frontiers | Research Topics

# Accelerating genetic gain for key traits using genome-wide association studies and genomic selection: Promising breeding tools for sustainable agriculture

**Edited by**
Sundeep Kumar, Philomin Juliana
and Dwijesh Chandra Mishra

**Coordinated by**
Neeraj Budhlakoti

**Published in**
Frontiers in Genetics
Frontiers in Plant Science

ISSN 1664-8714
ISBN 978-2-8325-4254-5
DOI 10.3389/978-2-8325-4254-5

## About Frontiers

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

## Frontiers journal series

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

## Dedication to quality

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

## What are Frontiers Research Topics?

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

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

# Accelerating genetic gain for key traits using genome-wide association studies and genomic selection: Promising breeding tools for sustainable agriculture

**Topic editors**

Sundeep Kumar – National Bureau of Plant Genetic Resources, Indian Council of Agricultural Research (ICAR), India

Philomin Juliana – International Maize and Wheat Improvement Center, Mexico

Dwijesh Chandra Mishra – Indian Agricultural Statistics Research Institute, Indian Council of Agricultural Research, India

**Topic coordinator**

Neeraj Budhlakoti – Indian Agricultural Statistics Research Institute, Indian Council of Agricultural Research, India

**Citation**

Kumar, S., Juliana, P., Mishra, D. C., Budhlakoti, N., eds. (2024). *Accelerating genetic gain for key traits using genome-wide association studies and genomic selection: Promising breeding tools for sustainable agriculture*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-4254-5

# Table of contents

frontiers | Frontiers in Genetics

TYPE Editorial
PUBLISHED 20 December 2023
DOI 10.3389/fgene.2023.1351870



OPEN ACCESS

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

*CORRESPONDENCE
Sundeep Kumar,
Sundeep.Kumar@icar.gov.in

RECEIVED 07 December 2023
ACCEPTED 08 December 2023
PUBLISHED 20 December 2023

CITATION
Mishra DC, Budhlakoti N, Juliana P and Kumar S (2023), Editorial: Accelerating genetic gain for key traits using genome-wide association studies and genomic selection: promising breeding tools for sustainable agriculture.
*Front. Genet.* 14:1351870.
doi: 10.3389/fgene.2023.1351870



# Editorial: Accelerating genetic gain for key traits using genome-wide association studies and genomic selection: promising breeding tools for sustainable agriculture

Dwijesh Chandra Mishra[1], Neeraj Budhlakoti[1], Philomin Juliana[2] and Sundeep Kumar[3]*

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

KEYWORDS

**GWAS, genetic gain, sustainable agriculture, genomic selection (GS), breeding programs**

Editorial on the Research Topic
Accelerating genetic gain for key traits using genome-wide association studies and genomic selection: promising breeding tools for sustainable agriculture

The pursuit of sustainable agriculture is undergoing a transformative shift through the integration of advanced genomic tools, namely, Genome-Wide Association Studies (GWAS) and Genomic Selection (GS). These tools offer a revolutionary approach to accelerating genetic gain for key traits in crop breeding programