was carried out at duration points (0 h, 36 h, 72 h, 120 h, and 10 d) following root rot induction using the underground stem of resistant variety Jishuzi203 and susceptible variety Jishuzi563. These plants were provided by the Institute of Cereal and Oil Crops, Hebei Academy of Agriculture and Forestry Sciences (Shijiazhuang, China) and planted in the natural root rot disease nursery in Xiong'an New Area, China. qRT-PCR was conducted on a CFX Opus 384 Real-Time PCR system (Bio-Rad, USA) by the ChamQ Universal SYBR qPCR Master Mix (Vazyme, China). The relative expression level of the target gene was presented as fold change compared with the internal control using the $2^{-\Delta ct}$ method, and data were analyzed with Duncan's multiple range test ($p < 0.05$). Three biological replications were performed for each test. A gene of *I. batatas* ADP-ribosylation factor (ARF, GenBank JX177359) was used as an internal control. The specific primers used for the qRT-PCR analysis were listed in Supplementary Table S2. Based on related investigations (Zhang et al., 2017; Dong et al., 2019; Zhu et al., 2019) ribonucleic nucleic acid sequencing (RNA-seq) data for *IbGLRs* were collected. RNA-seq data for *ItfGLRs* and *ItbGLRs* were collected through Sweetpotato Genomics Resource (http://sweetpotato.plantbiology.msu.edu/) (viewed: 22 March 2023). The fragments per kilobase of exon per million mapped fragments (FPKM) method was used to calculate the *GLRs* expression levels. Tbtools was used to build the heat maps.

3 Results

3.1 Characterization of *GLRs* in sweet potato and two diploid relatives

37 *GLRs* for sweet potato (named *Ib*) and 32 for each of two diploid relatives (named *Itf* and *Itb*) were identified through a combination of three methods used by Dai et al. (2022). The sequences from *I. batatas* were used for the analysis of the physicochemical features of *GLRs* (Table 1). The coding sequence length of *IbGLRs* varied from 1977 bp (*IbGLR7*) to 7326 bp (*IbGLR36*). The molecular weights of IbGLRs ranged from 73.627 to 270.105 kDa, the amino acid lengths extended from 658 to 2441 amino acids, and their isoelectric points varied from 5.49 (IbGLR32) to 9.31 (IbGLR30). Most IbGLRs were stable, only IbGLR2/10/11/13/14/15/21 and IbGLR30 were unstable (instability index > 40). 17 hydrophobic proteins (the grand average of hydropathy [GRAVY] score > 0) and 20 hydrophilic proteins (GRAVY score < 0) were identified for the IbGLR family. Subcellular localization prediction assessment demonstrated that except for IbGLR1 (located on chloroplast) and IbGLR12 (located on chloroplast and nucleus), the remaining IbGLRs were located on the plasma membrane.

10.3389/fpls.2023.1255805

TABLE 1 Characterization for *IbGLRs* in sweet potato.

Gene name	Gene ID	CDS (bp)	Phosphorylation site			Amino acids (aa)	MW (kDa)	pI	Instability index	GRAVY	Subcellular locations
			Ser	Thr	Tyr						
IbGLR1	g624.t1	2313	13	4	3	770	85.398	5.98	36.00	−0.026	Chloroplast
IbGLR2	g641.t1	2778	16	7	3	925	103.146	7.30	46.59	0.008	Plasma membrane
IbGLR3	g2074.t1	2790	6	0	4	929	102.309	8.41	31.27	−0.070	Plasma membrane
IbGLR4	g10173.t1	2649	8	10	1	882	98.964	6.53	32.61	0.072	Plasma membrane
IbGLR5	g10174.t1	2460	21	11	6	819	92.102	8.04	36.39	0.036	Plasma membrane
IbGLR6	g10180.t1	2040	7	3	0	679	75.862	6.89	33.80	0.063	Plasma membrane
IbGLR7	g10181.t1	1977	5	0	1	658	73.627	6.43	33.86	0.040	Plasma membrane
IbGLR8	g16777.t1	2622	13	4	6	873	97.194	6.75	37.85	−0.015	Plasma membrane
IbGLR9	g29862.t1	2673	0	0	0	890	99.212	6.59	35.46	0.142	Plasma membrane
IbGLR10	g29863.t1	2562	9	13	0	853	95.261	6.30	41.75	0.187	Plasma membrane
IbGLR11	g29864.t1	2484	0	0	0	827	91.966	5.97	40.83	0.227	Plasma membrane
IbGLR12	g29866.t1	2316	2	2	1	771	86.103	8.25	35.63	0.019	Chloroplast and nucleus
IbGLR13	g29867.t1	2823	0	0	0	940	104.895	6.31	41.31	0.021	Plasma membrane
IbGLR14	g29868.t1	3507	4	0	3	1168	130.178	8.18	41.71	0.072	Plasma membrane
IbGLR15	g29869.t1	2811	13	5	6	936	104.367	5.58	42.27	−0.003	Plasma membrane
IbGLR16	g30933.t1	2532	11	5	1	843	93.383	6.15	38.76	−0.012	Plasma membrane
IbGLR17	g30934.t1	2922	0	0	0	973	108.564	6.44	33.24	0.010	Plasma membrane
IbGLR18	g34576.t1	3129	6	6	0	1042	115.899	7.37	34.06	−0.039	Plasma membrane
IbGLR19	g34577.t1	4431	0	0	0	1476	166.657	7.05	38.23	−0.116	Plasma membrane
IbGLR20	g37769.t1	2802	0	0	0	933	103.315	5.52	39.79	0.020	Plasma membrane
IbGLR21	g37945.t1	2643	0	0	0	880	97.320	6.79	41.35	0.090	Plasma membrane
IbGLR22	g49894.t1	3210	7	7	0	1069	119.444	9.08	37.40	−0.088	Plasma membrane
IbGLR23	g49895.t1	2331	12	10	2	776	86.988	6.51	34.15	−0.033	Plasma membrane
IbGLR24	g49902.t1	2601	22	11	7	866	96.941	7.94	36.89	−0.017	Plasma membrane
IbGLR25	g49903.t1	2652	0	1	4	883	98.686	6.12	32.15	−0.009	Plasma membrane
IbGLR26	g53987.t1	2652	0	0	0	883	98.219	6.07	34.81	−0.005	Plasma membrane
IbGLR27	g54255.t1	2541	2	1	2	846	93.832	8.32	35.25	−0.111	Plasma membrane
IbGLR28	g54256.t1	2631	11	13	0	876	97.683	8.88	33.69	−0.101	Plasma membrane
IbGLR29	g54698.t1	2802	21	7	6	933	104.088	8.72	32.90	0.059	Plasma membrane
IbGLR30	g54783.t1	3960	23	18	4	1319	149.363	9.31	41.56	−0.220	Plasma membrane
IbGLR31	g55118.t1	2628	0	0	0	875	97.429	6.44	33.13	0.051	Plasma membrane
IbGLR32	g55121.t1	2373	4	3	0	790	87.021	5.49	36.87	−0.018	Plasma membrane
IbGLR33	g55122.t1	2844	0	0	0	947	104.829	5.54	33.57	−0.008	Plasma membrane
IbGLR34	g55123.t1	2661	0	0	0	886	97.568	5.53	37.38	0.048	Plasma membrane
IbGLR35	g55125.t1	2622	0	0	0	873	96.299	8.51	33.15	−0.032	Plasma membrane
IbGLR36	g55127.t1	7326	0	0	9	2441	270.105	7.07	34.79	−0.033	Plasma membrane
IbGLR37	g60932.t1	2886	0	0	0	961	106.999	5.63	38.12	−0.050	Plasma membrane

CDS, coding sequence; Ser, serine; Thr, threonine; Tyr, tyrosine; MW, molecular weight; pI, isoelectric point.

10.3389/fpls.2023.1255805

Based on the physical locations of genes in the *I. batatas*, *I. trifida*, and *I. triloba* genomes, the chromosomal positions of *GLRs* in these crops were mapped. In *I. batatas*, *IbGLR*s were distributed unevenly on nine chromosomes. LG13 had the most *IbGLR* genes, with 11, followed by LG7 with seven. LG3, LG9, and LG12 had four *IbGLRs*. There were less than four *IbGLRs* on LG1, LG5, LG8, and LG15 (Figure 1A). In *I. trifida*, except for *ItfGLR31* and *ItfGLR32* located on Chr00 (not shown in Figure 1B), the other *ItfGLRs* were located on the same chromosomes (Chr02, Chr03, Chr05, Chr06, Chr07, Chr10, Chr11, Chr12, and Chr14) compared to those in *I. triloba*. Chr02 had the most *ItfGLR* genes, with 12, followed by Chr14 with four. There were less than four *ItfGLRs* on Chr03, Chr05, Chr06, Chr07, Chr10, Chr11, and Chr12 (Figure 1B). In *I. triloba*, Chr02 had the most *ItbGLR* genes, with 11, followed by Chr03 with six and Chr14 with five. There were less than four *ItbGLRs* on Chr05, Chr06 Chr07, Chr10, Chr11, and Chr12 (Figure 1C). Chr06, Chr10, Chr11, and Chr12 had the same numbers of *GLRs* in *I. trifida* and *I. triloba*. These findings imply that two diploid relatives differed in the proportion and distribution of *GLRs* on chromosomes in comparison to sweet potato.

3.2 Phylogenetic relationship of *GLRs* for sweet potato and two diploid relatives

Phylogenetic trees were constructed to elucidate the evolutionary interactions between 121 different GLRs from *Arabidopsis* (20), *I. batatas* (37), *I. trifida* (32), and *I. triloba* (32) (Figure 2). The GLRs for different species were categorized according to evolutionary distances into different subgroups, including five (Groups I to V) of *I. batatas*, four (Groups I to III, and V) of *I. trifida*, five (Groups I to V) of *I. triloba*, and three (Groups I to III) of *Arabidopsis*. The specific distributions for GLRs were as follows (total: *I. batatas*, *I. trifida*, *I. triloba*, *Arabidopsis*):

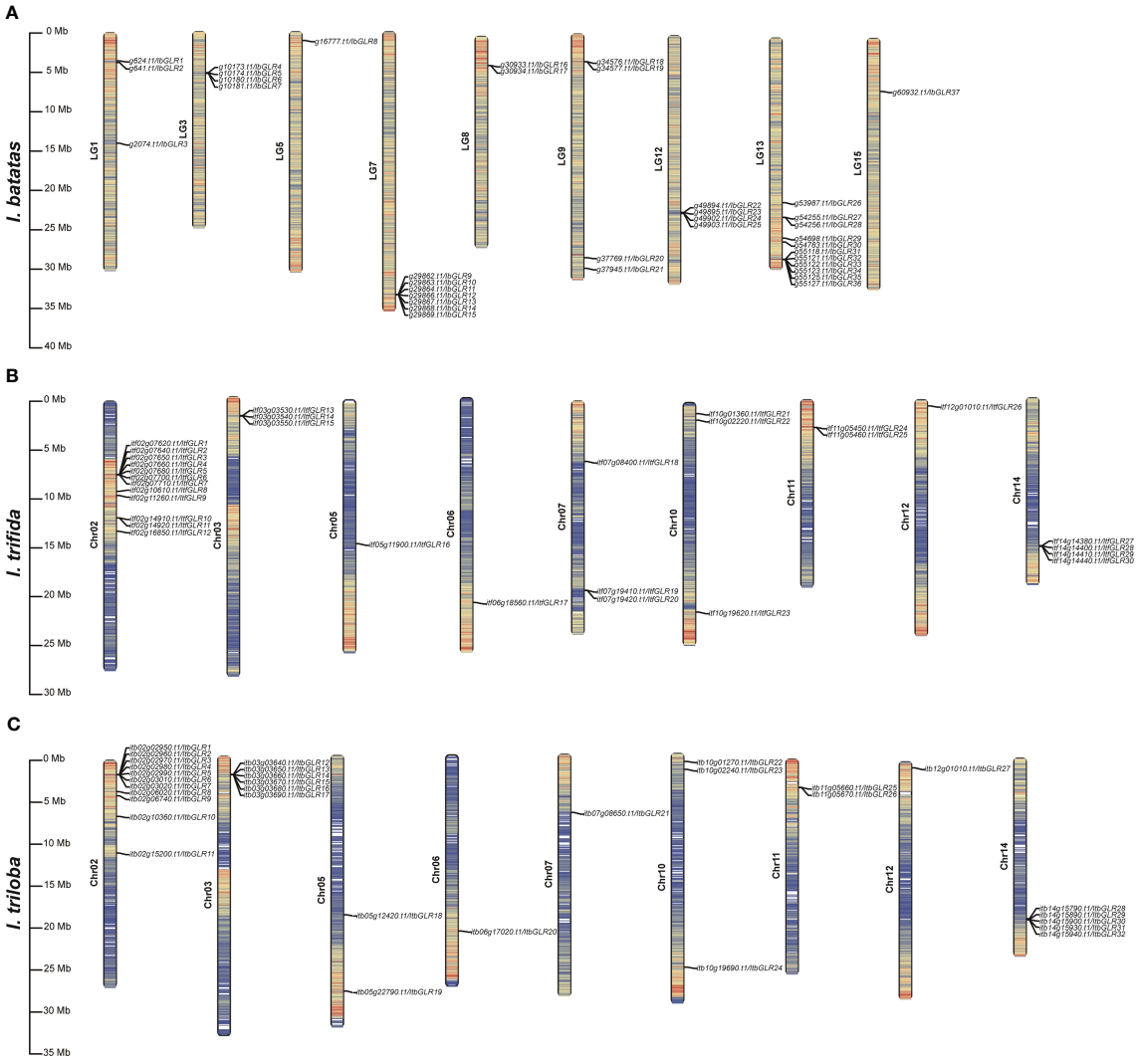


FIGURE 1
Chromosomal distribution and localization information of *GLR* genes in *Ipomoea batatas* **(A)**, *Ipomoea trifida* **(B)**, and *Ipomoea triloba* **(C)**. The bar chart represents chromosomes, with chromosome numbers on the left side and gene names on the right side for bar chart. The black line on the right side for bar chart marks the position of each *GLR* gene on the chromosome and represents it in units of Mbp.

Group I (37: 11, 12, 10, 4), Group II (27: 6, 6, 6, 9), Group III (47: 16, 10, 14, 7), Group IV (2: 1, 0, 1, 0), and Group V (8: 3, 4, 1, 0) (Figure 2; Supplementary Table S1). All AtGLRs were found to have homologous proteins in *I. batatas*, *I. trifida*, and *I. triloba*, but IbGLR18/19/27/28, ItfGLR10/11/23/32, and ItbGLR10/24 showed no homology with *Arabidopsis* GLRs. Except for Groups II and IV, the numbers and types of GLRs distributed in the *I. batatas* subgroups were different from those in *I. trifida*, *I. triloba*, and *Arabidopsis*.

3.3 GLR conserved motifs and exon–intron structural assessments for sweet potato and two diploid relatives

The MEME website was used to analyze the sequence motifs of 37 *IbGLRs*, 32 *ItfGLRs*, and 32 *ItbGLRs*, and motif 1 to 5 were identified as five conserved motifs (Figure 3A). All GLR sequences were used to produce the sequences for the five most conserved motifs, shared between sweet potato and two diploid relatives (Supplementary Figure S1). Despite being substantially similar, the GLRs for *I. batata*, *I. trifida* together with *I. triloba* could have different numbers and types of conserved domains. For example, ItfGLR1 contained motifs 2 to 5, ItbGLR1 contained motifs 1 to 5, and IbGLR31 contained motifs 1 to 5. ItfGLR4 contained motifs 1 to 5, ItbGLR4 contained motifs 1 to 5, and IbGLR34 contained motifs 1 to 3 and 5. ItfGLR8 comprised motifs 1 to 5, ItbGLR8 contained motifs 3 to 5, and IbGLR30 contained motifs 1 to 5 (Figure 3A). Additionally, the receptor domains (atrial natriuretic factor [ANF]_receptor and periplasmic_binding_protein) and four transmembrane helix domains (Lig_chan domains) of GLRs are closely related to plant functions. Most GLRs of *I. batata*, *I. trifida*, and *I. triloba* (30 IbGLRs, 28 ItfGLRs, and 29 ItbGLRs) contained ANF_receptor and Lig_chan domains. However, ItbGLR31 only contained a Lig_chan domain. IbGLR7/18/23/28, ItfGLR32, and ItbGLR24 contained

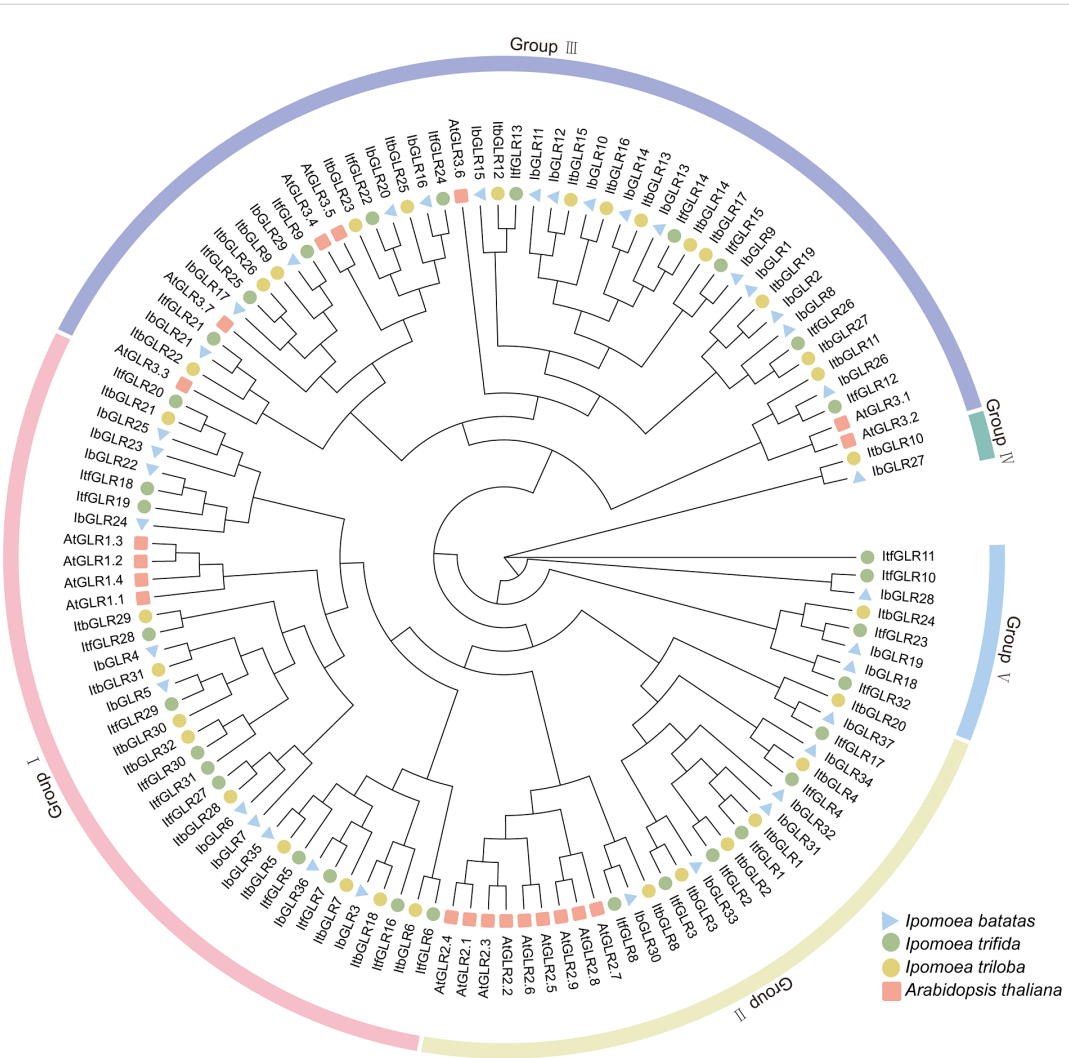


FIGURE 2
Phylogenetic relationship analysis of GLRs in *Arabidopsis*, *Ipomoea batatas*, *Ipomoea trifida*, and *Ipomoea triloba*. The 20 AtGLRs for *Arabidopsis* are represented by red squares. The 37 IbGLRs for *I.batatas* are represented by blue triangles. The 32 ItfGLRs for *I. trifida* are represented by green circles. The 32 ItbGLRs in *I. triloba* are represented by yellow circles.

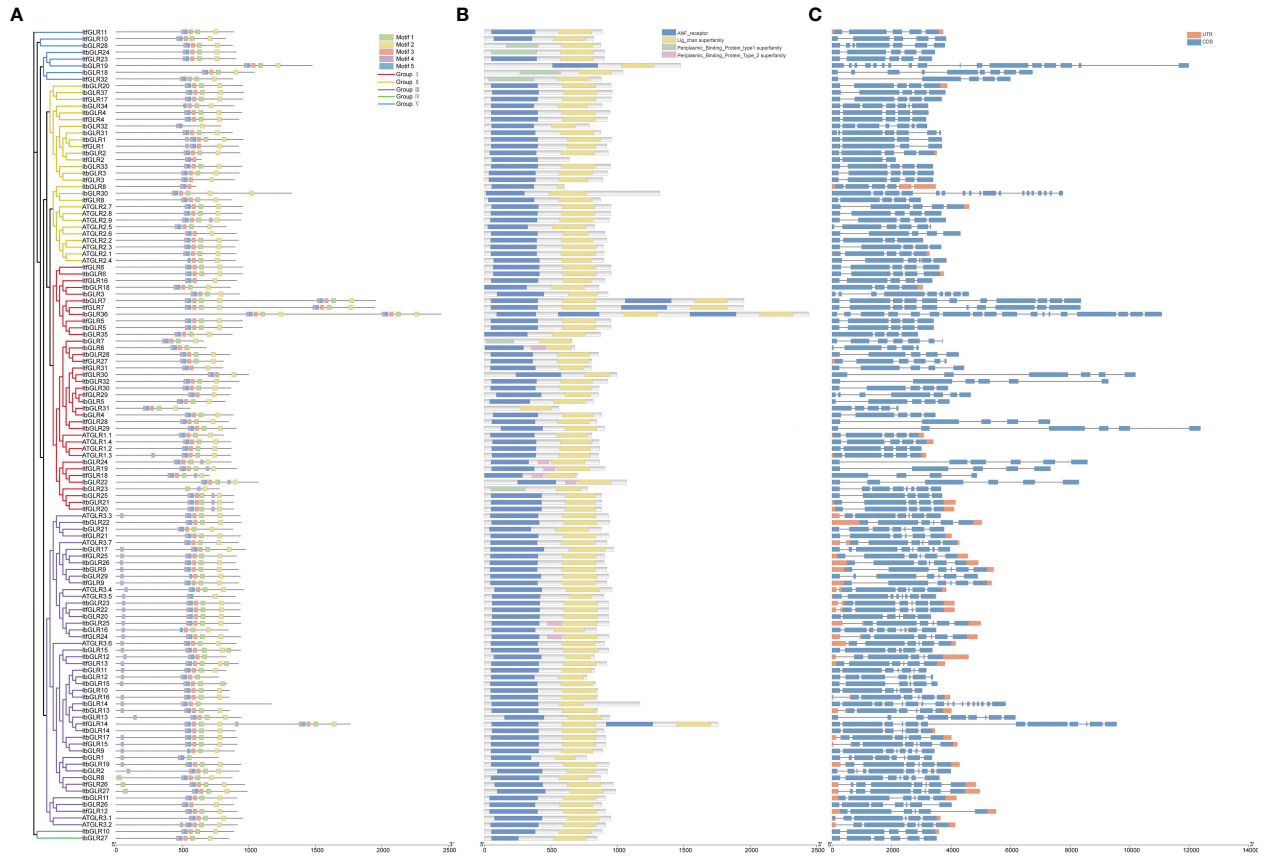


FIGURE 3
The exon–intron structural and conserved motifs analysis of GLR family in *Ipomoea batatas*, *Ipomoea trifida*, and *Ipomoea triloba*. **(A)** GLRs were distributed in five subgroups by phylogenetic tree (left), and motif 1-5 are shown in different colors (right). **(B)** The GLRs conserved domain structures. The blue boxes, yellow boxes, green boxes, and pink boxes represent the ANF_receptor domain, Lig_chan domain, periplasmic_binding_protein_type1 domain and periplasmic_binding_protein_type2 domain, respectively. **(C)** The GLRs exon-intron structures. The black lines, blue boxes, and orange boxes represent the introns, exons, and UTRs, respectively.

periplasmic_binding_protein_type1 and Lig_chan domains, and IbGLR6/22/24, ItfGLR18/19/24, and ItbGLR25 contained ANF_receptor, periplasmic_binding_protein_type2, and Lig_chan domains (Figure 3B).

Exon-intron architectures were examined in order to gain a better understanding of structural variation among the *GLRs* (Figure 3C). The number of exons in the GLR genes varied from 3 to 18. In detail, *GLRs* of Group I contained 4 to 18 exons, *GLRs* of Group II contained 3 to 18 exons, *GLRs* of Group III contained 6 to 17 exons, *GLRs* of Group IV contained 5 to 6 exons, and *GLRs* of Group V carried 5 to 17 exons (Figure 3C). Additionally, exon-intron architecture for several homologous *GLRs* in *I. batatas* was identified as possibly different from the counterparts in *I. trifida* and *I. triloba*. Such as, *IbGLR36* carried 18 exons, but *ItfGLR7* and *ItbGLR7*, the corresponding homologous genes of *IbGLR36* contained 10 and 11 exons in Group I, *IbGLR30* contained 18 exons, but *ItfGLR8* and *ItbGLR8* both contained five exons in Group II, and *IbGLR13* carried nine exons, but *ItfGLR14* and *ItbGLR14* contained 12 and six exons in Group III (Figure 3C). These findings suggest that sweet potato genome potentially underwent a lineage-specific differentiation event involving *GLR* gene family members.

3.4 *Cis*-elements assessment for *IbGLR* promoters in sweet potato

A *cis*-element, such as the sequence upstream of GLRs, could play a significant role in plant development and stress responses. Hence, we used a 1500 bp DNA sequence upstream of *IbGLRs* for *cis*-element analysis in *I. batatas*. According to the predicted functions, such components were separated into five groups (core, hormone, developmental, light, and abiotic/biotic) (Figure 4). The CAAT-box and TATA-box core promoter elements were present in all 37 *IbGLRs*. There were 19 to 47 CAAT-box and 17 to 129 TATA-box core promoter elements in 37 *IbGLRs*. In addition to *IbGLR35*, other *IbGLRs* were found to have several hormone elements, such as the P-box for gibberellic acid (GA)-responsive elements, TCA for SA-responding elements, AuxRR-core together with TGA-element for indole-3-acetic acid (IAA)-responsive elements, CGTCA-motif and TGACG-motif for MeJA-responsive elements/ABRE for ABA-responsive elements (Figure 4).

For development-related elements, the CAT-box related to meristem formation (found in *IbGLR1/4/8/10/11/18/21/26/27/30/36*), O2-site, a zein metabolism-regulatory element (observed in

10.3389/fpls.2023.1255805

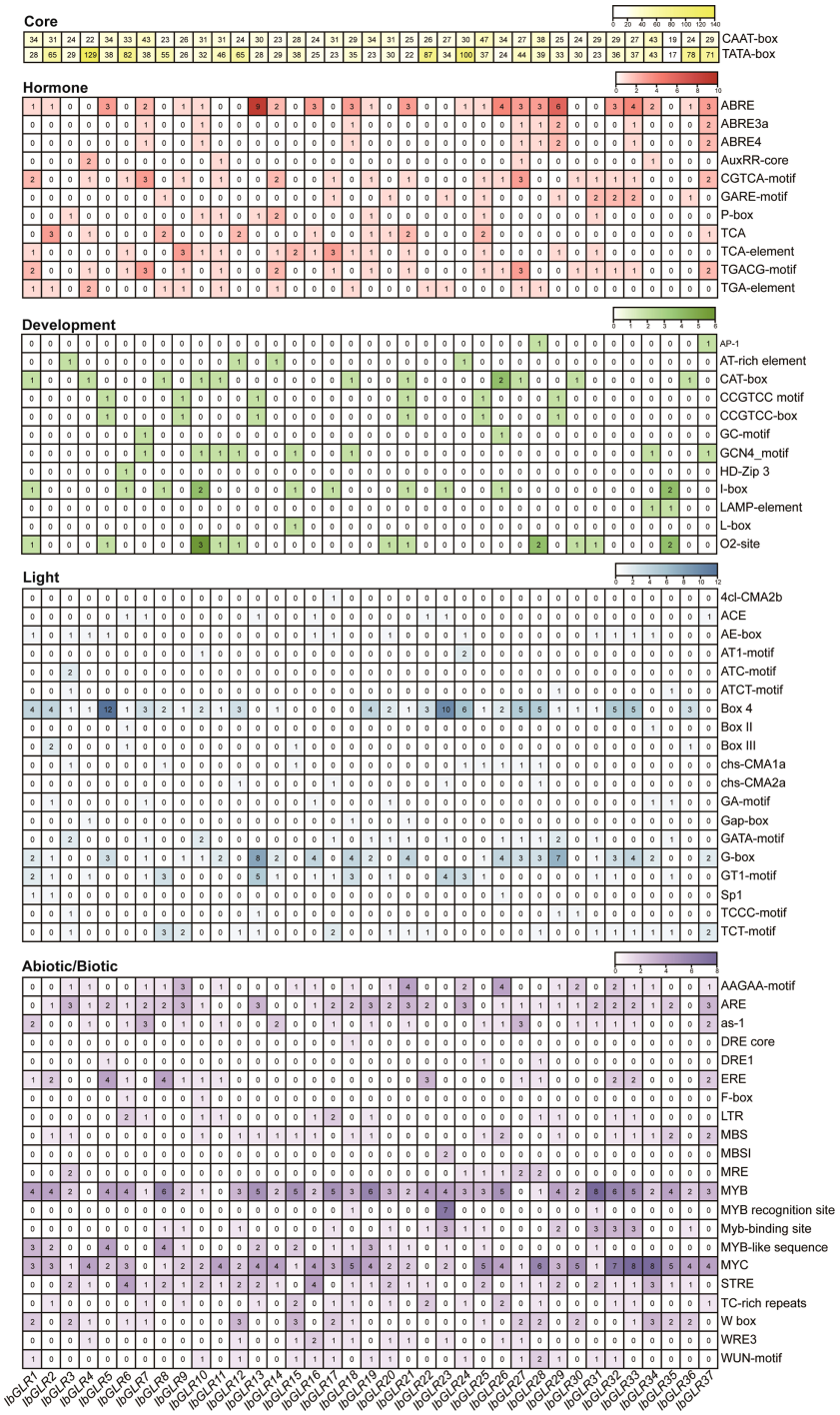


FIGURE 4
The *cis*-element analysis of *IbGLR*s in *Ipomoea batatas*. Yellow, red, green, blue, and purple represents core elements, hormone elements, development elements, light elements, and abiotic/biotic elements, respectively. For the same color, a darker color indicates a greater number for *cis*-elements.

IbGLR1/5/10/11/12/20/21/28/30/31/35), I-box responding to light (found in *IbGLR1/6/8/10/15/17/21/23/26/35*), and GCN4_motif participating in regulating seed-specific expression (found in *IbGLR7/10/11/12/15/18/34/37*) were abundant in promoters of *IbGLRs*. The promoter regions of *IbGLR2/16/19/22/32/33* did not contain any development-related elements (Figure 4). Similarly, the promoters of 37 *IbGLRs* contained a number of light-responsive elements. *IbGLRs* were found to be rich in BOX4, G-boxes, GT1-motifs, and TCT-motifs. The ATC-motif was only present in *IbGLR3*, and 4cl-CMA2b only in *IbGLR17* (Figure 4).

Moreover, several abiotic elements, the MYC, MYB, and DRE core responded to drought stress, LTR, MBS, and W box responded to salt stress and biotic elements, the WUN, WRE3, and W box motifs were found in most *IbGLRs* (Figure 4). This indicated that *IbGLRs* may have different roles under different stress conditions. The DRE core was only found in the promoter region of *IbGLR18*.

The results demonstrate that *IbGLRs* take part in regulating the growth and development of plants, hormone crosstalk, and adaptation to abiotic and biotic stress in sweet potato.

3.5 Protein interaction networking for IbGLRs within sweet potato

An IbGLR interaction network was built through orthologous proteins from *Arabidopsis* to examine possible regulatory networking for IbGLRs (Figure 5). The prediction of protein interactions suggested that some IbGLRs (IbGLR2/5/7/12/18/20/25) could interact with other IbGLRs. Moreover, IbGLR7, IbGLR17, and IbGLR20 were determined to interact with CNGC18, which was reported to regulate germination of pollen grains and the growth of pollen tubes (Chang et al., 2007; Cao et al., 2016; Gu et al., 2017). IbGLR17, IbGLR20, and IbGLR25 were determined to interact with GF14, which participates in transport, growth, metabolism, and the drought stress reactions (Li et al., 2016; Hajibarat et al., 2022). IbGLR2, IbGLR7, and IbGLR17 were determined to interact with TPK1, which participates in potassium transportation, fruit ripening, and quality formation (Lu, 2011; Latza et al., 2013; Wang et al., 2018). IbGLR2/3/12/20 were determined to interact with TPC1, which mediates Ca^{2+} release (Moccia et al., 2021; Navarro-Retamal et al., 2021; Pottosin and Dobrovinskaya, 2022), and IbGLR2/3/5/6/7/12/17/18/20 and IbGLR25 were determined to interact with AGB1, which regulates salt stress tolerance, fungal and bacterial immunity, and plant growth (Takahashi et al., 2018; Yu and Assmann, 2018; Zhang et al., 2020; Qi et al., 2021; Afrin et al., 2022). These findings indicate that IbGLRs have a distinct function in mediating biotic and abiotic stress and influence plant development in sweet potatoes.

3.6 *GLR* expression for sweet potato and two diploid relatives

3.6.1 Tissue-specific expression assessment

IbGLR expression levels were examined using RNA-seq data in four typical *I. batatas* tissues (leaves, stems, fibrous roots, and tuberous roots), in order to investigate their potential biological functions in plant developmental processes (Figure 6A).

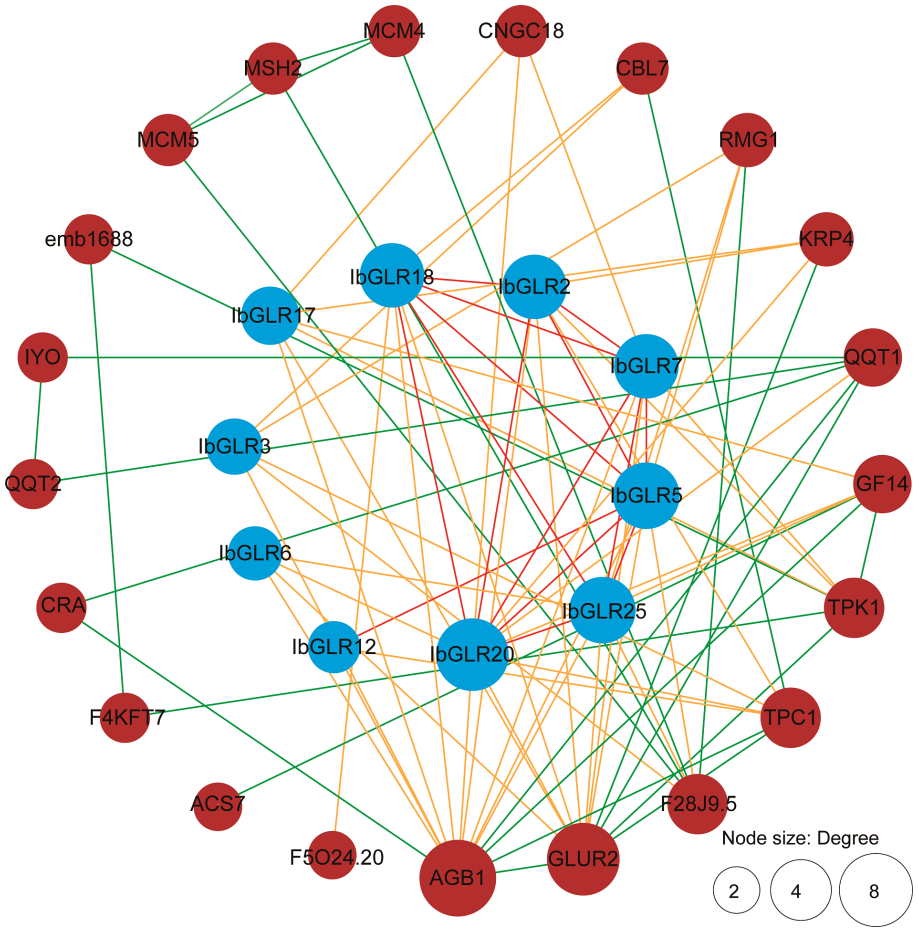


FIGURE 5
IbGLR functional interaction network analysis in *Ipomoea batatas*. The network nodes represent proteins. The number of interacting proteins is represented by node size. The lines represent the interactions between proteins. The interactions between different GLRs are represented by red lines. The interactions between GLRs and other proteins are represented by orange lines. The interactions between proteins other than GLR are represented by green lines.

Interestingly, different *IbGLRs* had distinct patterns of expression in the four tissues, with some *IbGLRs* showing tissue-specific expression. *IbGLR6/13/26/32* and *IbGLR35* had a high expression in leaves, whereas the expression of *IbGLR1/2* and *IbGLR4* was high in the stems. *IbGLR8/12/18/19/21/23/25/28/33/34/36* and *IbGLR37* highly expressed in fibrous roots, *IbGLR29* and *IbGLR31* highly expressed in tuberous roots, *IbGLR3* and *IbGLR27* highly expressed in leaves and fibrous roots, *IbGLR11/14/15/16/17/20/22* and *IbGLR24* highly expressed in leaves and stems, and *IbGLR10* and *IbGLR30* highly expressed in stems and fibrous roots. These findings imply that *IbGLRs* may have different roles in sweet potato tissue development.

Additionally, the expression patterns of *ItfGLRs* and *ItbGLRs* were studied *via* re-analyzing RNA-seq data published in six tissues (flower buds, flowers, leaves, stems, root 1, and root 2) (Wu et al., 2018) (Figures 6B, C). In *I. trifida*, *ItfGLR9* was highly expressed in leaves (FPKM value = 15.50) and stems (FPKM value = 14.12). In *I. triloba*, *ItbGLR8*, and *ItbGLR27* were only expressed in flowers and leaves, respectively, whereas the other *ItbGLRs* were expressed in different tissues. *ItbGLR9* exhibited a high expression in leaves. *ItbGLR22* was highly expressed in root 1 and root 2. While *ItbGLR25* was highly expressed in stems and flower buds, *ItbGLR26* exhibited a high expression in leaves and flower buds. These findings suggest that different *GLR* genes exhibit diverse growth regulatory roles in sweet potato and two diploid relatives, as well as different tissue expression patterns.

3.6.2 Expression assessment for *GLRs* during various stages of root development

The expression patterns for *IbGLRs* in the roots of Xushu22 plants were studied at five developmental stages using RNA-seq data (Dong et al., 2019). The results indicated that among the 37 *IbGLRs*, 22 *IbGLRs* shared similar expression patterns, with a higher transcriptional level in fibrous roots (diameter of approximately 1 mm) than roots from other growth stages. Within this group, *IbGLR21* had the highest level of expression (FPKM value = 16.38) (Figure 7). The expression of *IbGLR29* was the highest (FPKM value = 31.83) in the initial tuberous root (diameter of approximately 1 cm), and the expression of *IbGLR2* was the highest (FPKM value = 6.51) in tuberous roots (diameter of

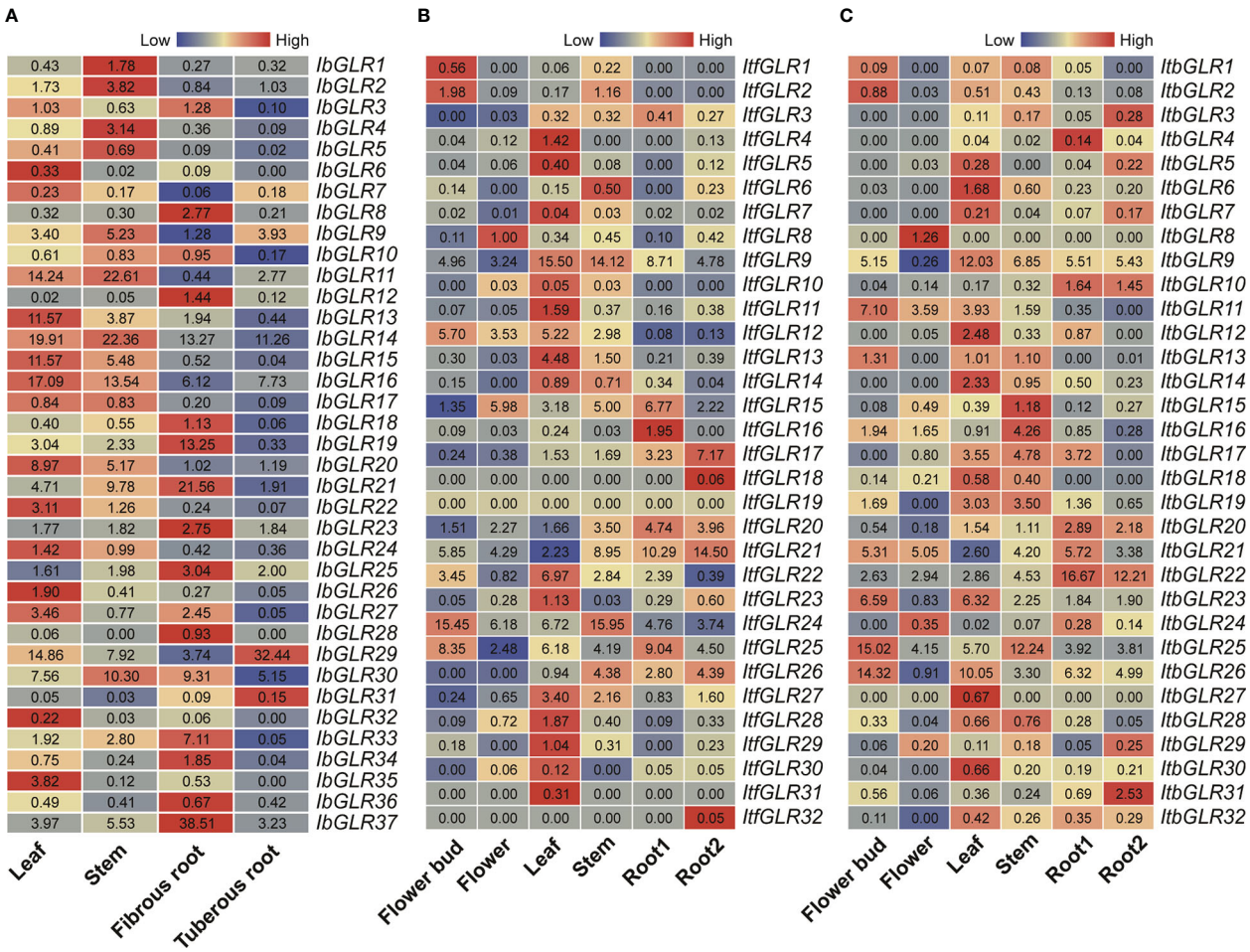


FIGURE 6
GLR expression study in different tissues for different cultivars. **(A)** Represents expression analysis for *Ipomoea batatas* based on the leaf, stem, fibrous root, and tuberous root. **(B, C)** Represent gene expression patterns for *Ipomoea trifida* and *Ipomoea triloba* based on the flower bud, flower, leaf, stem, root 1, and root 2. The FPKM values are displayed in the boxes. The color bar only represents the expression values of one *GLR* gene in different tissues.

approximately 3 cm). Meanwhile, the expression of *IbGLR14* was highest (FPKM value = 12.69) in tuberous roots (diameter of approximately 5 cm), and the expression of *IbGLR16* was the highest (FPKM value = 11.31) in tuberous roots (diameter of approximately 10 cm), which was consistent with the tuberous root of approximately 5 cm in diameter. Overall, the expression patterns imply different contributions of *IbGLRs* in the root development of sweet potato.

3.6.3 Expression assessment for *GLRs* in resistant and susceptible varieties under root rot treatment conditions

To study the possible role of *IbGLRs* during interactions with the sweet potato root rot pathogen, we analyzed the expression levels of *IbGLRs* at different time points after root rot induction in Jishuzi563 (root rot-sensitive variety) and Jishuzi203 (root rot-resistant variety) using RNA-seq data (Figure 8). Subsequently, we further evaluated the expression levels of some *IbGLRs* using qRT-PCR (Supplementary Figure S2). Without root rot infection, the expression of *IbGLR10/11/12/16/18* was repressed at 36 h, 72 h, 120 h, and 10 d compared with that at 0 h in Jishuzi563 and Jishuzi203. After root rot infection, only *IbGLR18* expression was repressed in Jishuzi563 infected by the root rot but was induced in Jishuzi203 infected by the root rot. This result was consistent with qRT-PCR results. It has been speculated that *IbGLR18* may participate in sweet potato resistance to root rot. In addition, the RNA-seq results showed that *IbGLR14/26/37* expression levels were upregulated in Jishuzi563 and Jishuzi203 in response to root rot, and the expression levels in the resistant variety Jishuzi203 were higher than those in susceptible variety Jishuzi563, which was consistent with real-time quantitative results. qRT-PCR results showed that *IbGLR14* manifested the highest level of expression at 120 h, *IbGLR26* at 72 h, and *IbGLR37* at 10 d. The outcome suggests that different *IbGLRs* may function differently in developing disease resistance at different stages during the interaction between sweet potato and root rot pathogens.

3.6.4 Expression assessment for *GLRs* in response to salt and drought stresses in hexaploid sweet potato and two diploid relatives

To better understand the capacity of *IbGLRs* to combat abiotic stress, RNA-seq data from a salt-tolerant line (ND98) and a salt-sensitive variety (Lizixiang) were used to compare the expression patterns of *IbGLRs* under salt stress. Meanwhile, the RNA-seq data from Xu55-2, a drought-resistant cultivar, was studied under drought stress (Zhang et al., 2017; Zhu et al., 2019). Salt stress enhanced the expression of 13 *IbGLRs* in the ND98 line (the expression for *IbGLR1/11/14/19/24/33/34* was highest at 12 h, while that of *IbGLR10/18/22/30/35/37* was highest at 48 h). For these genes, *IbGLR19* expression in salt sensitive variety Lizixiang was repressed by salt stress (Figure 9A). In Xu55-2, the expression of 26 *IbGLRs* was induced rapidly following drought stress and reached its maximum at the early stage (≤3 h). Meanwhile, *IbGLR21* expression was repressed by polyethylene glycol (PEG) treatment, whereas *IbGLR9/10/13/15/19/29/31/32* expression levels were induced (Figure 9B). Thus, it may be speculated that *IbGLR19* is related to salt and drought tolerance in sweet potato.

Additionally, the evaluation of *ItfGLRs* and *ItbGLRs* expression in *I. trifida* and *I. triloba* following drought and salt treatment was conducted using RNA-seq data (Wu et al., 2018). Expression levels of *ItfGLR1/2/18/31* were not induced, whereas those for *ItfGLR3/23/24* and *ItbGLR3/8/10/11/21/31* were induced (Figure 10). These results indicate that there are differences in the expression of *GLRs* in sweet potato and two diploid relatives under drought and salt stress.

3.6.5 Comparative investigation of *ItfGLRs* and *ItbGLRs* expression in response to hormone and temperature stress

Ultimately, an assessment of expression patterns for *ItfGLRs* and *ItbGLRs* under ABA, GA, and IAA treatments was carried out through RNA-seq data from *I. trifida* and *I. triloba* (Wu et al., 2018). In *I. trifida*, the expression of *ItfGLR18* and *ItfGLR31* was not induced by any of the hormones, whereas the expression levels of five *ItfGLRs* were induced by ABA, those of 11 *ItfGLRs* were induced by GA, and those of two *ItfGLRs* were induced by IAA. *ItfGLR6* and *ItfGLR30* expression was induced by all hormones, and *ItfGLR1/5/7/8/9/11/13/15/16/20/23/25* and *ItfGLR27* expression was suppressed by the three hormones (Figure 11A). In the case of *I. triloba*, expression levels of 15 *ItbGLRs* were induced by ABA, those of four *ItbGLRs* were induced by GA, and those of five *ItbGLRs* were induced by IAA. *ItbGLR18* and *ItbGLR19* expression was induced by all hormones, whereas *ItbGLR5/6/7/13/23/25/26* and *ItbGLR32* expression was repressed by all hormones (Figure 11B). The

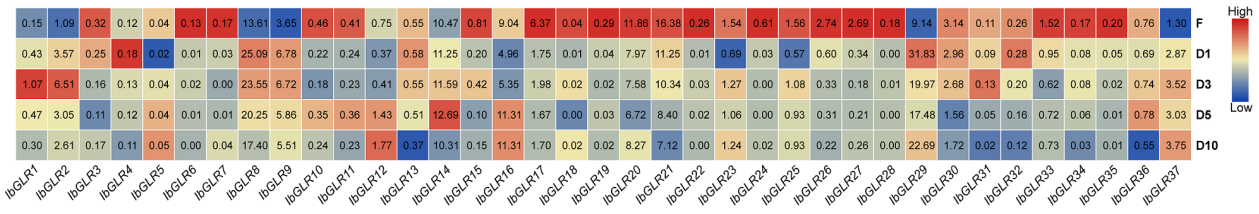


FIGURE 7
The expression analysis of *IbGLRs* at different stages of root development. F, D1, D3, D5 and D10 represent fibrous root (diameter of approximately 1 mm), initial tuberous root (diameter of approximately 1 cm), tuberous root (diameter of approximately 3 cm), tuberous root (diameter of approximately 5 cm), and tuberous root (diameter of approximately 10 cm), respectively. The FPKM values are displayed in the boxes. The color bar only represents the expression values of one *GLR* gene in different stages of root development.

10.3389/fpls.2023.1255805

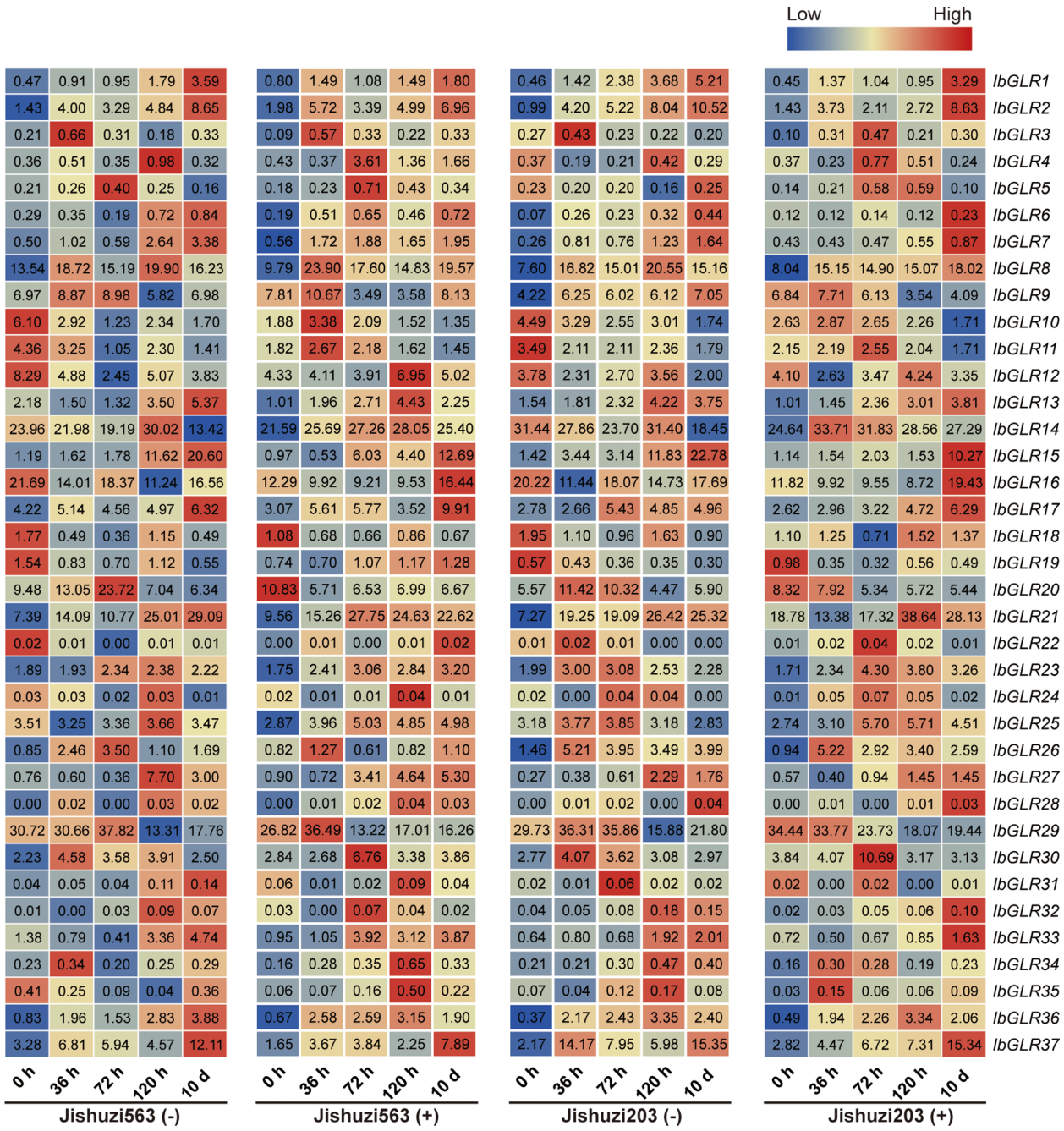


FIGURE 8
IbGLRs expression analysis in response to root rot. Jishuzi563(-): a root rot sensitive variety without root rot; Jishuzi563(+): a root rot sensitive variety with root rot; Jishuzi203(-): a root rot resistant variety without root rot; Jishuzi203(+): a root rot resistant variety with root rot. The FPKM values are displayed in the boxes. The color bar only represents the expression values of one *GLR* gene at different times after root rot induction in different varieties.

findings suggest that *GLRs* in the two diploid relatives participate in various hormonal pathways and take part in the crosstalk among different hormones.

Additionally, the pattern of expression for *ItfGLRs* and *ItbGLRs* was analyzed utilizing *I. trifida* and *I. triloba* RNA-seq data, respectively, at 10/4°C (day/night) and 35/35°C (day/night) treatment (Wu et al., 2018). In *I. trifida*, *ItfGLR4/8/17/23/27* and *ItfGLR28* expression was induced by 10/4°C (day/night), whereas *ItfGLR16* expression was induced by 35/35°C (day/night) temperatures comparison with the control levels (Figure 12A). In *I. triloba*, *ItbGLR3/8/20/24/29* and *ItbGLR32* expression was induced by 10/4°C (day/night) and *ItfGLR31* expression was induced by 35/35°C (day/night) compared with control levels temperatures (Figure 12B). These results show that different *GLRs* participate in the responses of *I. trifida* and *I. triloba* to different temperature stresses.

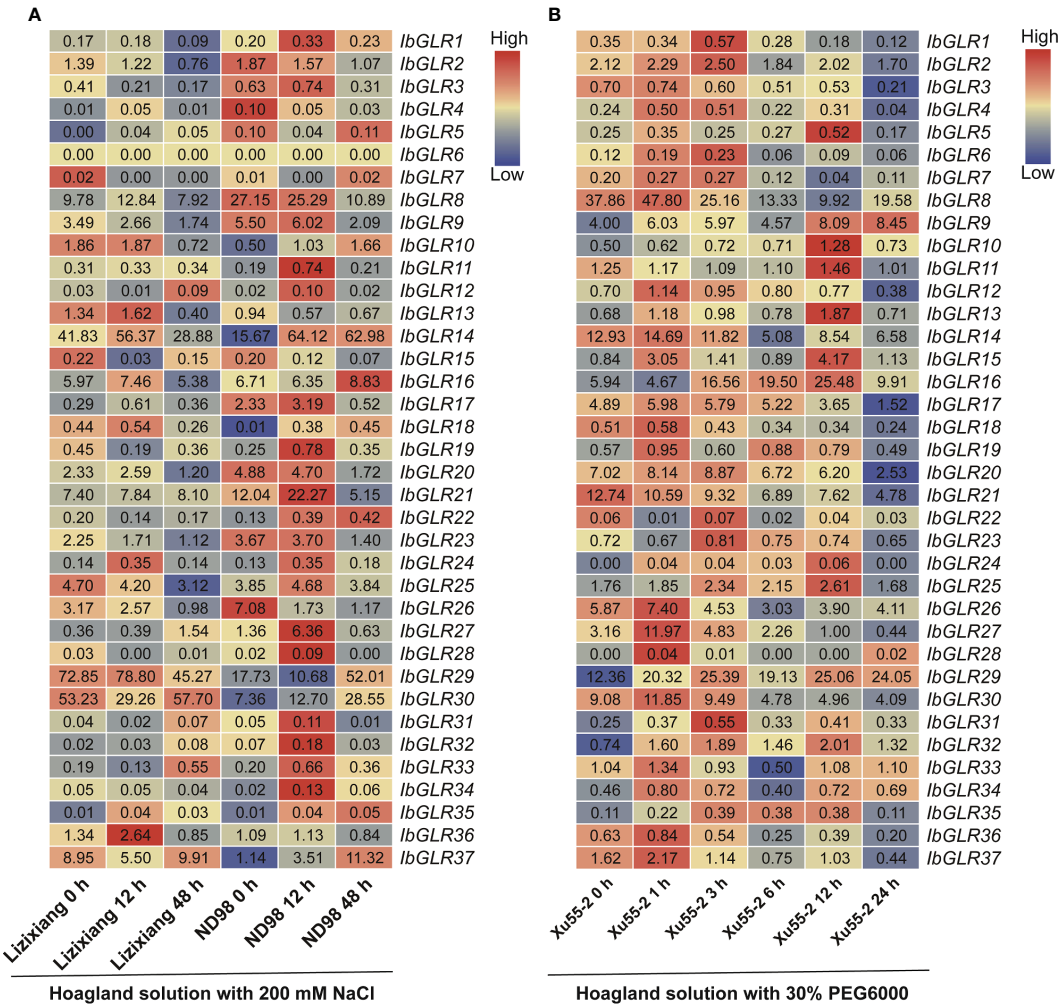


FIGURE 9
IbGLRs expression analysis in response to NaCl and PEG treatments. **(A)** Expression analysis for *IbGLRs* under NaCl treatment conditions in salt-sensitive variety Lizixiang and salt-tolerant line ND98. **(B)** Expression analysis for *IbGLRs* under PEG treatment conditions in drought-tolerant variety Xu55-2. The FPKM values are displayed in the boxes. The color bar only represents the expression values of one *IbGLR* in different treatments.

4 Discussion

Plant GLRs are mainly located on vacuolar and plasma membranes (Davenport, 2002). They are potential candidates for the plasma membrane-level regulation for calcium influx (Mäser et al., 2001) and could be amino acid sensors in plants (Tapken et al., 2013). *Arabidopsis*, rice, and other model plants have been investigated for the functions of *GLRs*, however, until now, there have been no reports on the identification of the *GLR* gene family members in sweet potato. Because of the complicated genetic background of cultivated sweet potato, research on the sweet potato gene family is limited in the cultivated variety, mostly focusing on its two diploid relatives (Chen et al., 2019a; Chen et al., 2019b; Li et al., 2019; Lu et al., 2019; Wan et al., 2020; Zhu et al., 2020; Huang et al., 2021; Dai et al., 2022; Nie et al., 2023). Through this investigation, we conducted an in-depth study of the characteristics of *GLR* genes and compared their expression patterns in response to different types of biotic and abiotic stresses using genomic sequences for the hexaploid sweet potato and two diploid relatives. The genome-wide study of *GLR* genes provides vital guidance for the further studying their functions.

4.1 Evolution of *GLR* genes for sweet potato and two diploid relatives

Overall, *I. trifida* and *I. triloba* had the same number of *GLRs* (32 *ItfGLRs* and 32 *ItbGLRs*) identified, however they were less numerous than those (37 *IbGLRs*) in *I. batatas*. This is consistent with the fact that *I. trifida* and *I. triloba* are diploid and have a close genetic relationship. The evolution and differentiation of chromosomes were revealed by genomic alignment (Mukherjee et al., 2018). *I. batatas, I. trifida*, and *I. triloba* differed in terms of the distribution and proportion of *GLRs* in individual chromosomes. *GLRs* were located on nine chromosomes for *I. batatas* and *I. triloba*, however, *GLRs* were located on ten chromosomes for *I. trifida*. If different members of the same family are located within the same or adjacent intergenic regions,

10.3389/fpls.2023.1255805

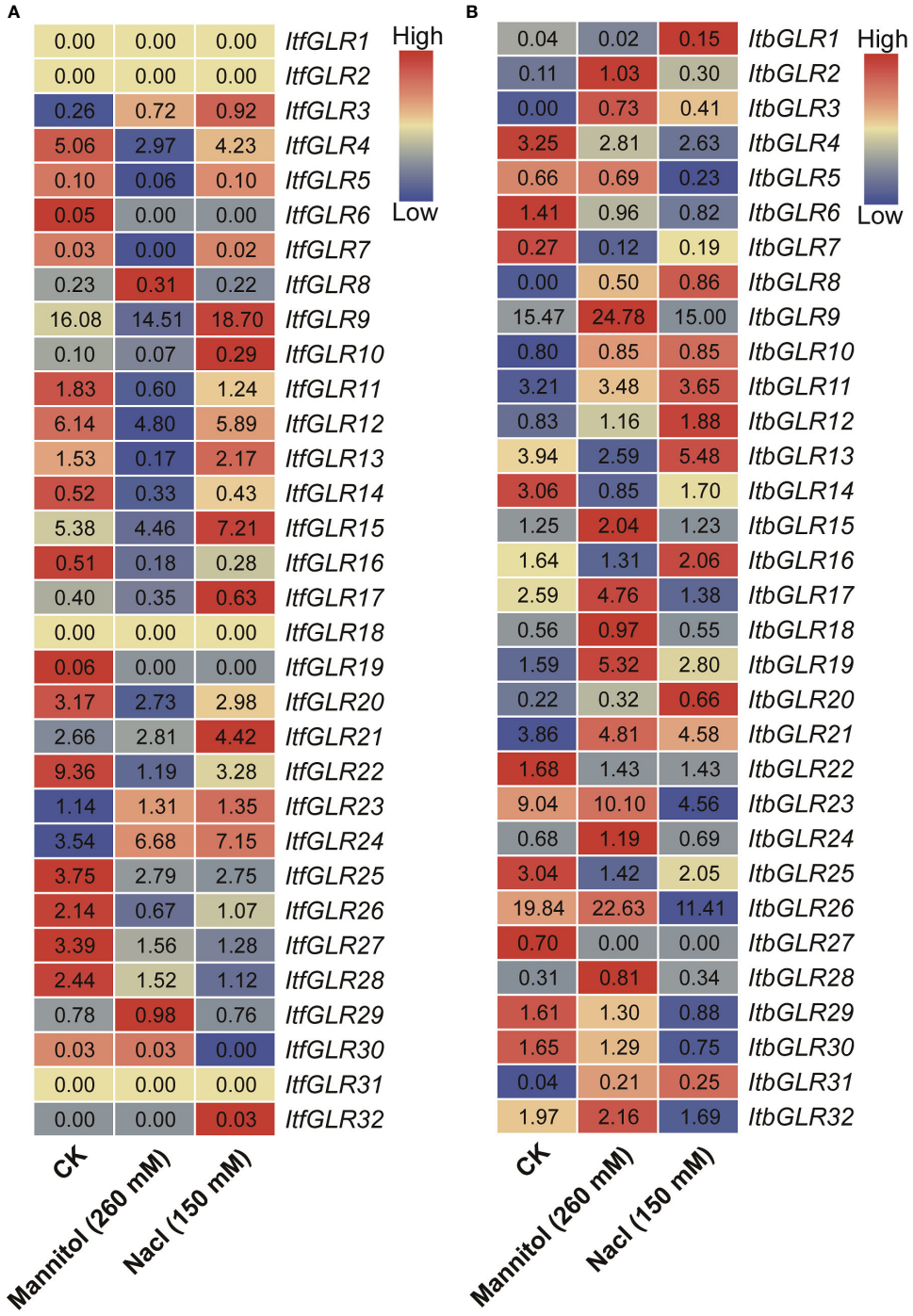


FIGURE 10
GLRs expression analysis in response to mannitol and NaCl stresses in *Ipomoea trifida* **(A)** and *Ipomoea triloba* **(B)**. The FPKM values are displayed in the boxes. The color bar only represents the expression values of one *GLR* gene in different treatments.

they have tandem repeat relationships (Freeling, 2009). Based on this standard, eight, six, and four tandem repeats were identified in *I. batatas*, *I. trifida*, and *I. triloba*, respectively (Figure 1). This indicates that tandem repeats might be one of the reasons for the amplification of *GLR* genes in *I. batatas*, *I. trifida*, and *I. triloba*.

This research is based on the GLRs for *I. batatas, I. trifida, and I. triloba*, which were segregated into five groups (Groups I to V), four groups (Groups I to III and V), and five groups (Groups I to V), respectively. The fourth and fifth subgroups consisted of 10 GLRs that were absent in *Arabidopsis*. The type and number of GLRs in various subgroups for sweet potato and two diploid relatives are different from those in *Arabidopsis*. These findings imply that the genomes may have experienced lineage-specific differentiation for *GLR* gene family.

Introns are important components of eukaryotic protein-coding genes, eliminated during messenger RNA precursor molecule splicing. They are characterized by their clear organization and

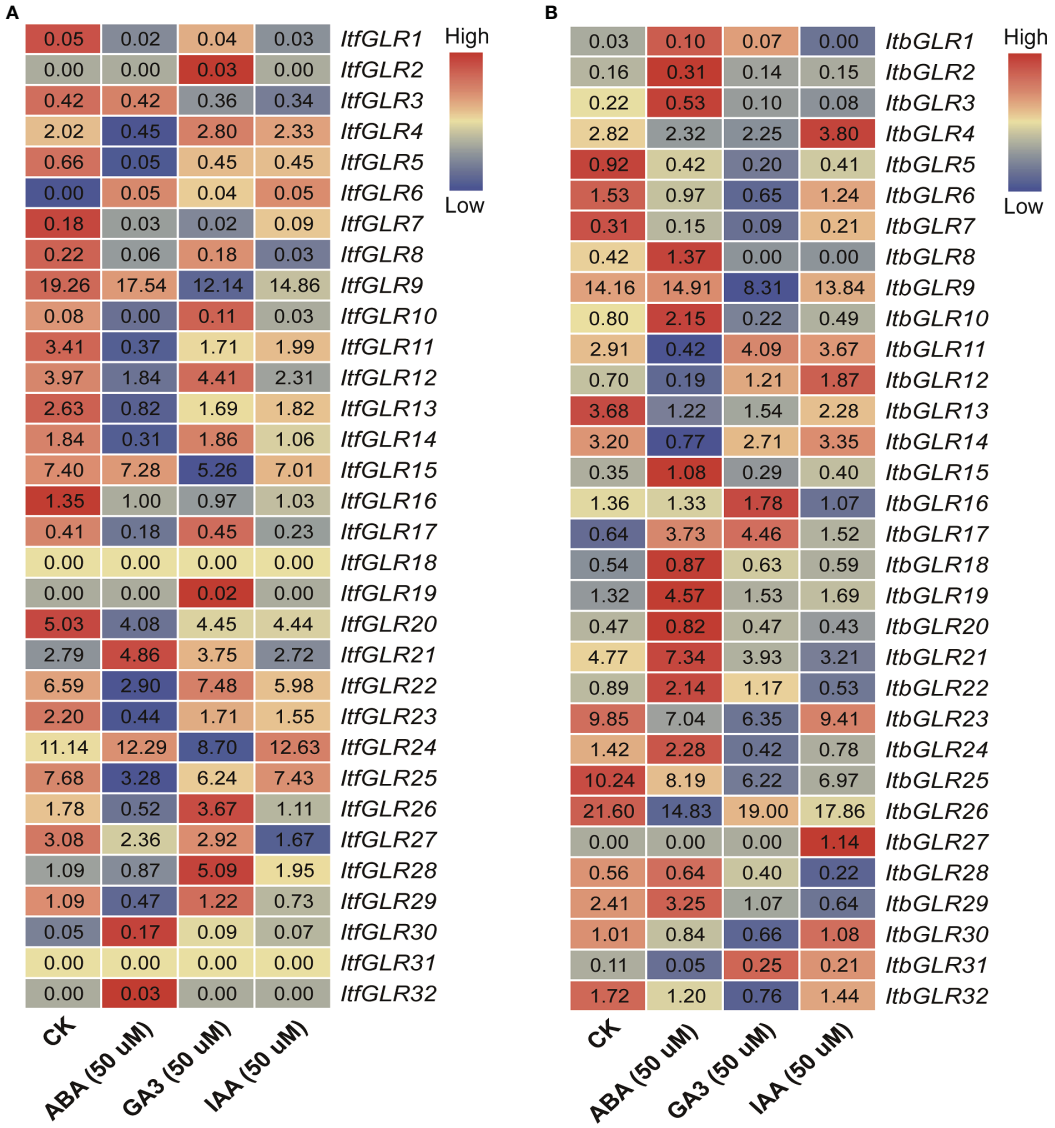


FIGURE 11
GLRs expression analysis in response to abscisic acid (ABA), gibberellin (GA), and indolent-3-acetic acid (IAA) treatments in *Ipomoea trifida* **(A)** and *Ipomoea triloba* **(B)**. The FPKM values are displayed in the boxes. The color bar only represents the expression values of one *GLR* gene in different treatments.

abundance in eukaryotic genes (Rogozin et al., 2005; Mukherjee et al., 2018). Because of the presence of introns, gene expression in eukaryotic cells is much more complex than that in prokaryotic cells. Under stressful conditions, introns assume a key role in regulating cell growth (Morgan et al., 2019). Herein, in comparison with *I. trifida* and *I. triloba*, *I. batatas* had distinct exon-intron patterns for some homologous *GLRs* (Figure 3C). For example, *IbGLR36*, which was expressed in fibrous roots, contained 17 introns, whereas its homologous genes *ItfGLR7* and *ItbGLR7* contained nine and ten introns, respectively, and were expressed in leaves. Moreover, *IbGLR30*, which was expressed in the stem, contained 17 introns, whereas *ItfGLR8* and *ItbGLR8* contained four and five introns, respectively, and were expressed in flowers (Figures 3C, 6). The corresponding differences in exon–intron structure between sweet potato and its two diploid relatives might lead to different functions for *GLRs* in various aspects of plant growth and development (Pang et al., 2018; Ma et al., 2019; Ma J et al., 2020).

4.2 Hormone crosstalk roles of *GLRs* for sweet potato and two diploid relatives

Plant *GLRs* are actively involved in hormone biosynthesis and signal transduction concerning the regulation of plant growth and responses to stress (Weiland et al., 2015). *PpGLR1* is involved in ABA-mediated growth regulation in *Physcomitrium patens* (Wang et al., 2022). Meanwhile, *RsGluR* could serve as a defensive mechanism against pathogen infection by triggering MeJA biosynthesis (Kang et al., 2006). In this work, the promoters of the *IbGLRs* contained 11 hormone-responsive elements, 12

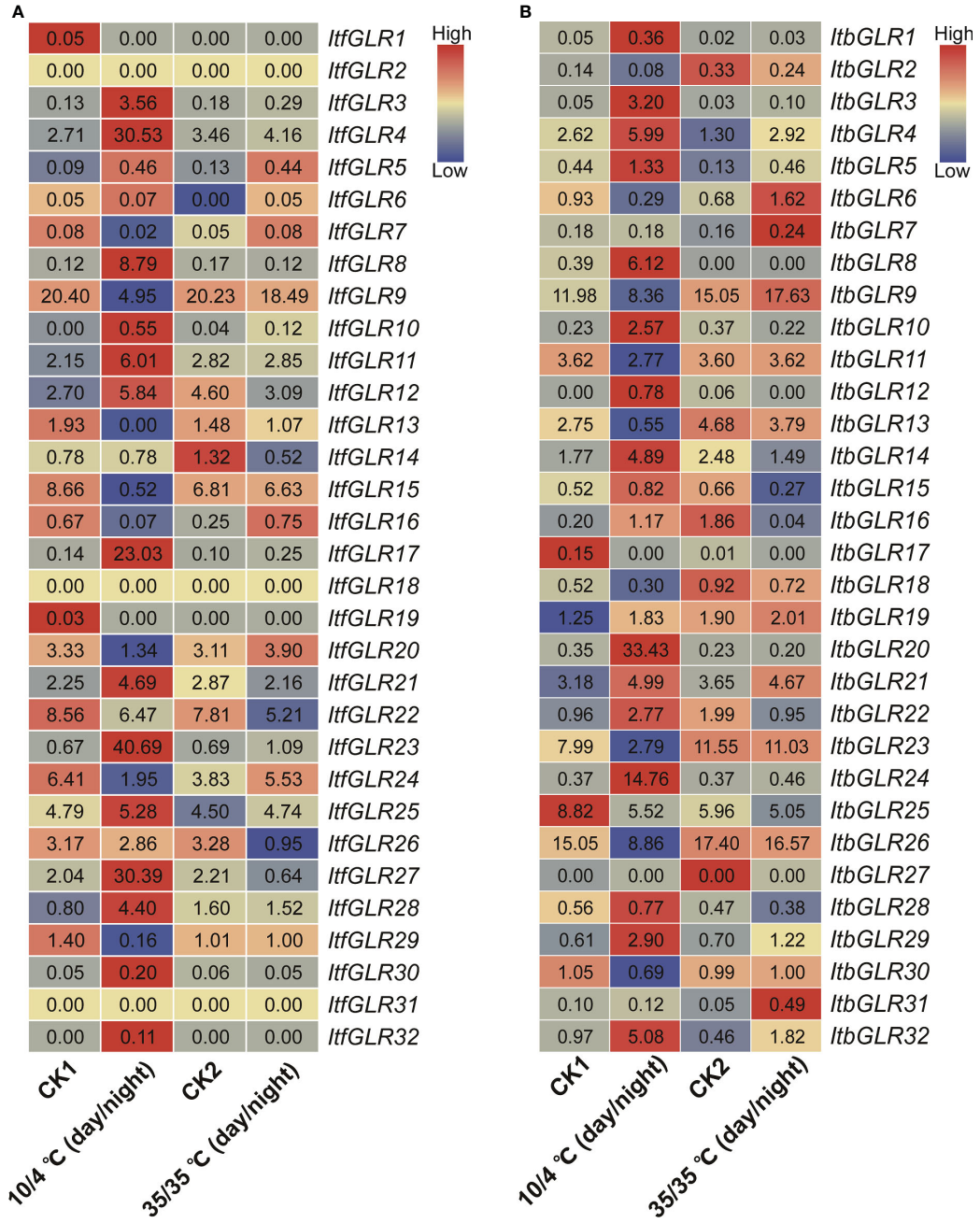


FIGURE 12
GLRs expression analysis under 10/4°C (day/night) and 35/35°C (day/night) treatments in *Ipomoea trifida* **(A)** and *Ipomoea triloba* **(B)**. CK1, cold control; CK2, heat control. FPKM values are shown in the boxes. The color bar only represents the expression values of one *GLR* gene in different treatments.

developmental elements, 19 light-responsive elements, and 21 abiotic/biotic-response elements. In addition to *IbGLR35*, other *IbGLRs* comprised at least one hormone element (Figure 4). *IbGLR3* only contained the P-box of the GA-responsive elements, whereas its homologous gene *ItbGLR18* was determined to be induced by ABA. *IbGLR4* was found to contain the TGACG-motif and CGTCA-motif for MeJA-responsive elements, TCA for SA-responsive elements, and the AuxRR-core together with TGA-element for IAA-responsive elements, whereas its homologous gene, *ItbGLR31*, was determined to be induced by GA and IAA. *IbGLR21* contained ABRE for ABA-responsive elements, the TGACG-motif together with CGTCA-motif for MeJA-responsive elements, and TCA for SA-responsive elements, whereas its homologous genes, *ItfGLR21* and *ItbGLR22*, were both determined to be induced by ABA and GA. *IbGLR30* carried the TGACG-motif and CGTCA-motif for MeJA-responsive elements, whereas its homologous gene, *ItbGLR8*, was determined to be induced by ABA (Figures 4, 11). Such results demonstrate *GLRs* involvement in the interplay of several hormones and the participation of homologous *GLR* genes in numerous hormone

pathways in sweet potato and two diploid relatives. Further research is needed to fully understand how *GLRs* control hormonal crosstalk.

4.3 Functions of *GLRs* in biotic stress responses and root growth of sweet potato and two diploid relatives

GLR plays an indispensable role in the response of plants to different biological stresses (He et al., 2016). Liu and colleagues discovered a point mutation for *GhGLR4.8* exon that can increase the resistance of upland cotton to *Fusarium* wilt, while knocking out the *GhGLR4.8* lead the defense ability of cotton cell wall against *G. hirsutum Fov race* 7 to weaken (Liu et al., 2021). *AtGLR3.3* also has a valuable function in the defense response induced by *Pseudomonas syringae* pv *tomato* DC3000. An *AtGLR3.3* mutant was shown to exhibit high sensitivity to DC3000, and the response of defense genes was reduced in mutant lines induced by pathogenic bacteria (Li et al., 2013). Sweet potato root rot infection usually begins at the underground stem and tip or middle of fibrous roots. Severely diseased plants do not produce tuberous roots, whereas slightly diseased plants produce tuberous roots with diseased spots (Ma Z et al., 2020).

Transcriptome and qRT-PCR analysis showed that *IbGLR18* expression was induced in Jishuzi203 (root rot-resistant variety) and repressed in Jishuzi563 (root rot-sensitive variety) after root rot infection. In addition, the *IbGLR18* expression in fibrous roots was higher in comparison to that in other tissues (leaf, stem, and tuberous root). Furthermore, *IbGLR18* expression was also higher in the fibrous root (diameter of root 1 mm) than that in tuberous roots (diameter of root 1/3/5/10 cm) in the five developmental stages of the Xushu22 root. Therefore, we speculate that *IbGLR18* might be involved in resistance to root rot and the formation of fibrous roots in sweet potato, but its function should be verified in the future studies.

Usually, tuberous roots are the main harvesting tissues for hexaploid sweet potato. *I. trifida* and *I. triloba* are not capable of forming tuberous roots (Wu et al., 2018). The expression for *IbGLR29* in tuberous roots was higher than that in other tissues (leaf, stem, and fibrous roots) of sweet potato, and its homolog genes *ItfGLR9* and *ItbGLR9* were highly expressed in the leaf. Further, in the five developmental stages of Xushu22 roots, the expression of *IbGLR29* in tuberous roots (root diameter of 1/3/5/10 cm) was high in comparison to that in fibrous roots (diameter of approximately 1 mm). *IbGLR29* may therefore be involved in the generation of tuberous roots in sweet potato.

4.4 Functions of *GLRs* in abiotic stress response of sweet potato and two diploid relatives

GLRs play a crucial role in mediating plants responses to abiotic environmental stress (He et al., 2016). The degree of expression of *ZmGLR2.3/3.1* demonstrated a significant increase after drought stress (Zhou et al., 2021). In this work, MYB and MYC were found to respond to drought stress, MBS to salt stress, and LTR to cryogenic stress, and these were identified in the *IbGLR19* promoter, which was rapidly expressed after drought stress and reached its expression maximum at the early stage (1 h). Moreover, its expression was upregulated by NaCl treatment in ND98 (salt-tolerant line) and downregulated in Lizixiang (salt-sensitive variety, Figures 4, 9). *ItfGLR23*, the homolog gene of *IbGLR19*, was induced by both drought and salt treatments in *I. trifida* (Figure 10). These results suggest that *IbGLR19* may contribute to abiotic stress. Additionally, *I. trifida* and *I. triloba* may be utilized for searching and identifying functional genes, particularly those that provide tolerance or resistance to biotic and abiotic stresses, which may have been lost during the domestication of cultivated sweet potato (Cao et al., 2016). In *I. trifida* and *I. triloba*, several genes (*ItfGLR8* and its homologous gene *ItbGLR8*, *ItfGLR17* and its homologous gene *ItbGLR20*, *ItfGLR23* and its homologous gene *ItbGLR24*, and *ItfGLR28* and its homologous gene *ItbGLR29*) exhibited the same expression patterns and were induced under 10/4°C (day/night) (Figure 12). *ItfGLR3* and its homologous gene *ItbGLR3* were induced by drought and salt treatments (Figure 10). In summary, the *GLRs* induced in *I. trifida* and *I. triloba* could serve as candidate genes for enhancing the abiotic stress resistance of sweet potato.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: NCBI SRA (http://www.ncbi.nlm.nih.gov/Traces/sra), accession numbers SAMN10755180, SAMN10755181, SAMN10755182, SAMN10755183, SAMN10755184, SAMN10755185, SAMN10755186, SAMN10755187, SAMN10755188, SAMN10755189, SAMN10755190, SAMN10755191, SAMN10755192, SAMN10755193, SAMN10755194 and SRP092215.

Author contributions

YH: Writing – original draft, Writing – review & editing. ZD: Writing – original draft. JH: Writing – original draft. MH: Writing – original draft. ZW: Writing – original draft. WJ: Writing – original draft. ZG: Writing – original draft. XL: Writing – review & editing. LL: Writing – review & editing. ZM: Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the National Natural Science Foundation of China (32301957), the China Agriculture Research System

(CARS-10-Sweetpotato), the Basic Research Funds of Hebei Academy of Agriculture and Forestry Sciences (2021060207), the Key Research and Development Project of Hebei Province (22322911D), the Modern Agriculture Industrial Technology System Innovation Team Project of Hebei Province (HBCT2023060202), and the HAAFS Science and Technology Innovation Special Project (2022KJCXZX-LYS-12). The funding bodies were not involved in the design of the study, in the collection, analysis and interpretation of data or in writing the manuscript.

Acknowledgments

We are grateful to Huan Zhang (College of Agronomy and Biotechnology, China Agricultural University) for her suggestions on our manuscript.

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1255805/full#supplementary-material

References

Afrin, T., Costello, C. N., Monella, A. N., Kørner, C. J., and Pajerowska-Mukhtar, K. M. (2022). The interplay of GTP-binding protein AGB1 with ER stress sensors IRE1a and IRE1b modulates *Arabidopsis* unfolded protein response and bacterial immunity. *Plant Signal Behav.* 17, 2018857. doi: 10.1080/15592324.2021.2018857

Ahmed, I., Kumar, A., Bheri, M., Srivastava, A. K., and Pandey, G. K. (2023). Glutamate receptor like channels: Emerging players in calcium mediated signaling in plants. *Int. J. Biol. Macromol.* 234, 123522. doi: 10.1016/j.ijbiomac.2023.123522

Aouini, A., Matsukura, C., Ezura, H., and Asamizu, E. (2012). Characterisation of 13 glutamate receptor-like genes encoded in the tomato genome by structure, phylogeny and expression profiles. *Gene* 493, 36–43. doi: 10.1016/j.gene.2011.11.037

Bian, C., Demirer, G. S., and Brady, S. M. (2022). GLRs: mediating a defense-regeneration tradeoff in plants. *Dev. Cell* 57, 417–418. doi: 10.1016/j.devcel.2022.02.007

Cao, Q., Li, A., Chen, J., Sun, Y., Tang, J., Zhang, A., et al. (2016). Transcriptome sequencing of the sweet potato progenitor (*Ipomoea Trifida* (HBK) G. Don.) and discovery of drought tolerance genes. *Trop. Plant Biol.* 9, 63–72. doi: 10.1007/s12042-016-9162-7

Chang, F., Yan, A., Zhao, L., Wu, W., and Yang, Z. (2007). A putative calcium-permeable cyclic nucleotide-gated channel, CNGC18, regulates polarized pollen tube growth. *J. Integr. Plant Biol.* 49, 1261–1270. doi: 10.1111/j.1672-9072.2007.00524.x

Chen, J., Jing, Y., Zhang, X., Li, L., Wang, P., Zhang, S., et al. (2016). Evolutionary and expression analysis provides evidence for the plant glutamate-like receptors family is involved in woody growth-related function. *Sci. Rep.* 6, 32013. doi: 10.1038/srep32013

Chen, Y., Zhu, P., Wu, S., Lu, Y., Sun, J., Cao, Q., et al. (2019a). Identification and expression analysis of GRAS transcription factors in the wild relative of sweet potato *Ipomoea trifida*. *BMC Genomics* 20, 911. doi: 10.1186/s12864-019-6316-7

Chen, Y., Zhu, P., Zhang, L., Wu, S., Ma, D., Cao, Q., et al. (2019b). Genome-wide identification, structural and gene expression analysis of the bZIP transcription factor family in sweet potato wild relative *Ipomoea trifida*. *BMC Genet.* 20, 41. doi: 10.1186/s12863-019-0743-y

Chiu, J. C., Brenner, E. D., DeSalle, R., Nitabach, M. N., Holmes, T. C., and Coruzzi, G. M. (2002). Phylogenetic and expression analysis of the glutamate-receptor-like gene family in *Arabidopsis thaliana*. *Mol. Biol. Evol.* 19, 1066–1082. doi: 10.1093/oxfordjournals.molbev.a004165

Cho, D., Kim, S. A., Murata, Y., Lee, S., Jae, S. K., Nam, H. G., et al. (2009). De-regulated expression of the plant glutamate receptor homolog *AtGLR3.1* impairs long-term Ca^{2+}-programmed stomatal closure. *Plant J.* 58, 437–449. doi: 10.1111/j.1365-313X.2009.03789.x

Dai, Z., Yan, P., He, S., Jia, L., Wang, Y., Liu, Q., et al. (2022). Genome-wide identification and expression analysis of SWEET family genes in sweet potato and its two diploid relatives. *Int. J. Mol. Sci.* 23, 15848. doi: 10.3390/ijms232415848

Davenport, R. (2002). Glutamate receptors in plants. *Ann. Bot.* 90, 549–557. doi: 10.1093/abo/mcf228

Dong, T., Zhu, M., Yu, J., Han, R., Tang, C., Xu, T., et al. (2019). RNA-Seq and iTRAQ reveal multiple pathways involved in storage root formation and development in sweet potato (*Ipomoea batatas* L.). *BMC Plant Biol.* 19, 136. doi: 10.1186/s12870-019-1731-0

Food and Agriculture Organization of the United Nations. (2021). *Sweetpotato*. Available at: https://www.fao.org/faostat/zh/#search/Sweet%20potatoes (Accessed March 15, 2023).

Forde, B. G., and Lea, P. J. (2007). Glutamate in plants: metabolism, regulation, and signalling. *J. Exp. Bot.* 58, 2339–2358. doi: 10.1093/jxb/erm121

Freeling, M. (2009). Bias in plant gene content following different sorts of duplication: tandem, whole-genome, segmental, or by transposition. *Annu. Rev. Plant Biol.* 60, 433–453. doi: 10.1146/annurev.arplant.043008.092122

Gu, L. L., Gao, Q. F., and Wang, Y. F. (2017). Cyclic nucleotide-gated channel 18 functions as an essential $Ca2^{+}$ channel for pollen germination and pollen tube growth in *Arabidopsis*. *Plant Signal. Behav.* 12, e1197999. doi: 10.1080/15592324.2016.1197999

Hajibarat, Z., Saidi, A., and Hajibarat, Z. (2022). Genome-wide identification of *14-3-3* gene family and characterization of their expression in developmental stages of *Solanum tuberosum* under multiple biotic and abiotic stress conditions. *Funct. Integr. Genomics* 22, 1377–1390. doi: 10.1007/s10142-022-00895-z

He, M., Sun, Y., Chen, X., Shi, D., Li, D., Chen, Y., et al. (2016). Current research advances on glutamate receptors (GLRs) in plants. *Acta Bot. Sin.* 51, 827–840. doi: 10.11983/CBB15212

Hernández-Coronado, M., Araujo, P. C. D., Ip, P. L., Nunes, C. O., Rahni, R., Wudick, M. M., et al. (2022). Plant glutamate receptors mediate a bet-hedging strategy between regeneration and defense. *Dev. Cell* 57, 451–465.e6. doi: 10.1016/j.devcel.2022.01.013

Hu, Y. R., Jiang, L. Q., Wang, F., and Yu, D. Q. (2013). Jasmonate regulates the inducer of CBF expression-C-repeat binding factor/DRE binding factor1 cascade and freezing tolerance in *Arabidopsis*. *Plant Cell* 25, 2907–2924. doi: 10.1105/tpc.113.112631

Huang, Z., Wang, Z., Li, X., He, S., Liu, Q., Zhai, H., et al. (2021). Genome-wide identification and expression analysis of JAZ family involved in hormone and abiotic stress in sweet potato and its two diploid relatives. *Int. J. Mol. Sci.* 22, 9786. doi: 10.3390/ijms22189786

Kang, J., Mehta, S., and Turano, F. J. (2004). The putative glutamate receptor 1.1 (AtGLR1.1) in *Arabidopsis thaliana* regulates abscisic acid biosynthesis and signaling to control development and water loss. *Plant Cell Physiol.* 45, 1380–1389. doi: 10.1093/pcp/pch159

Kang, S., Kim, H. B., Lee, H., Choi, J. Y., Heu, S., Oh, C. J., et al. (2006). Overexpression in *Arabidopsis* of a plasma membrane-targeting glutamate receptor from small radish increases glutamate-mediated Ca^{2+} influx and delays fungal infection. *Mol. Cells* 21, 418–427.

Kim, S. A., Kwak, J. M., Jae, S. K., Wang, M. H., and Nam, H. G. (2001). Overexpression of the *AtGluR2* gene encoding an *Arabidopsis* homolog of mammalian glutamate receptors impairs calcium utilization and sensitivity to ionic stress in transgenic plants. *Plant Cell Physiol.* 42, 74–84. doi: 10.1093/pcp/pce008

Kohl, M., Wiese, S., and Warscheid, B. (2011). Cytoscape: Software for visualization and analysis of biological networks. *Methods Mol. Biol.* 696, 291–303. doi: 10.1007/978-1-60761-987-1_18

Kong, D., Ju, C., Parihar, A., Kim, S., Cho, D., and Kwak, J. M. (2015). *Arabidopsis* glutamate receptor homolog 3.5 modulates cytosolic Ca^{2+} level to counteract effect of abscisic acid in seed germination. *Plant Physiol.* 167, 1630–1642. doi: 10.1104/pp.114.251298

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Latza, A., Mehlmerb, N., Zapf, S., Mueller, T. D., Wurzinger, B., Pfister, B., et al. (2013). Salt stress triggers phosphorylation of the *Arabidopsis* vacuolar K^+ channel TPK1 by calcium-dependent protein kinases (CDPKs). *Mol. Plant* 6, 1274–1289. doi: 10.1093/mp/sss158

Lescot, M., Dehais, P., Thijs, G., Marchal, K., Moreau, Y., Peer, Y. V., et al. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* 30, 325–327. doi: 10.1093/nar/30.1.325

Li, F., Wang, J., Ma, C., Zhao, Y., Wang, Y., Hasi, A., et al. (2013). Glutamate receptor-like channel3.3 is involved in mediating glutathione-triggered cytosolic calcium transients, transcriptional changes, and innate immunity responses in *Arabidopsis*. *Plant Physiol.* 162, 1497–1509. doi: 10.1104/pp.113.217208

Li, T., Sun, Y., Ruan, Y., Xu, L., Hu, Y., Hao, Z., et al. (2016). Potential role of D-myo-inositol-3-phosphate synthase and 14-3-3 genes in the crosstalk between *Zea* mays and *Rhizophagus intraradices* under drought stress. *Mycorrhiza* 26, 879–893. doi: 10.1007/s00572-016-0723-2

Li, Y., Zhang, L., Zhu, P., Cao, Q., Sun, J., Li, Z., et al. (2019). Genome-wide identification, characterisation and functional evaluation of *WRKY* genes in the sweet potato wild ancestor *Ipomoea trifida* (H.B.K.) G. Don. under abiotic stresses. *BMC Genet.* 20, 90. doi: 10.1186/s12863-019-0789-x

Liu, S., Zhang, X., Xiao, S., Ma, J., Shi, W., Qin, T., et al. (2021). A single-nucleotide mutation in a glutamate receptor-like gene confers resistance to *Fusarium* wilt in *Gossypium hirsutum*. *Adv. Sci.* 8, 2002723. doi: 10.1002/advs.202002723

Lu, L. (2011). In silico cloning and bioinformatic analysis of *TPK1* Gene in tobacco. *Sci. Agric. Sin.* 44, 28–35. doi: 10.3864/j.issn.0578-1752.2011.01.004

Lu, Y., Sun, J., Yang, Z., Zhao, C., Zhu, M., Ma, D., et al. (2019). Genome-wide identification and expression analysis of glycine-rich RNA-binding protein family in sweet potato wild relative *Ipomoea trifida*. *Gene* 686, 177–186. doi: 10.1016/j.gene.2018.11.044

Luan, S., and Wang, C. (2021). Calcium signaling mechanisms across kingdoms. *Annu. Rev. Cell Dev. Biol.* 37, 311–340. doi: 10.1146/annurev-cellbio-120219-035210

Ma, J., Deng, S., Jia, Z., Sang, Z., Zhu, Z., Zhou, C., et al. (2020). Conservation and divergence of ancestral *AGAMOUS/SEEDSTICK* subfamily genes from the basal angiosperm *Magnolia wufengensis*. *Tree Physiol.* 40, 90–107. doi: 10.1093/treephys/tpz091

Ma, Z., Gao, W., Liu, L., Liu, M., Zhao, N., Han, M., et al. (2020). Identification of QTL for resistance to root rot in sweetpotato (*Ipomoea batatas* (L.) Lam) with SSR linkage maps. *BMC Genomics* 21, 366. doi: 10.1186/s12864-020-06775-9

Ma, D. F., Li, Q., Cao, Q. H., Niu, F. X., Xie, Y. P., Tang, J., et al. (2012). Development and prospect of sweetpotato industry and its technologies in China. *Jiangsu J. Agric. Sci.* 28, 969–973.

Ma, R., Song, W., Wang, F., Cao, A., Xie, S., Chen, X., et al. (2019). A Cotton (*Gossypium hirsutum*) *Myo*-Inositol-1-Phosphate Synthase (*GhMIPS1D*) gene promotes root cell elongation in *Arabidopsis*. *Int. J. Mol. Sci.* 20, 1224. doi: 10.3390/ijms20051224

Mäser, P., Thomine, S., Schroeder, J. I., Ward, J. M., Hirschi, K., Sze, H., et al. (2001). Phylogenetic relationships within cation transporter families of *Arabidopsis*. *Plant Physiol.* 126, 1646–1667. doi: 10.1104/pp.126.4.1646

Mayer, M. L. (2006). Glutamate receptors at atomic resolution. *Nature* 440, 456–462. doi: 10.1038/nature04709

Meng, Y. S., and Lai, Q. X. (2019). Application and prospects of ornamental sweetpotato. *J. Zhejiang Agric. Sci.* 60, 2181–2184,2244. doi: 10.16178/j.issn.0528-9017.20191206

Michard, E., Lima, P. T., Borges, F., Silva, A. C., Portes, M. T., Carvalho, J. E., et al. (2011). Glutamate receptor-like genes form Ca^{2+} channels in pollen tubes and are regulated by pistil D-serine. *Science* 332, 434–437. doi: 10.1126/science.1201101

Moccia, F., Zuccolo, E., Nezza, F. D., Pellavio, G., Faris, P. S., Negri, S., et al. (2021). Nicotinic acid adenine dinucleotide phosphate activates two-pore channel TPC1 to mediate lysosomal Ca^{2+} release in endothelial colony-forming cells. *J. Cell Physiol.* 236, 688–705. doi: 10.1002/jcp.29896

Morgan, J. T., Fink, G. R., and Bartel, D. P. (2019). Excised linear introns regulate growth in yeast. *Nature* 565, 606–611. doi: 10.1038/s41586-018-0828-1

Mukherjee, D., Saha, D., Acharya, D., Mukherjee, A., Chakraborty, S., and Ghosh, T. C. (2018). The role of introns in the conservation of the metabolic genes of *Arabidopsis thaliana*. *Genomics* 110, 310–317. doi: 10.1016/j.ygeno.2017.12.003

Navarro-Retamal, C., Schott-Verdugo, S., Gohlke, H., and Dreyer, I. (2021). Computational analyses of the AtTPC1 (*Arabidopsis* Two-Pore Channel 1) permeation pathway. *Int. J. Mol. Sci.* 22, 10345. doi: 10.3390/ijms221910345

Nie, N., Huo, J., Sun, S., Zuo, Z., Chen, Y., Liu, Q., et al. (2023). Genome-wide characterization of the PIFs family in sweet potato and functional identification of *IbPIF3.1* under drought and *Fusarium* wilt stresses. *Int. J. Mol. Sci.* 24, 4092. doi: 10.3390/ijms24044092

Pang, X., Wei, Y., Cheng, Y., Pan, L., Ye, Q., Wang, R., et al. (2018). The tryptophan decarboxylase in *Solanum lycopersicum*. *Molecules* 23, 998. doi: 10.3390/molecules23050998

Philippe, F., Verdu, I., Morère-Le Paven, M. C., Limami, A. M., and Planchet, E. (2019). Involvement of *Medicago truncatula* glutamate receptor-like channels in nitric oxide production under short-term water deficit stress. *J. Plant Physiol.* 236, 1–6. doi: 10.1016/j.jplph.2019.02.010

Pottosin, I., and Dobrovinskaya, O. (2022). Major vacuolar TPC1 channel in stress signaling: what matters, K^+, Ca^{2+} conductance or an ion–flux independent mechanism? *Stress Biol.* 2, 31. doi: 10.1007/s44154-022-00055-0

Qi, X., Tang, W., Li, W., He, Z., Xu, W., Fan, Z., et al. (2021). *Arabidopsis* G-Protein β subunit AGB1 negatively regulates DNA binding of MYB62, a suppressor in the gibberellin pathway. *Int. J. Mol. Sci.* 22, 8270. doi: 10.3390/ijms22158270

Rogozin, I. B., Sverdlov, A. V., Babenko, V. N., and Koonin, E. V. (2005). Analysis of evolution of exon-intron structure of eukaryotic genes. *Brief. Bioinform.* 6, 118–134. doi: 10.1093/bib/6.2.118

Singh, S. K., Chien, C. T., and Chang, I. F. (2016). The *Arabidopsis* glutamate receptor-like gene *GLR3.6* controls root development by repressing the Kip-related protein gene *KRP4*. *J. Exp. Bot.* 67, 1853–1869. doi: 10.1093/jxb/erv576

Singh, A., Kanwar, P., Yadav, A. K., Mishra, M., Jha, S. K., Baranwal, V., et al. (2014). Genome-wide expressional and functional analysis of calcium transport elements during abiotic stress and development in rice. *FEBS J.* 281, 894–915. doi: 10.1111/febs.12656

Sobolevsky, A. I., Rosconi, M. P., and Gouaux, E. (2009). X-ray structure, symmetry and mechanism of an AMPA-subtype glutamate receptor. *Nature* 462, 745–756. doi: 10.1038/nature08624

Su, Y. J., Dong, L. X., Wang, J., Dai, X. B., Zhang, A., Zhao, D. L., et al. (2018). Genetic diversity in vegetable and ornamental sweetpotato germplasm. *J. Plant Genet. Res.* 19, 57–64. doi: 10.13430/j.cnki.jpgr.2018.01.007

Takahashi, T., Murano, T., and Ishikawa, A. (2018). SOBIR1 and AGB1 independently contribute to nonhost resistance to *Pyricularia oryzae* (syn. *Magnaporthe oryzae*) in *Arabidopsis thaliana*. *Biosci. Biotechnol. Biochem.* 82, 1922–1930. doi: 10.1080/09168451.2018.1498727

Tapken, D., Anschütz, U., Liu, L. H., Huelsken, T., Seebohm, G., Becker, D., et al. (2013). A plant homolog of animal glutamate receptors is an ion channel gated by multiple hydrophobic amino acids. *Sci. Signal.* 6, ra47. doi: 10.1126/scisignal.2003762

Tian, W., Wang, C., Gao, Q., Li, L., and Luan, S. (2020). Calcium spikes, waves and oscillations in plant development and biotic interactions. *Nat. Plants* 6, 750–759. doi: 10.1038/s41477-020-0667-6

Traynelis, S. F., Wollmuth, L. P., McBain, C. J., Menniti, F. S., Vance, K. M., Ogden, K. K., et al. (2010). Glutamate receptor ion channels: structure, regulation, and function. *Pharmacol. Rev.* 62, 405–496. doi: 10.1124/pr.109.002451

Vincill, E. D., Clarin, A. E., Molenda, J. N., and Spalding, E. P. (2013). Interacting glutamate receptor-like proteins in phloem regulate lateral root initiation in *Arabidopsis*. *Plant Cell* 25, 1304–1313. doi: 10.1105/tpc.113.110668

Wan, R., Liu, J., Yang, Z., Zhu, P., Cao, Q., and Xu, T. (2020). Genome-wide identification, characterisation and expression profile analysis of *DEAD-box* family genes in sweet potato wild ancestor *Ipomoea trifida* under abiotic stresses. *Genes Genom.* 42, 325–335. doi: 10.1007/s13258-019-00910-x

Wan, S., Wang, W., Zhou, T., Zhang, Y., Chen, J., Xiao, B., et al. (2018). Transcriptomic analysis reveals the molecular mechanisms of *Camellia sinensis* in response to salt stress. *Plant Growth Regul.* 84, 481–492. doi: 10.1007/s10725-017-0354-4

Wang, S., Song, M., Guo, J., Huang, Y., Zhang, F., Cheng, X., et al. (2018). The potassium channel *FaTPK1* plays a critical role in fruit quality formation in strawberry (*Fragaria* × *ananassa*). *Plant Biotechnol. J.* 16, 737–748. doi: 10.1111/pbi.12824

Wang, Y., Yu, D., Zhao, H., Jiang, L., Gao, L., Song, Y., et al. (2022). A glutamate receptor-like gene is involved in ABA-mediated growth control in *Physcomitrium (Physcomitrella)* patens. *Plant Signal. Behav.* 17, e2145057. doi: 10.1080/15592324.2022.2145057

Weiland, M., Mancuso, S., and Baluska, F. (2015). Signalling via glutamate and GLRs in *Arabidopsis thaliana*. *Funct. Plant Biol.* 43, 1–25. doi: 10.1071/FP15109

Wu, S., Lau, K. H., Cao, Q., Hamilton, J. P., Sun, H., Zhou, C., et al. (2018). Genome sequences of two diploid wild relatives of cultivated sweetpotato reveal targets for genetic improvement. *Nat. Commun.* 9, 4580. doi: 10.1038/s41467-018-06983-8

Wudick, M. M., Michard, E., Nunes, C. O., and Feijó, J. A. (2018a). Comparing plant and animal glutamate receptors: common traits but different fates? *J. Exp. Bot.* 69, 4151–4163. doi: 10.1093/jxb/ery153

Wudick, M. M., Portes, M., Michard, T. E., Rosas-Santiago, P., Lizzio, M. A., Nunes, C. O., et al. (2018b). CORNICHON sorting and regulation of GLR channels underlie pollen tube Ca^{2+} homeostasis. *Science* 360, 533–536. doi: 10.1126/science.aar6464

Xiang, C., Shen, S. F., Ji, Z. X., Li, B., and Wu, L. H. (2020). Pedigree and quality traits of Zheshu sweetpotato varieties for table use and food processing use. *J. Nucl. Agric. Sci.* 34, 36–44. doi: 10.11869/j.issn.100-8551.2020.01.0036

Xie, Y. Z., Guo, X. D., Jia, Z. D., Ma, P. Y., Bian, X. F., and Yu, Y. (2018). Progresses and prospects on edible sweetpotato breeding in China. *Jiangsu J. Agric. Sci.* 34, 1419–1424. doi: 10.3969/j.issn.1000-4440.2018.06.030

Yan, M., Chen, M., Zhang, W., Hu, Y., Lu, Y., He, S., et al. (2022). Identification and expression characteristics analysis of *GLRs* gene family of *Erigeron breviscapus*. *Mol. Plant Breed.* 20, 2842–2853. doi: 10.13271/j.mpb.020.002842

Yang, J., Moeinzadeh, M. H., Kuhl, H., Helmuth, J., Xiao, P., Haas, S., et al. (2017). Haplotype-resolved sweet potato genome traces back its hexaploidization history. *Nat. Plants* 3, 696–703. doi: 10.1038/s41477-017-0002-z

Yang, L., Zhao, Y., Wu, X., Zhang, Y., Fu, Y., Duan, Q., et al. (2022). Genome wide identification and expression analysis of *BraGLRs* reveal their potential roles in abiotic stress tolerance and sexual reproduction. *Cells* 11, 3729. doi: 10.3390/cells11233729

Yu, Y., and Assmann, S. M. (2018). Inter-relationships between the heterotrimeric Gβ subunit AGB1, the receptor-like kinase FERONIA, and RALF1 in salinity response. *Plant Cell Environ.* 41, 2475–2489. doi: 10.1111/pce.13370

Yu, B., Wu, Q., Li, X., Zeng, R., Min, Q., and Huang, J. (2022). GLUTAMATE RECEPTOR-like gene *OsGLR3.4* is required for plant growth and systemic wound signaling in rice (*Oryza sativa*). *New Phytol.* 233, 1238–1256. doi: 10.1111/nph.17859

Zeng, H., Zhao, B., Wu, H., Zhu, Y., and Chen, H. (2020). Comprehensive in silico characterization and expression profiling of nine gene families associated with calcium transport in soybean. *Agronomy* 10, 1539. doi: 10.3390/agronomy10101539

Zhang, J., Cui, T., Su, Y., Zang, S., Zhao, Z., Zhang, C., et al. (2022). Genome-wide identification, characterization, and expression analysis of glutamate receptor-like gene (*GLR*) family in sugarcane. *Plants (Basel).* 11, 2440. doi: 10.3390/plants11182440

Zhang, T., Xu, P., Wang, W., Wang, S., Caruana, J. C., Yang, H., et al. (2020). Corrigendum: *Arabidopsis* G-protein β subunit AGB1 interacts with BES1 to regulate brassinosteroid signaling and cell elongation. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.01122

Zhang, H., Zhang, Q., Zhai, H., Li, Y., Wang, X., Liu, Q., et al. (2017). Transcript profile analysis reveals important roles of jasmonic acid signalling pathway in the response of sweet potato to salt stress. *Sci. Rep.* 7, 40819. doi: 10.1038/srep40819

Zheng, Y., Luo, L. D., Wei, J. J., Chen, Q., Yang, Y. P., Hu, X. Y., et al. (2018). The glutamate receptors *AtGLR1.2* and *AtGLR1.3* increase cold tolerance by regulating jasmonate signaling in *Arabidopsis thaliana*. *Biochem. Biophys. Res. Commun.* 506, 895–900. doi: 10.1016/j.bbrc.2021.06.061

Zhou, Z. L., Tang, J., Cao, Q. H., Zhao, D. L., and Zhang, A. (2020). Formation laws of quality characters in starch sweetpotato cultivars and its correlation with main agronomic characters. *Jiangsu J. Agric. Sci.* 36, 277–283. doi: 10.3969/j.issn.1000-4440.2020.02.004

Zhou, S., Zhang, L., Lv, Q., and Huang, J. (2021). Identification and analysis of *GLR* family genes in maize. *J. Maize Sci.* 29, 35–42. doi: 10.13597/j.cnki.maize.science.20210206

Zhu, P., Dong, T., Xu, T., and Kang, H. (2020). Identification, characterisation and expression analysis of *MADS-box* genes in sweetpotato wild relative *Ipomoea trifida*. *Acta Physiol. Plant* 42, 163. doi: 10.1007/s11738-020-03153-6

Zhu, S., and Gouaux, E. (2017). Structure and symmetry inform gating principles of ionotropic glutamate receptors. *Neuropharmacology* 112, 11–15. doi: 10.1016/j.neuropharm.2016.08.034

Zhu, H., Zhou, Y., Zhai, H., He, S., Zhao, N., and Liu, Q. (2019). Transcriptome profiling reveals insights into the molecular mechanism of drought tolerance in sweetpotato. *J. Integr. Agric.* 18, 9–23. doi: 10.1016/S2095-3119(18)61934-3

frontiers | Frontiers in Plant Science

TYPE Review
PUBLISHED 08 January 2024
DOI 10.3389/fpls.2023.1334620

OPEN ACCESS

EDITED BY
Jian Sun,
Jiangsu Normal University, China

REVIEWED BY
Jun Cui,
Hunan Normal University, China
Giampiero Cai,
University of Siena, Italy

*CORRESPONDENCE
Huaijun Si
hjsi@gsau.edu.cn

†These authors have contributed equally to this work

RECEIVED 07 November 2023
ACCEPTED 15 December 2023
PUBLISHED 08 January 2024

CITATION
Jin X, Wang Z, Li X, Ai Q, Wong DCJ, Zhang F, Yang J, Zhang N and Si H (2024) Current perspectives of lncRNAs in abiotic and biotic stress tolerance in plants. *Front. Plant Sci.* 14:1334620. doi: 10.3389/fpls.2023.1334620

Current perspectives of lncRNAs in abiotic and biotic stress tolerance in plants

Xin Jin[1,2†], Zemin Wang[1,2†], Xuan Li[2], Qianyi Ai[2], Darren Chern Jan Wong[3], Feiyan Zhang[2], Jiangwei Yang[1,2], Ning Zhang[1,2] and Huaijun Si[1,2*]

[1]State Key Laboratory of Aridland Crop Science, Gansu Agricultural University, Lanzhou, China, [2]College of Life Science and Technology, Gansu Agricultural University, Lanzhou, China, [3]Division of Ecology and Evolution, Research School Research of Biology, The Australian National University, Acton, ACT, Australia

Abiotic/biotic stresses pose a major threat to agriculture and food security by impacting plant growth, productivity and quality. The discovery of extensive transcription of large RNA transcripts that do not code for proteins, termed long non-coding RNAs (lncRNAs) with sizes larger than 200 nucleotides in length, provides an important new perspective on the centrality of RNA in gene regulation. In plants, lncRNAs are widespread and fulfill multiple biological functions in stress response. In this paper, the research advances on the biological function of lncRNA in plant stress response were summarized, like as Natural Antisense Transcripts (NATs), Competing Endogenous RNAs (ceRNAs) and Chromatin Modification etc. And in plants, lncRNAs act as a key regulatory hub of several phytohormone pathways, integrating abscisic acid (ABA), jasmonate (JA), salicylic acid (SA) and redox signaling in response to many abiotic/biotic stresses. Moreover, conserved sequence motifs and structural motifs enriched within stress-responsive lncRNAs may also be responsible for the stress-responsive functions of lncRNAs, it will provide a new focus and strategy for lncRNA research. Taken together, we highlight the unique role of lncRNAs in integrating plant response to adverse environmental conditions with different aspects of plant growth and development. We envisage that an improved understanding of the mechanisms by which lncRNAs regulate plant stress response may further promote the development of unconventional approaches for breeding stress-resistant crops.

KEYWORDS

long noncoding RNA, small RNA, competing endogenous RNAs, epigenetic, abiotic stress, biotic stress

Introduction

Ribonucleic acid (RNAs) are divided into coding RNAs and non-coding RNAs (ncRNAs). The ncRNAs constitute a class of RNA which includes microRNAs (miRNAs), small interfering RNAs (siRNAs), long non-coding RNAs (lncRNAs), and circular RNAs (circRNAs). The different types of aforementioned ncRNAs are involved in transcriptional and post-transcriptional regulation of gene expression, as well as regulation of RNA stability and translation (Wang et al., 2017; Qiu et al., 2019; Waititu et al., 2021). Instances where lncRNAs are transcribed by RNA polymerase II (Pol II) and/or Pol IV/V are also known (Rai et al., 2019). LncRNAs can be further classified as antisense lncRNAs or intron lncrnas (lincRNAs), depending on their genomic location (Wierzbicki et al., 2021).

LncRNA are heterogeneously regulated transcript that has emerged as a versatile regulator in a variety of development processes (Zhang et al., 2021a; Cao et al., 2022; Ren et al., 2022; Ye et al., 2022). In *Brassica rapa*, 47 *cis*-acting lncRNAs and 451 *trans*-acting lncRNAs were identified during pollen development and fertilization. These lncRNAs were notably co-expressed with their target genes (Huang et al., 2018). In tomato, the silencing of two lncRNAs (lncRNA1459 and lncRNA1840) delayed fruit ripening (Zhu et al., 2015). Moreover, Li et al. (2018) found that tomato fruit ripening was significantly inhibited with CRISPR/cas9-mediated mutation of lncRNA1459. In cucumber, lncRNA *CsM10* is specifically expressed in male cultivated cucumber species and may play a role in sexual differentiation (Cho et al., 2006). Beyond developmental processes, emerging findings indicate that lncRNAs also fulfill important roles in abiotic and biotic stress (Moison et al., 2021; Song et al., 2021; Ye et al., 2022). In many plants, numerous stress-related lncRNAs have been found by high throughput sequence analyzes (Wu et al., 2019; Gong et al., 2021; Li et al., 2022). Similar to protein-coding genes, lncRNAs display stress-responsive differential expression in addition to tissue- and developmental-specific expression patterns. Furthermore, small RNAs may also influence lncRNA expression in abiotic stress. In wheat, the lncRNAs –*TalnRNA9* and *TalnRNA12* – could be regulated by 24 nt siRNAs (Xin et al., 2011).

LncRNAs and miRNAs are both non-coding RNAs, and lncRNAs have the potential to act as competing endogenous RNAs (ceRNAs) to regulate the expression of target genes (Shuai et al., 2014). Additionally, lncRNAs can be targeted by miRNAs and may also serve as an endogenous target mimic (eTM) and/or possible miRNA precursors (Wu et al., 2013; Li et al., 2017). By analyzing competing relationships between mRNAs and lncRNAs, genome-scale lncRNA-mediated ceRNAs networks have been inferred in many plants (Fan et al., 2018; He et al., 2020; Ren et al., 2022). For instance, lncRNA regulates the expression of heat stress response genes (e.g. HEAT-SHOCK PROTEIN, *HSPs* and HEAT STRESS TF, *HSFs*) by competitively binding to bra-miR159a or bra-miR172a in Chinese cabbage (Wang et al., 2019a). Certain lncRNAs also function as natural antisense transcripts (NATs) that are transcribed in an orientation opposite to that of protein-coding genes. Examples include Arabidopsis *MAS* and potato *StFLORE* (Zhao et al., 2018; Ramirez et al., 2021). Some ncRNAs produced by PolIV are precursors of siRNAs (Movahedi et al. (2015) but may also participate in scaffolding RNA by modulating the chromatin framework. Examples include the Arabidopsis lncRNA *COOLAIR* (Csorba et al., 2014) and *APOLO* (Moison et al., 2021). Many stress-responsive lncRNAs have also been identified by screening the genome-wide binding peaks of ChIP-seq data (Di et al., 2014). For example, 107 stress-responsive lncRNAs were inferred from the binding signals of PIF4 and PIF5 TFs ChIP-seq data. Many conserved sequence and enriched structural motifs in different functional lncRNAs groups were also identified. They include a UUC motif responding to salt and a AU-rich stem-loop correlated with cold stress response (Di et al., 2014). As such, these conserved elements might also be responsible for the stress-responsive functions of lncRNAs.

In light of their emerging roles, lncRNA holds promise as a viable target for enhancing agronomic traits and stress tolerance in crop breeding endeavors. In this review, we systematically summarize the plant lncRNAs involved in stress responses and their crucial, but often overlooked, roles in plants.

Abiotic stress-associated lncRNAs in plants

Adverse environmental conditions, such as extreme temperature, drought, and high salinity, negatively impact plant growth and development. In response to abiotic stress, a myriad of stress-responsive proteins and regulatory factors are often highly induced (Zhu, 2016; Zhang et al., 2022a). Similarly, lncRNAs fulfilling roles in the regulation of abiotic stress responses share such characteristics (Wang et al., 2019a; Yang et al., 2022; Ye et al., 2022). The roles of many lncRNAs in the context of abiotic stress responses have now been well investigated. (Figures 1–4). This is in part enabled by the advent of next-generation sequencing technologies that have provided the identification of many lncRNAs from various plant species, growth and development processes (e.g. root development, flowering time control), and stress conditions (Zhao et al., 2018; Gong et al., 2021; Moison et al., 2021; Li et al., 2022). In these studies, we also found that some transcription factors are widely involved in the regulation of lncRNA in different stress responses, such as lncRNA-NAC in drought, salt stress, and temperature stress. the implication is that lncrnas participate in the common regulatory network in various abiotic stresses by regulating some key members of the network.

LncRNAs response to extreme temperature stress in plants

Low temperature

Low-temperature stress constitutes a prominent factor influencing plant growth and development, leading to a substantial reduction in crop production (Gong et al., 2020; Ding and Yang, 2022; Zhang et al., 2022a). In plants, C-REPEAT (CRT)-

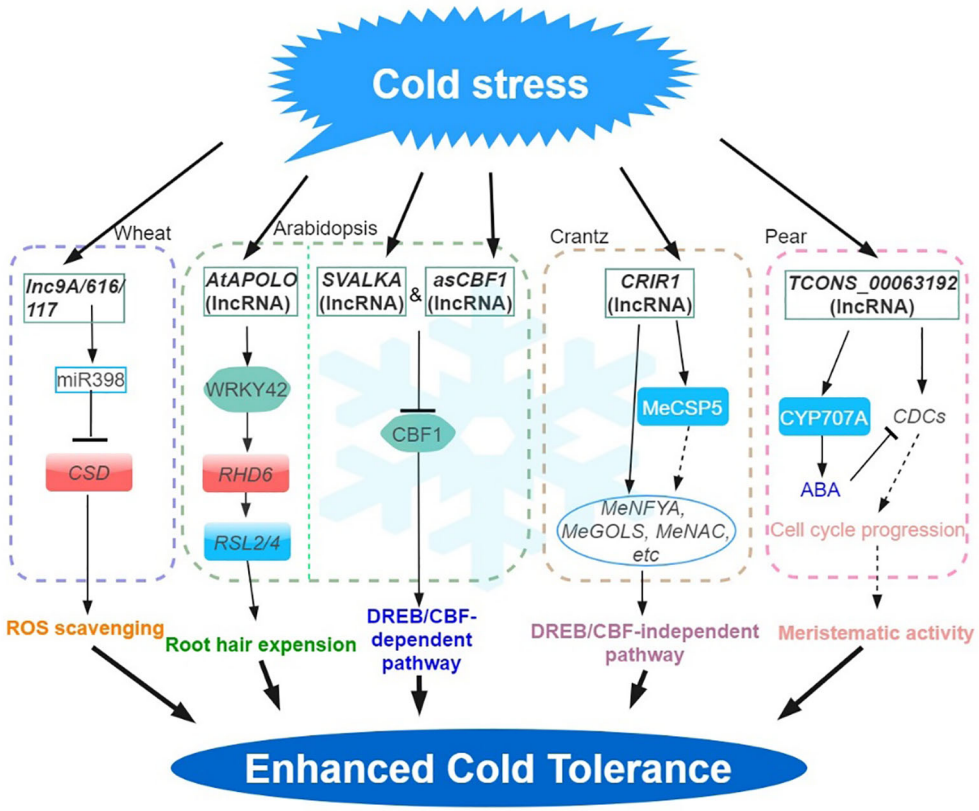


FIGURE 1
Regulatory network of the lncRNAs involved in cold stress. Cold stress-responsive lncRNAs in Arabidopsis, wheat, crantz and pear are shown in the chart. The arrows and hammer represent positive and negative regulation, respectively. APOLO, lncRNA AUXIN-REGULATED PROMOTER LOOP; MeGOLS, GALACTINOL SYNTHASES; DREB, DEHYDRATION-RESPONSIVE ELEMENT (DRE) BINDING FACTOR; CBF, C-REPEAT (CRT)-BINDING FACTOR; CSP, COLD SHOCK DOMAIN-CONTAINING PROTEIN; ROS, REACTIVE OXYGEN SPECIES.

BINDING FACTOR/dehydra-RESPONSIVE ELEMENT (DRE) -BINDING FACTOR (CBF/DREB) -mediated transcriptional regulatory cascade pathway is the main low temperature signaling pathway. This mechanism is key to inducing a series of cold response (COR) genes (Shi et al., 2015). In Arabidopsis, *asCBF1* – a cryptic lncRNA (overlapping *CBF1* on the antisense strand) produced by RNAPII read-through transcription of the lncRNA, *SVALKA* – represses *CBF1* and promotes cold acclimation. This *SVALKA-asCBF1* cascade pathway provides a negative feedback for plants to adapt to low temperatures (by maximizing cold tolerance while reducing adaptation costs) through the strict control of *CBF1* expression and timing (Kindgren et al., 2018). In grapes, 487 and 326 lncRNAs were significantly up- and down-regulated during cold stress, respectively. Among them, 203 differentially expressed lncRNAs were predicted to regulate 326 target genes. Among them, many differentially expressed lncRNAs were predicted to target stress response-related genes such as *CBF* and *WRKY* transcription factors, late embryogenesis abundant (LEAs), and peroxisome biogenesis-related genes (Wang et al., 2019b).

In cassava, a total of 316 lncRNAs were characterized as cold-related lncRNAs (Li et al., 2022). Functional analysis shows that the trans-acting lncRNA, CRIR1 can confer cold stress tolerance by interacting with a cold shock domain protein (MeCSP) to improve translation efficiency at low temperatures. Many cold-related genes (e.g. *MeNAC*, *MeNF-YA* and *MeGOLS*, GALACTINOL SYNTHASES), were upregulated in WT and *CRIR1*-OE plants during cold stress, however, CBF-dependent pathway cold stress-related marker genes (e.g. *MeCBFs*, *MeCOR* and *MeICE1*) were not differentially expressed (DE). This suggested that *CRIR1* enhanced plant cold tolerance via a CBF-independent pathway, likely via the auxin signaling (Li et al., 2022).

In *Arabidopsis*, pre-mRNA alternative splicing (AS) and lncRNA expression changes have been associated with stress response. For example, by exploiting a new Arabidopsis transcriptome prediction (named AtRTD2), regulatory and post-transcriptional regulation of lncRNA gene expression in response to cold stress has been possible (Calixto et al., 2019). Briefly, the analysis of a comprehensive RNA-seq time-series dataset identified 135 lncRNA genes with cold-dependent DE and/or differential AS of lncRNAs. By carrying out experiments on plants cultivated under different temperatures and employing high-resolution time-course analysis spanning the initial three hours before temperature reduction, it was demonstrated that the AS of specific lncRNAs exhibits a remarkable sensitivity to even minute temperature fluctuations. This finding suggested that the expression of these lncRNAs are subject to stringent regulation (Calixto et al., 2019). In

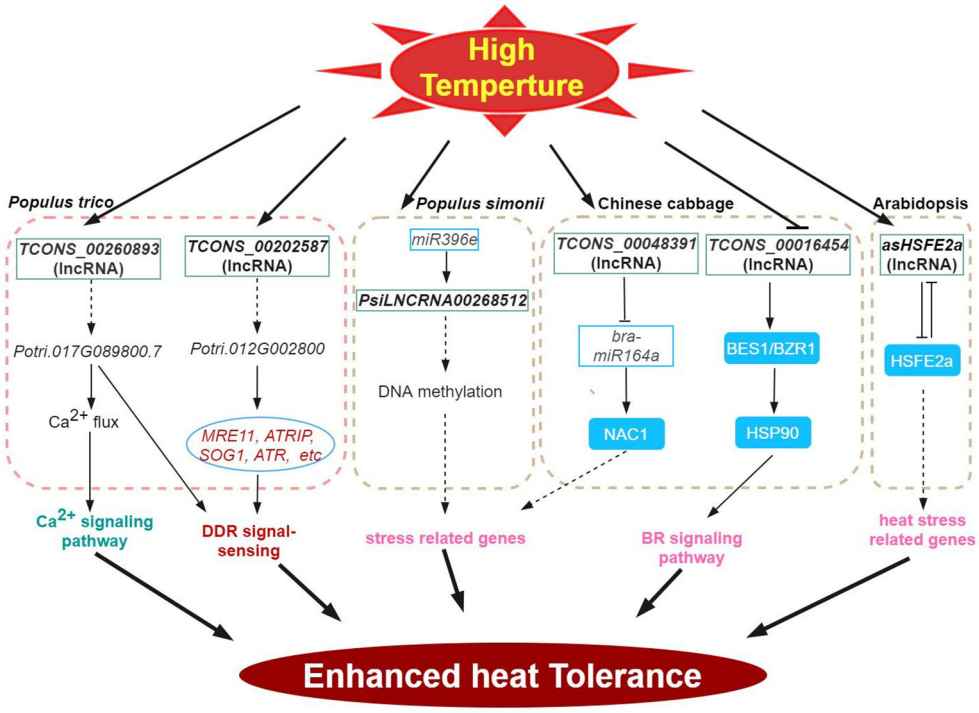


FIGURE 2
High-temperature stress signaling and responses of lncRNAs in plants. HSP, HEAT-SHOCK PROTEIN; BES1, BRASSINOSTEROID INSENSITIVE 1EMS-SUPRESSOR 1; BZR1, BRASSINAZOLE RESISTANT 1; HSF, HEAT STRESS TF.

addition, the identification of a cold-repressive lincRNA159 has been made possible using strand-specific RNA-seq. Functional analysis shows that lincRNA159 is the target mimics of miR164, and 16 of the 682 lncRNAs were identified to act as ceRNA in cassava (Li et al., 2017). *MiR398* is reported to respond to diverse abiotic stresses (Leng et al., 2017). In wheat, the ceRNAs –*lncR9A*, *lncR117*, and *lncR616* – interact with *miR398*, to regulate the expression of *COPPER/ZINC SUPEROXIDE DISMUTASE 1* (*CSD1*) and improve tolerance to cold stress (Lu et al., 2020).

Natural antisense transcripts (NATs) is another type of lncRNAs playing significant roles in various physiological processes. NAT lncRNAs can serve as scaffolds to facilitate chromatin conformation changes, DNA methylation, and histone modifications (Chen et al., 2012). The induction of *MAS*, a NAT lncRNA transcribed from the *MADS AFFECTING FLOWERING4 (MAF4)* locus, occurs as a result of cold and leads to the upregulation of *MAF4* expression at the transcriptional level. The interaction of MAS with WDR5a (a core component of the COMPASS-like complexes) further mediates the recruitment of WDR5a to *MAF4*, thus, enhancing H3K4me3 (Zhao et al., 2018). These findings demonstrate the involvement of NAT-lncRNAs in the regulation of gene expression during the vernalization response. Some plant lncRNAs also mediate the regulatory pathway of chromatin modification (lncR2Epi) (Wang et al., 2016). Previous studies showed that cold-induced lncRNA *COOLAIR* can inhibit the expression of *FLOWERING LOCUS C* (*FLC*) in *Arabidopsis* under cold stress through the lncR2Epi regulatory pathway (Heo et al., 2013; Csorba et al., 2014). During vernalization, *COOLAIR* and *COLD ASSISTED INTRONIC NONCODING RNA* (*COLDAIR*) are produced by the *FLC* gene. *COOLAIR* is a NAT lncRNAs of *FLC*, and the expression level increased rapidly in the early vernalization stage, effect on *FLC* transcriptional shutdown is independent histone H3 lysine 27 trimethylation (H3K27me3) accumulation, but mediates reduction in H3K36me3 at *FLC* during vernalization (Kim and Sung, 2012; Csorba et al., 2014). In contrast, *COLDAIR* is transcribed by the first intron of the *FLC* gene, is induced at a later stage of low temperature conditions and continuously increases after vernalization (Kim and Sung, 2012). In further analysis, *COLDWRAP* (*COLD OF WINTER-INDUCED NONCODING RNA FROM THE PROMOTER*), an *FLC* promoter-derived lncRNA, is induced by vernalization and functions with the *COLDAIR* (Kim and Sung, 2017). *COLDAIR* cooperates with *COLDWRAP* resulting in the repression of *FLC* gene by forming an intragenic chromatin loop, leading to the establishment of high H3K27me3 (Swiezewski et al., 2009; Kim and Sung, 2017). In Arabidopsis, the interaction of the lncRNA *AUXIN-REGULATED PROMOTER LOOP* (*APOLO*), with WRKY42 coordinates the activation of *RHD6* by binding to the *RHD6* promoter, which in turn modulates local chromatin 3D conformation, leading to root hair cell elongation in response to low temperatures (Moison et al., 2021). Furthermore, bud dormancy serves as a crucial and intricate protection mechanism employed by plants in winter, like grapes, apples and pears (Masocha et al., 2020; Wang et al., 2020; Li et al., 2021). In pear, the potential roles of lncRNAs in regulating the dormancy release by interactions with mRNA and miRNAs that are associated with the ABA degradation

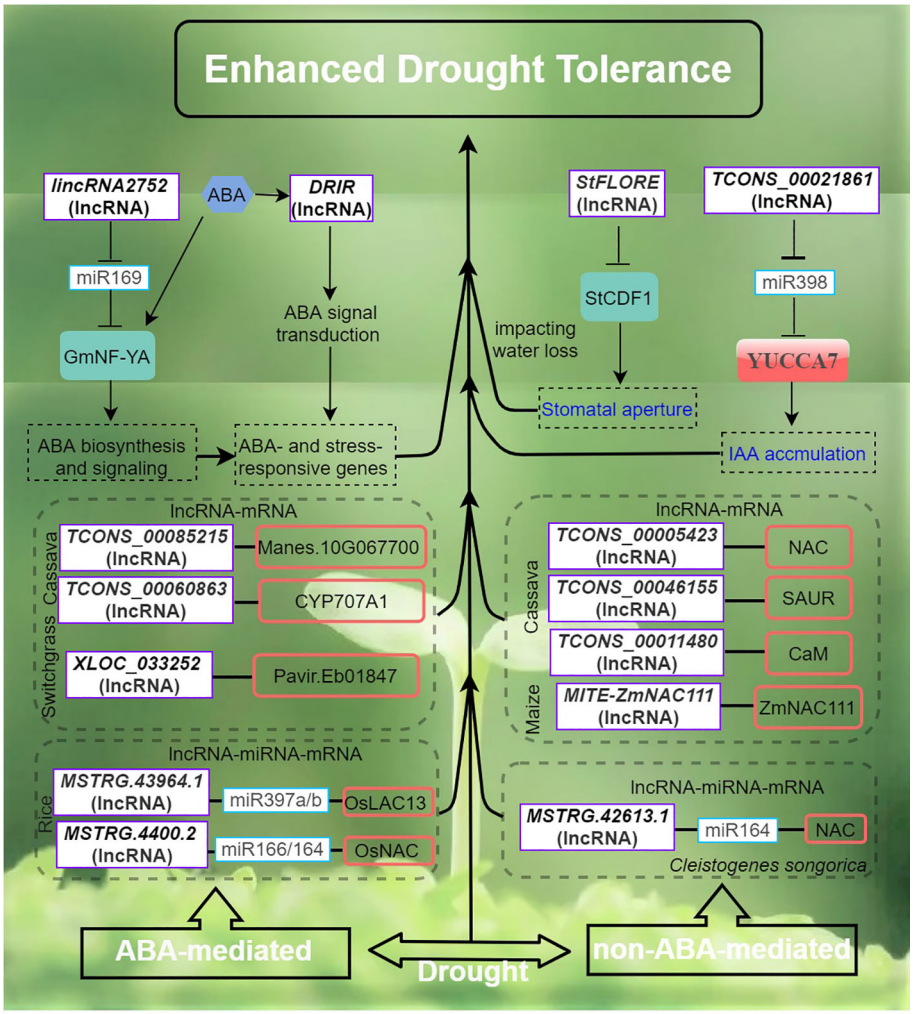


FIGURE 3
Drought stress signaling and responses of lncRNAs in plants are primarily by ABA-dependent and ABA-independent pathways. The arrows and hammer represent positive and negative regulation, respectively. ABA, abscisic acid; StCDF1, CYCLING DOF FACTOR 1; DRIR, DROUGHT INDUCED lncRNA; StFLORE, The StCDF1 locus also codes for an antisense lncRNA; YUCCA7, flavin monooxygenase gene.

pathway have also been proposed following simulated chilling temperature regimes to induce buds to enter different dormancy states (Li et al., 2021).

High temperature

Like low temperature, heat stress significantly impacts the growth and development of plants, thus, limiting global crop productivity (Gong et al., 2020; Ding and Yang, 2022; Lesk et al., 2022; Zhang et al., 2022a) At high temperatures, plants often induce genes encoding reactive-oxygen-scavenging (ROS) enzymes and the regulatory proteins (e.g. TFs and protein kinases), among others (general plant heat stress review here). Accumulating evidence has shown the importance of lncRNAs in plant responses to heat stress. Several heat stress response co-expression networks between lncRNA and mRNA have been inferred by analyzing cis-acting lncRNAs (Bhatia et al., 2020; He et al., 2020; Hu et al., 2022). Consistently, many TF genes were enriched in such networks providing support for the regulatory role of lncRNAs in the regulation of TFs during heat stress response potential targets of heat stress-responsive lncRNAs (204 transcripts) have also been inferred in poplar (Song et al., 2020). In wheat, Xin et al. (2011) characterized 77 lncRNAs during heat stress, and one candidate, *TahlnRNA27* was identified as a putative miRNA precursor. lncRNAs can be further classified into poly(A)+ and poly(A)- transcript type. Compared with poly(A)+ lncRNAs, the poly(A)- lncRNAs exhibit comparatively shorter transcripts and lower levels of expression and show notable specificity in their expression patterns in response to stresses (Di et al., 2014). In *Arabidopsis*, 245 DE poly(A)+ and 58 poly(A)- lncRNAs were responsive under various stress, however, of these, only 15 were heat-responsive (Di et al., 2014). In B. rapa, many of the 192 predicted target genes of the 34 DE lncRNAs identified under heat stress were also heat responsive (Song et al., 2016a). Comparative phenotype and transcriptome analysis of the heat-resistant elite maize inbred line under heat stress revealed 993 heat-responsive DE lncRNAs with close to a thousand predicted cis-regulated target genes that share

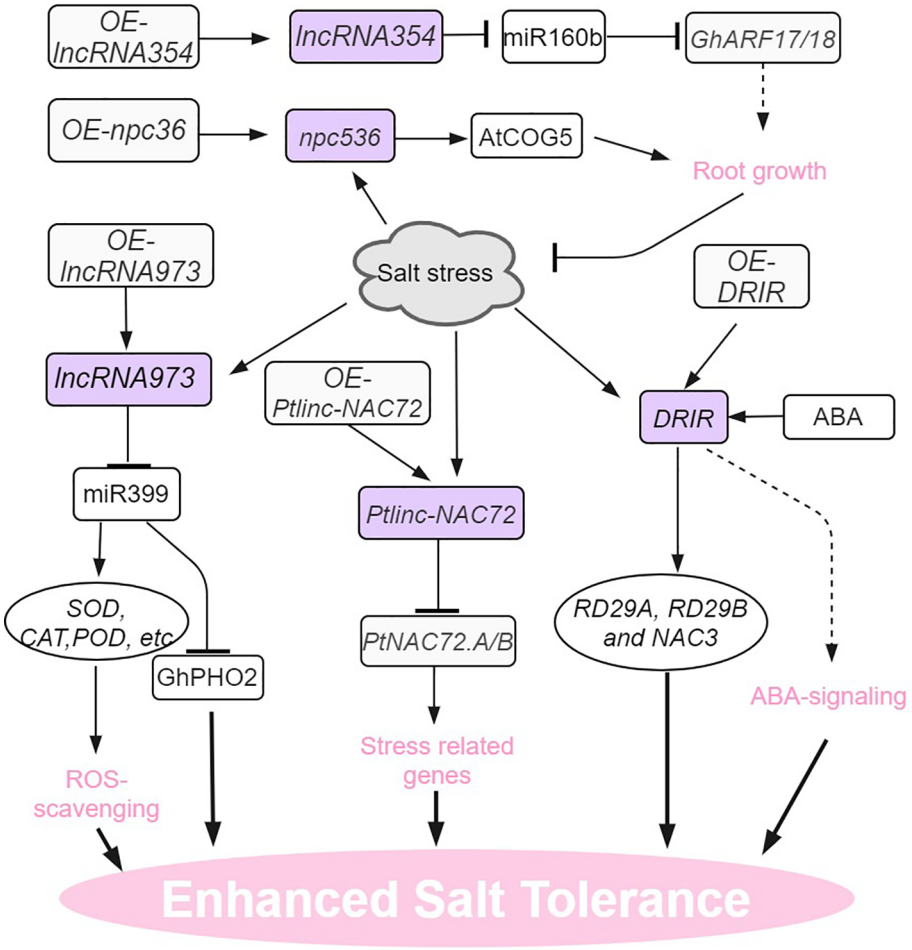


FIGURE 4
Salt stress signaling and responses of lncRNAs in plants. AtCOG5, CONSERVED OLIGOMERIC GOLGI COMPLEX 5; DRIR, DROUGHT INDUCED lncRNA; RD29, RESPONSIVE TO DESSICATION 29. The arrows and hammer represent positive and negative regulation, respectively.

strong co-expression (Hu et al., 2022). Many important biological processes and pathways involved in heat response including stress response, hormone signaling, metabolism, photosynthesis and spliceosome were also highly enriched (Hu et al., 2022).

DNA methylation plays an important role in various plant stress responses. In an abiotic stress-tolerant poplar (*Populus simonii*), comparative methylome and gene expression analysis under four abiotic stresses (cold, osmotic, heat, and salt stress) enabled the prioritization of stress-specific differentially methylated regions (SDMRs) containing 17 lncRNA genes. In particular, the epigenetic pathway-mediated regulation of SDMR162 region encompassed the *miR396e* gene and *PsiLNCRNA00268512*. Stable predicted interaction energies between *PsiLNCRNA00268512* and *miR396e* indicate that *PsiLNCRNA00268512* may serve as a target mimic regulating *miR396e* levels (Song et al., 2016b). In Chinese cabbage, heat-responsive lncRNA (*TCONS_00048391*) regulates the expression of *bra-miR164a* by acting as an endogenous target mimic (eTM) of *bra-miR164a*. Furthermore, this lncRNA is predicted to function as a sponge for miRNA binding or as a ceRNA that targets *NAC1* (Wang et al., 2019a).

The HEAT STRESS TF (HSFs) is the central regulator of heat stress response (Scharf et al., 2012; Zhu, 2016; Chen et al., 2018). For example, Arabidopsis *HSFB2a* is expressed in the female gametophyte and regulates vegetative and gametophyte development processes (Wunderlich et al., 2014). Transgenic experiments revealed that the NAT lncRNA of *HSFB2a*, *asHSFB2a* specifically counteracts *HSFB2a* expression when overexpressed. And showed improved biomass production. Conversely, overexpression of *HSFB2a* impaired biomass production. Following heat stress, the repressive effects of *HSFB2a* is counteracted by the antisense regulation of *HSFB2a* thereby restoring growth and further development (Wunderlich et al., 2014). In another example, the HSP90 participates in the regulation of BR signaling and BR responsive genes through the interaction with BRASSINOSTEROID INSENSITIVE 1EMS-SUPRESSOR 1 (BES1)/BRASSINAZOLE RESISTANT 1 (BZR1) TF, (Shigeta et al., 2015). In the Chinese cabbage, the lncRNA *TCONS_00016454* may be involved in conferring heat tolerance via BR signaling as this lncRNA showed antagonistic expression patterns with its target, a *BES1/BZR1* homolog in response to high temperatures in this system (Wang et al., 2019a).

In plants, intracellular concentration of free Ca^{2+} exhibits fast and stimulus-specific patterns (Zhang et al., 2022a). In *Populus simonii*, the lncRNA *TCONS_00260893* is involved in the regulation of Ca^{2+} flux and Ca^{2+} signaling under stress conditions by interfering with the target gene transcription encoding a cyclic nucleotide-regulated ion channel family protein (Potri.017G089800) via RNA interference and/or as RNA scaffolds. Heterologous expression of Potri.017G089800 in Arabidopsis enhanced photosynthetic protection and recovery, inhibited membrane peroxidation, and suppressed DNA damage in response to heat stress (Song et al., 2020). Potential ceRNA functions of lncRNAs and circRNAs in response to high-temperature stress have also been presented for cucumber (He et al., 2020).

LncRNAs response to drought stress in plants

Drought poses a significant global concern in agriculture as it typically leads to reduced crop yield and quality (Gupta et al., 2020; Zhang et al., 2022a). LncRNAs are believed to play crucial roles in the regulation of gene expression in plants exposed to drought stress. To date, various lncRNAs regulatory pathways of plant drought tolerance and the functions of plant lncRNAs and their ceRNA pairs in response to drought have been elucidated (Shuai et al., 2014; Qin and Xiong, 2019). In rice, strand-specific sequencing and small RNA sequencing have enabled the identification of putative lncRNA, miRNA and mRNA associated with drought resistance and the construction of an integrated ceRNA network (Yang et al., 2022). A total of 191 lncRNAs and 32 miRNAs were found to be DE in drought-stressed rice. Interestingly, up to 16 drought-specific lncRNAs were predicted to target one drought-responsive gene, Os05g0586700 potentially implicated in plant membrane repair during drought (Yang et al., 2022). In another study, a rice drought-responsive ceRNA network, consisting of 40 lncRNAs, 23 miRNAs, and 103 mRNAs, was also presented (Chen et al., 2021). Particularly, lncRNA TCONS_00021861 could regulate YUCCA7 by sponging miR528-3p, thereby validating its potential role as ceRNA. This interaction subsequently triggers the activation of the IAA biosynthetic pathway, leading to the acquisition of drought stress tolerance.

In peanuts, integrated ceRNA networks have also been established with comparative transcriptomics between varieties with divergent levels of drought tolerance (Ren et al., 2022). A total of 673 ceRNA pairs were inferred in the drought-tolerant variety, and the resulting ceRNA network analysis identified six lncRNAs, namely *MSTRG.70535.2*, *MSTRG.86570.2*, *MSTRG.86570.1*, *MSTRG.100618.1*, *MSTRG.81214.2*, and *MSTRG.30931.1*, which were recognized as central nodes in the network. Components of this specific ceRNA network may have significant implications in enhancing the drought tolerance of peanuts (Ren et al., 2022). Similarly, comparative transcriptome analysis of two Tibetan wild barley with contrasting drought tolerance revealed 535 DE lncRNAs with distinct expression profiles under drought stress. A total 503 and 776 potential lncRNA-mRNA pairs in *cis*- and *trans*-acting mode, respectively were confidently inferred. Significant enrichment in molecular functions such as plant hormone signal transduction, kinase signaling, and ascorbate/aldarate metabolism characterized the majority of predicted target genes. One serine/threonine-protein kinase SMG1 in particular was predicted to be targeted by multiple lncRNA and thus hypothesized to undepin the drought-tolerant cultivar (Qiu et al., 2019). Similarly, in rapeseed, a comparative transcriptomic analysis of lncRNA changes between drought-tolerant and drought-sensitive cultivars responding to water deficit stress has also been performed (Tan et al., 2020). LncRNA-mRNA interaction network analysis identified 34 TFs corresponding to 126 DE lncRNAs in the drought-tolerant cultivar, and 45 TFs corresponding to 359 DE lncRNAs in the drought-sensitive ones These increased understanding of plant lncRNAs in response to drought stress holds significant value for advancing the investigation of lncRNA functionality and underlying processes in the context of drought stress.

The activation of stress-responsive genes is facilitated by both ABA and dehydration-induced primary and secondary signals, which operate through a complex network of ABA-dependent and ABA-independent signaling pathways (Figure 3). This network encompasses various processes, including sensing, signal transduction, and gene activation (Xiong et al., 2002; Zhu, 2016). In Arabidopsis, *DROUGHT INDUCED lncRNA* (*DRIR*), is specifically and strongly induced by drought, salt stress and ABA treatment, while remaining lowly expressed level under non-stress conditions (Qin et al., 2017). The T-DNA insertion mutant $drir^D$ and transgenic *DRIR* overexpressing lines in Arabidopsis showed that *DRIR* exerts a positive role by restricting transpirational water loss and enhancing drought tolerance. Furthermore, the $drir^D$ mutant and the OE-seedlings exhibited heightened sensitivity to ABA compared to the wild type (WT). Transcriptome analysis revealed significant modulation of genes associated with ABA signaling, water transport, and other stress-relief processes in both the $drir^D$ and the DRIR overexpressing lines (Qin et al., 2017).

The potato CYCLING DOF FACTOR 1, *StCDF1* in conjunction with its lncRNA counterpart, *StFLORE* (StCDF1 locus encodes for an antisense lncRNA), plays a role in the regulation of water loss by influencing stomatal development and diurnal opening (Ramirez et al., 2021). Both naturally occurring and CRISPR-Cas9 mutations in the *StFLORE* transcript result in plants that exhibit heightened sensitivity to water-limiting situations. Conversely, a higher level of *StFLORE* expression achieved either through the overexpression of *StFLORE* or the downregulation of *StCDF1*, leads to enhanced drought tolerance by mitigating water loss. Taken together, the *StCDF1–StFLORE* locus plays a significant role in both vegetative reproduction and water homeostasis (Ramirez et al., 2021). In *Populus trichocarpa*, drought-responsive *lincRNA2752* could reduce the level of *ptc-miR169* by acting as a target mimic of *ptc-miR169* (Shuai et al., 2014). *MiR169* is known to play an important role in drought stress by regulating TF NF-YA in plants (Ni et al., 2013). Therefore, the regulation network of *lincrna2752* on drought resistance may involve both *miR169*- and NF-YA-associated mechanisms. In maize, tissue- and developmental-stage spatio-temporal transcriptional dynamics of drought stress-responsive

lncRNAs have also been investigated (Pang et al., 2019). Interestingly, the developmental stage showing the most drought-responsive lncRNA DE involves the transition from the vegetative stage to the reproductive stage. New insights into lncRNA-mediated gene regulation in drought response in this system were also gained. Particularly, one lncRNA-mRNA pair involving - *vpp4* (encoding a vacuolar (H^+)-pumping ATPase subunit homolog) and its adjacent lncRNA *MSTRG.6838.1* sharing strong expression correlation was prioritized as one of many mechanisms underpinning the drought response (Pang et al., 2019).

LncRNAs response to salt stress in plants

Salt stress is also a prominent environmental factor that has a significant impact on the growth and development of plants (Zhu, 2016; Gong et al., 2020; Zhang et al., 2022a). In *Populus trichocarpa*, multi-tissue (leaves, stems and roots) RNA-seq enabled the identification of 1183 salt-induced DE lncRNAs (Ye et al. (2022). One lncRNA in particular, *Ptlinc-NAC72* plays a negative role in the salt tolerance of this species. When exposed to long-term salt stress, *Ptlinc-NAC72* exhibited the capacity to cis- and trans-regulate salt-responsive genes. For example, *Ptlinc-NAC72* can directly upregulate *PtNAC72.A/B* expression by recognizing the GAAAAA tandem elements in the 5'-UTR of *PtNAC72.A/B*. Induction of *PtNAC72.A/B* in turn confers resilience of the plant to salt stress (Ye et al., 2022). In arabidopsis, *npc536* was identified as a salt-responsive lncRNA, which is observed to be upregulated in both roots and leaves when subjected to salt treatment. Furthermore, the overexpression of *npc536* has been found to promote root growth under conditions of salt stress, presumably through its involvement in regulating the translation of *CONSERVED OLIGOMERIC GOLGI COMPLEX 5* (*AtCOG5*) mRNA (Ben Amor et al., 2009).

In upland cotton, endogenous target mimic (eTM) analysis indicates that *LncRNA354* had a potential binding site for *miR160b* (Zhang et al., 2021a). Overexpression of lncRNA354 in transgenic Arabidopsis plants reduced tolerance to salt stress by affecting the expression of *miR160b* and target auxin responsive factors (*ARFs*). The transgenic plants overexpressing *GhARF17/18* target genes also shared this salt-sensitive phenotype. On the contrary, silencing *lncRNA354* and targets *GhARF17/18* showed taller plants and enhanced the tolerance to salt stress (Zhang et al., 2021a). In addition to its role in drought tolerance, Arabidopsis lncRNA *DRIR* is also salt-stress inducible and confers salinity stress tolerance (Qin et al., 2017). In salt-treated $drir^D$ and DRIR overexpressing seedlings, the expression of stress-related genes like *P5CS1*, *RD29A*, *RD29B*,and *NAC3* increased dramatically. Unlike drought conditions, no ABA signaling pathway genes (like *ABI5*) were found indicating that *DRIR* regulates stress-related genes through different mechanisms during specific stress conditions (Qin et al., 2017). In cotton, salt-stress inducible *lncRNA973* increased salinity stress tolerance by regulating the expression of *miR399* and its target gene *GhPHO2*. Knocking out *lncRNA973* in cotton seedling roots significantly increased miR399 expression (Zhang et al., 2019).

LncRNAs response to biotic stress in plants

The involvement of lncRNAs in various abiotic stress responses has been well-documented, while their contributions to biotic stress response are gradually emerging. In wheat, Xin et al. (2011) identified 125 lncRNAs during powdery mildew infection and heat stress. Among them, 48 lncRNAs showed responsiveness only to powdery mildew infection (designated TapmlnRNA.) while 23 lncRNAs demonstrated responsiveness to both powdery mildew infection and heat stress (designated TalnRNA). Four lncRNAs (*TalnRNA5*, *TapmlnRNA8*, *TapmlnRNA19*, and *TahlnRNA27*) were shown to form stable hairpin structures and were identified as putative precursors of miRNA. Among them, *TapmlnRNA8* and *TapmlnRNA19* showed specificity toward powdery mildew infection (Xin et al., 2011). More recently, disease-responsive (e.g. to Rice Black-Streaked Dwarf Virus and Rice Stripe Virus) lncRNA profile was investigated in rice (Cao et al., 2022). A total of 21 lncRNAs were found to be up-regulated in response to both virus and putative targets including an excess of 1,000 co-regulated transcripts involved in transcriptional regulation, plant hormone signal transduction, phenylpropanoid biosynthesis, and plant-pathogen interaction terms (Cao et al., 2022).

In tomato plant, predicted ceRNA networks involving *Phytophthora infestans* resistance have been presented (Cui et al., 2020). One lncRNA in particular, *LncRNA40787* was strongly up-regulated upon infection with *P. infestans*. Additionally, *LncRNA40787* was predicted as a ceRNAs of *miR394* as it contains the endogenous target mimics (eTM) structure of *miR394*. Follow-up studies revealed miR394 plays a negative role in tomato resistance against *P. infestans* whereas lncRNA40787 positively regulates tomato defense response by acting as an eTM for miR394 to regulate *Leaf Curling Responsiveness* (*LCR*) by activating genes involved in JA biosynthesis components (Zhang et al., 2021b). A ceRNA network consisting of 5 lncRNAs, 5 miRNAs, 4 circRNAs, and 15 mRNAs has also been constructed to provide the candidate ceRNAs in *Paulownia* and an in-depth understanding of the molecular mechanism responsible for Paulownia witches' broom disease progression (Fan et al., 2018).

Recently, a pathogen infection-responsive lncRNA, termed salicylic acid biogenesis controller 1 (*SABC1*), was discovered to fine-tune salicylic acid biosynthesis, which in turn balances plant immunity and growth (Liu et al., 2022). CURLY LEAF (CLF), an essential component of polycomb repressive complex 2 (PRC2) responsible for catalyzing H3K27me3, was identified in *SABC1* RNA immunoprecipitation. Further molecular mechanism analysis indicates *SABC1* associates with the CLF protein and assists the recruitment of CLF-PRC2 to its neighboring NAC TF gene *NAC3* locus, thereby decreasing *NAC3* transcription via H3K27me3. In healthy plants, *SABC1* exerts a suppressive effect on SA production and plant immunity by downregulating the expression of *NAC3* and a key enzyme catalyzing salicylic acid biosynthesis, *isochorismate synthase 1*, however, upon pathogen infection, *SABC1* is downregulated and relieves the derepression of

SABC1-mediated plant immune response (Liu et al., 2022). In cotton, Zhang et al. (2022b) leveraged co-expression network analysis to identify two herbivory-responsive lncRNA hub genes, lncA07 and lncD09, predicted to play a regulatory role in defense against aphid infestation. Functional analysis of these two lncRNA by CRISPR/Cas9 revealed that the phytohormone jasmonic acid was significantly decreased in CRISPR/Cas9 knock-out mutant of *lncD09* and *lncA07*. Furthermore, three candidate genes (*Ghir_A01G022270*, *Ghir_D04G014430*, and *Ghir_A01G022270*) were implicated in the regulation of the JA-mediated signaling pathway, which in turn underpins susceptibility to insect infestation (Zhang et al., 2022b).

Conclusions and perspectives

The advent of next-generation sequencing technologies and complex bioinformatics analysis have enabled the identification of diverse lncRNAs and their predicted target genes in model plants to emerging crops but also those involved in various abiotic and biotic stress. Coupled with the confirmation of predicted lncRNA function in various plant systems using various reverse genetics approaches exceptional insights for the study of the complex lncRNA-mediated regulatory network controlling plant response tolerance have been gained. Here, we have highlighted a significant body of literature demonstrating the key role of lncRNAs in regulating various aspects of plant stress (abiotic and biotic stress), including stress perception, signal transduction, and downstream responsive pathways. These lncRNAs play multiple roles in plant stress through the different mechanistic strategies (as enhancers, decoys, scaffolds, and/or guides) that are tightly integrated into various plant stress response networks, including the regulation of transcription machinery and chromatin modification in promoter regions, modulation of TF-binding affinity in enhancer regions, RNA–DNA hybridization and modulation of chromatin topological structures, among others. Moreover, many conserved sequence motifs and structural motifs were found to be enriched in stress-responsive lncRNAs, suggesting their potential involvement in the stress-responsive functions of lncRNAs, however, detailed molecular mechanisms of stress regulation by these plant lncRNAs are very limited. Presently, validation of lncRNA function primarily relies on overexpression techniques, nonetheless, CRISPR/Cas holds great potential as an efficient and flexible method for investigating the function of lncRNAs (e.g., by creating functional loss- and/or gain-of-function mutants). It is anticipated that future investigations on lncRNA will continue to contribute valuable knowledge regarding plant stress along with viable strategies for the improvement of crop stress tolerance. With the ultimate goal to ensure global food security for future generations, it will be exciting to know whether lncRNAs hold the key to achieving a harmonious equilibrium between conferring resistance (i.e. an adequate level of stress response) and ensuring crop development, yield, and quality are not compromised.

Author contributions

XJ: Writing – original draft, Writing – review & editing. ZW: Writing – original draft, Writing – review & editing. XL: Writing – original draft. QA: Writing – original draft. DW: Writing – review & editing. FZ: Writing – review & editing. JY: Writing – review & editing. NZ: Writing – review & editing. HS: Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by The Scientific Research Start-up Fund by Gansu Agricultural University (GAU-KYQD-2019-06 and GAU-KYQD-2020-26); The National Natural Science Foundation of China (32260518); Technology Fund by Science and Technology Department of Gansu Province (21JR7RA844 and 22JR5RA882) and The Research Program Sponsored by State Key Laboratory of Aridland Crop Science, Gansu Agricultural University (No. GSCS-2022-03 and GSCS-2020-07).

Conflict of interest

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

Publisher's note

References

Ben Amor, B., Wirth, S., Merchan, F., Laporte, P., d'Aubenton-Carafa, Y., Hirsch, J., et al. (2009). Novel long non-protein coding RNAs involved in Arabidopsis differentiation and stress responses. *Genome Res.* 19 (1), 57–69. doi: 10.1101/gr.080275.108

Bhatia, G., Singh, A., Verma, D., Sharma, S., and Singh, K. (2020). Genome-wide investigation of regulatory roles of lncRNAs in response to heat and drought stress in Brassica juncea (Indian mustard). *Environ. Exp. Bot.* 171, 103922e. doi: 10.1016/j.envexpbot.2019.103922

Calixto, C. P. G., Tzioutziou, N. A., James, A. B., Hornyik, C., Guo, W., Zhang, R., et al. (2019). Cold-dependent expression and alternative splicing of arabidopsis long non-coding RNAs. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.00235

Cao, W., Cao, J., Gao, J., Wang, R., Li, Y., Li, C., et al. (2022). Genome-wide identification and association analysis for virus-responsive lncRNAs in rice (Oryza sativa L.). *Plant Growth Regul.* 98 (1), 65–76. doi: 10.1007/s10725-022-00833-w

Chen, B., Feder, M. E., and Kang, L. (2018). Evolution of heat-shock protein expression underlying adaptive responses to environmental stress. *Mol. Ecol.* 27 (15), 3040–3054. doi: 10.1111/mec.14769

Chen, D., Yuan, C., Zhang, J., Zhang, Z., Bai, L., Meng, Y., et al. (2012). PlantNATsDB: a comprehensive database of plant natural antisense transcripts. *Nucleic Acids Res.* 40 (Database issue), D1187–D1193. doi: 10.1093/nar/gkr823

Chen, J., Zhong, Y., and Qi, X. (2021). LncRNA TCONS_00021861 is functionally associated with drought tolerance in rice (Oryza sativa L.) via competing endogenous RNA regulation. *BMC Plant Biol.* 21 (410), 1–12. doi: 10.1186/s12870-021-03195-z

Cho, J., Koo, D.-H., Nam, Y.-W., Han, C.-T., Lim, H.-T., Bang, J.-W., et al. (2006). Isolation and characterization of cDNA clones expressed under male sex expression conditions in a monoecious cucumber plant (Cucumis sativus L. cv. Winter Long). *Euphytica* 146 (3), 271–281. doi: 10.1007/s10681-005-9023-1

Csorba, T., Questa, J. I., Sun, Q., and Dean, C. (2014). Antisense COOLAIR mediates the coordinated switching of chromatin states at FLC during vernalization. *Proc. Natl. Acad. Sci. United States America* 111 (45), 16160–16165. doi: 10.1073/pnas.1419030111

Cui, J., Jiang, N., Hou, X., Wu, S., Zhang, Q., Meng, J., et al. (2020). Genome-Wide Identification of lncRNAs and Analysis of ceRNA Networks During Tomato Resistance to Phytophthora infestans. *Phytopathology* 110 (2), 456–464. doi: 10.1094/PHYTO-04-19-0137-R

Di, C., Yuan, J., Wu, Y., Li, J., Lin, H., Hu, L., et al. (2014). Characterization of stress-responsive lncRNAs in Arabidopsis thaliana by integrating expression, epigenetic and structural features. *Plant J.* 80 (5), 848–861. doi: 10.1111/tpj.12679

Ding, Y., and Yang, S. (2022). Surviving and thriving: How plants perceive and respond to temperature stress. *Dev. Cell* 57 (8), 947–958. doi: 10.1016/j.devcel.2022.03.010

Fan, G., Wang, Z., Zhai, X., and Cao, Y. (2018). ceRNA cross-talk in paulownia witches' Broom disease. *Int. J. Mol. Sci.* 19, (8). doi: 10.3390/ijms19082463

Gong, H., You, J., Zhang, X., Liu, Y., Zhao, F., Cui, X., et al. (2021). Genome-wide identification and functional analysis of long non-coding RNAs in sesame response to salt stress. *J. Plant Biol.* 64 (6), 555–565. doi: 10.1007/s12374-021-09324-3

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

Gupta, A., Rico-Medina, A., and Caño-Delgado, A. I. (2020). The physiology of plant responses to drought. *Science* 368 (6488), 266. doi: 10.1126/science.aaz7614

He, X., Guo, S., Wang, Y., Wang, L., Shu, S., and Sun, J. (2020). Systematic identification and analysis of heat-stress-responsive lncRNAs, circRNAs and miRNAs with associated co-expression and ceRNA networks in cucumber (Cucumis sativus L.). *Physiologia Plantarum* 168 (3), 736–754. doi: 10.1111/ppl.12997

Heo, J. B., Lee, Y. S., and Sung, S. (2013). Epigenetic regulation by long noncoding RNAs in plants. *Chromosome Res.* 21 (6-7), 685–693. doi: 10.1007/s10577-013-9392-6

Hu, X., Wei, Q., Wu, H., Huang, Y., Peng, X., Han, G., et al. (2022). Identification and characterization of heat-responsive lncRNAs in maize inbred line CM1. *BMC Genomics* 23 (1), 208. doi: 10.1186/s12864-022-08448-1

Huang, L., Dong, H., Zhou, D., Li, M., Liu, Y., Zhang, F., et al. (2018). Systematic identification of long non-coding RNAs during pollen development and fertilization in Brassica rapa. *Plant J.* 96 (1), 203–222. doi: 10.1111/tpj.14016

Kim, D. H., and Sung, S. (2012). Environmentally coordinated epigenetic silencing of FLC by protein and long noncoding RNA components. *Curr. Opin. In Plant Biol.* 15 (1), 51–56. doi: 10.1016/j.pbi.2011.10.004

Kim, D. H., and Sung, S. (2017). Vernalization-triggered intragenic chromatin loop formation by long noncoding RNAs. *Dev. Cell* 40 (3), 302–312 e304. doi: 10.1016/j.devcel.2016.12.021

Kindgren, P., Ard, R., Ivanov, M., and Marquardt, S. (2018). Transcriptional read-through of the long non-coding RNA SVALKA governs plant cold acclimation. *Nat. Commun.* 9 (1), 4561. doi: 10.1038/s41467-018-07010-6

Leng, X., Wang, P., Zhu, X., Li, X., Zheng, T., Shangguan, L., et al. (2017). Ectopic expression of CSD1 and CSD2 targeting genes of miR398 in grapevine is associated with oxidative stress tolerance. *Funct. Integr. Genomics* 17 (6), 697–710. doi: 10.1007/s10142-017-0565-9

Lesk, C., Anderson, W., Rigden, A., Coast, O., Jägermeyr, J., McDermid, S., et al. (2022). Compound heat and moisture extreme impacts on global crop yields under climate change. *Nat. Rev. Earth Environ.* 3 (12), 872–889. doi: 10.1038/s43017-022-00368-8

Li, L., Liu, J., Liang, Q., Zhang, Y., Kang, K., Wang, W., et al. (2021). Genome-wide analysis of long noncoding RNAs affecting floral bud dormancy in pears in response to cold stress. *Tree Physiol.* 41 (5), 771–790. doi: 10.1093/treephys/tpaa147

Li, R., Fu, D., Zhu, B., Luo, Y., and Zhu, H. (2018). CRISPR/Cas9-mediated mutagenesis of lncRNA1459 alters tomato fruit ripening. *Plant J.* 94 (3), 513–524. doi: 10.1111/tpj.13872

Li, S., Cheng, Z., Dong, S., Li, Z., Zou, L., Zhao, P., et al. (2022). Global identification of full-length cassava lncRNAs unveils the role of cold-responsive intergenic lncRNA 1 in cold stress response. *Plant Cell Environ.* 45 (2), 412–426. doi: 10.1111/pce.14236

Li, S., Yu, X., Lei, N., Cheng, Z., Zhao, P., He, Y., et al. (2017). Genome-wide identification and functional prediction of cold and/or drought-responsive lncRNAs in cassava. *Sci. Rep.* 7, 45981. doi: 10.1038/srep45981

Liu, N., Xu, Y., Li, Q., Cao, Y., Yang, D., Liu, S., et al. (2022). A lncRNA fine-tunes salicylic acid biosynthesis to balance plant immunity and growth. *Cell Host Microbe* 30 (8), 1124–1138 e1128. doi: 10.1016/j.chom.2022.07.001

Lu, Q., Guo, F., Xu, Q., and Cang, J. (2020). LncRNA improves cold resistance of winter wheat by interacting with miR398. *Funct. Plant Biol.* 47 (6), 544–557. doi: 10.1071/FP19267

Masocha, V. F., Li, Q., Zhu, Z., Chai, F., Sun, X., Wang, Z., et al. (2020). Proteomic variation in Vitis amurensis and V. vinifera buds during cold acclimation. *Scientia Hortic.* 263, 109143. doi: 10.1016/j.scienta.2019.109143

Moison, M., Martínez Pacheco, J., Lucero, L., Fonouni-Farde, C., Rodríguez-Melo, J., Mansilla, N., et al. (2021). The lncRNA APOLO interacts with the transcription factor WRKY42 to trigger root hair cell expansion in response to cold. *Mol. Plant* 14 (6), 937–948. doi: 10.1101/2020.07.13.188763

Movahedi, A., Sun, W., Zhang, J., Wu, X., Mousavi, M., Mohammadi, K., et al. (2015). RNA-directed DNA methylation in plants. *Plant Cell Rep.* 34 (11), 1857–1862. doi: 10.1007/s00299-015-1839-0

Pang, J., Zhang, X., Ma, X., and Zhao, J. (2019). Spatio-temporal transcriptional dynamics of maize long non-coding RNAs responsive to drought stress. *Genes (Basel)* 10, (2). doi: 10.3390/genes10020138

Qin, T., and Xiong, L. (2019). Subcellular localization and functions of plant lncRNAs in drought and salt stress tolerance. *Methods Mol. Cell. Biol.* 1933, 173–186. doi: 10.1007/978-1-4939-9045-0_9

Qin, T., Zhao, H., Cui, P., Albesher, N., and Xiong, L. (2017). A nucleus-localized long non-coding RNA enhances drought and salt stress tolerance. *Plant Physiol.* 175 (3), 1321–1336. doi: 10.1104/pp.17.00574

Qiu, C.-W., Zhao, J., Chen, Q., and Wu, F. (2019). Genome-wide characterization of drought stress responsive long non-coding RNAs in Tibetan wild barley. *Environ. Exp. Bot.* 164, 124–134. doi: 10.1016/j.envexpbot.2019.05.002

Rai, M. I., Alam, M., Lightfoot, D. A., Gurha, P., and Afzal, A. J. (2019). Classification and experimental identification of plant long non-coding RNAs. *Genomics* 111 (5), 997–1005. doi: 10.1016/j.ygeno.2018.04.014

Ramirez, G. L., Shi, L., Bergonzi, S. B., Oortwijn, M., Franco-Zorrilla, J. M., Solano-Tavira, R., et al. (2021). Potato CYCLING DOF FACTOR 1 and its lncRNA counterpart StFLORE link tuber development and drought response. *Plant J.* 105 (4), 855–869. doi: 10.1111/tpj.15093

Ren, J., Jiang, C., Zhang, H., Shi, X., Ai, X., Li, R., et al. (2022). LncRNA-mediated ceRNA networks provide novel potential biomarkers for peanut drought tolerance. *Physiologia Plantarum* 174 (1), e13610. doi: 10.1111/ppl.13610

Scharf, K.-D., Berberich, T., Ebersberger, I., and Nover, L. (20121819). The plant heat stress transcription factor (Hsf) family: Structure, function and evolution. *Biochim. Biophys. Acta (BBA) - Gene Regul. Mech.* 2), 104–119. doi: 10.1016/j.bbagrm.2011.10.002

Shi, Y., Ding, Y., and Yang, S. (2015). Cold signal transduction and its interplay with phytohormones during cold acclimation. *Plant Cell Physiol.* 56 (1), 7–15. doi: 10.1093/pcp/pcu115

Shigeta, T., Zaizen, Y., Sugimoto, Y., Nakamura, Y., Matsuo, T., and Okamoto, S. (2015). Heat shock protein 90 acts in brassinosteroid signaling through interaction with BES1/BZR1 transcription factor. *J. Plant Physiol.* 178, 69–73. doi: 10.1016/j.jplph.2015.02.003

Shuai, P., Liang, D., Tang, S., Zhang, Z., Ye, C. Y., Su, Y., et al. (2014). Genome-wide identification and functional prediction of novel and drought-responsive lincRNAs in Populus trichocarpa. *J. Exp. Bot.* 65 (17), 4975–4983. doi: 10.1093/jxb/eru256

Song, Y., Chen, P., Liu, P., Bu, C., and Zhang, D. (2020). High-Temperature-Responsive Poplar lncRNAs Modulate Target Gene Expression via RNA Interference and Act as RNA Scaffolds to Enhance Heat Tolerance. *Int. J. Mol. Sci.* 21, (18). doi: 10.3390/ijms21186808

Song, Y., Ci, D., Tian, M., and Zhang, D. (2016b). Stable methylation of a non-coding RNA gene regulates gene expression in response to abiotic stress in Populus simonii. *J. Exp. Bot.* 67 (5), 1477–1492. doi: 10.1093/jxb/erv543

Song, L., Fang, Y., Chen, L., Wang, J., and Chen, X. (2021). Role of non-coding RNAs in plant immunity. *Plant Commun.* 2 (3), 100180. doi: 10.1016/j.xplc.2021.100180

Song, X., Liu, G., Huang, Z., Duan, W., Tan, H., Li, Y., et al. (2016a). Temperature expression patterns of genes and their coexpression with LncRNAs revealed by RNA-Seq in non-heading Chinese cabbage. *BMC Genomics* 17, 297. doi: 10.1186/s12864-016-2625-2

Swiezewski, S., Liu, F., Magusin, A., and Dean, C. (2009). Cold-induced silencing by long antisense transcripts of an Arabidopsis Polycomb target. *Nature* 462 (7274), 799–802. doi: 10.1038/nature08618

Tan, X., Li, S., Hu, L., and Zhang, C. (2020). Genome-wide analysis of long non-coding RNAs (lncRNAs) in two contrasting rapeseed (Brassica napus L.) genotypes subjected to drought stress and re-watering. *BMC Plant Biol.* 20 (1), 81. doi: 10.1186/s12870-020-2286-9

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). doi: 10.3390/ijms22136980

Wang, Z., Chai, F., Zhu, Z., Kirabi Elias, G., Xin, H., Liang, Z., et al. (2020). The inheritance of cold tolerance in seven interspecific grape populations. *Scientia Hortic.* 266, 109260e. doi: 10.1016/j.scienta.2020.109260

Wang, P., Dai, L., Ai, J., Wang, Y., and Ren, F. (2019b). Identification and functional prediction of cold-related (lncRNA) in grapevine. *Sci. Rep.* 9 (1), 6638. doi: 10.1038/s41598-019-43269-5

Wang, A., Hu, J., Gao, C., Chen, G., Wang, B., Lin, C., et al. (2019a). Genome-wide analysis of long non-coding RNAs unveils the regulatory roles in the heat tolerance of Chinese cabbage (Brassica rapa ssp.chinensis). *Sci. Rep.* 9 (1), 5002. doi: 10.1038/s41598-019-41428-2

Wang, J., Meng, X., Dobrovolskaya, O. B., Orlov, Y. L., and Chen, M. (2017). Non-coding RNAs and their roles in stress response in plants. *Genomics Proteomics Bioinf.* 15 (5), 301–312. doi: 10.1016/j.gpb.2017.01.007

Wang, J., Meng, X., Yuan, C., Harrison, A. P., and Chen, M. (2016). The roles of cross-talk epigenetic patterns in Arabidopsis thaliana. *Briefings Funct. Genomics* 15 (4), 278–287. doi: 10.1093/bfgp/elv025

Wierzbicki, A. T., Blevins, T., and Swiezewski, S. (2021). Long noncoding RNAs in plants. *Annu. Rev. Plant Biol.* 72 (4), 245–271. doi: 10.1146/annurev-arplant-093020-035446

Wu, C., Ding, Z., Chen, M., Yang, G., Tie, W., Yan, Y., et al. (2019). Identification and functional prediction of lncRNAs in response to PEG and ABA treatment in cassava. *Environ. Exp. Bot.* 166, 103809e. doi: 10.1016/j.envexpbot.2019.103809

Wu, H. J., Wang, Z. M., Wang, M., and Wang, X. J. (2013). Widespread long noncoding RNAs as endogenous target mimics for microRNAs in plants. *Int. J. Mol. Sci.* 161 (4), 1875–1884. doi: 10.1104/pp.113.215962

Wunderlich, M., Gross-Hardt, R., and Schoffl, F. (2014). Heat shock factor HSFB2a involved in gametophyte development of Arabidopsis thaliana and its expression is controlled by a heat-inducible long non-coding antisense RNA. *Plant Mol. Biol.* 85 (6), 541–550. doi: 10.1007/s11103-014-0202-0

Xin, M., Wang, Y., Yao, Y., Song, N., Hu, Z., Qin, D., et al. (2011). Identification and characterization of wheat long non-protein coding RNAs responsive to powdery mildew infection and heat stress by using microarray analysis and SBS sequencing. *BMC Plant Biol.* 11 (1), 61. doi: 10.1186/1471-2229-11-61

Xiong, L., Schumaker, K. S., and Zhu, J. K. (2002). Cell signaling during cold, drought, and salt stress. *Plant Cell* 14 Suppl, S165–S183. doi: 10.1105/tpc.000596

Yang, X., Liu, C., Niu, X., Wang, L., Li, L., Yuan, Q., et al. (2022). Research on lncRNA related to drought resistance of Shanlan upland rice. *BMC Genomics* 23 (1), 336. doi: 10.1186/s12864-022-08546-0

Ye, X., Wang, S., Zhao, X., Gao, N., Wang, Y., Yang, Y., et al. (2022). Role of lncRNAs in cis- and trans-regulatory responses to salt in Populus trichocarpa. *Plant J.* 110, 978–993. doi: 10.1111/tpj.15714

Zhang, X., Dong, J., Deng, F., Wang, W., Cheng, Y., Song, L., et al. (2019). The long non-coding RNA lncRNA973 is involved in cotton response to salt stress. *BMC Plant Biol.* 19 (1), 459. doi: 10.1186/s12870-019-2088-0

Zhang, Y. Y., Hong, Y. H., Liu, Y. R., Cui, J., and Luan, Y. S. (2021b). Function identification of miR394 in tomato resistance to Phytophthora infestans. *Plant Cell Rep.* 40 (10), 1831–1844. doi: 10.1007/s00299-021-02746-w

Zhang, J., Li, J., Saeed, S., Batchelor, W. D., Alariqi, M., Meng, Q., et al. (2022b). Identification and functional analysis of lncRNA by CRISPR/cas9 during the cotton response to sap-sucking insect infestation. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.784511

Zhang, X., Shen, J., Xu, Q., Dong, J., Song, L., Wang, W., et al. (2021a). Long noncoding RNA lncRNA354 functions as a competing endogenous RNA of miR160b to regulate ARF genes in response to salt stress in upland cotton. *Plant Cell Environ.* 44 (10), 3302–3321. doi: 10.1111/pce.14133

Zhang, H., Zhu, J., Gong, Z., and Zhu, J. K. (2022a). Abiotic stress responses in plants. *Nat. Rev. Genet.* 23 (2), 104–119. doi: 10.1038/s41576-021-00413-0

Zhao, X., Li, J., Lian, B., Gu, H., Li, Y., and Qi, Y. (2018). Global identification of Arabidopsis lncRNAs reveals the regulation of MAF4 by a natural antisense RNA. *Nat. Commun.* 9 (1), 5056. doi: 10.1038/s41467-018-07500-7

Zhu, J. K. (2016). Abiotic stress signaling and responses in plants. *Cell* 167 (2), 313–324. doi: 10.1016/j.cell.2016.08.029

Zhu, B., Yang, Y., Li, R., Fu, D., Wen, L., Luo, Y., et al. (2015). RNA sequencing and functional analysis implicate the regulatory role of long non-coding RNAs in tomato fruit ripening. *J. Exp. Bot.* 66 (15), 4483–4495. doi: 10.1093/jxb/erv203

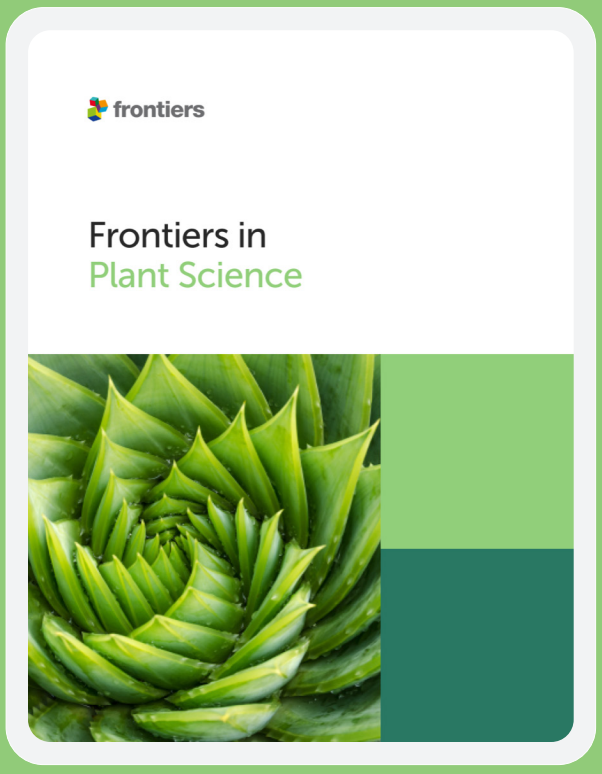
Frontiers in
Plant Science

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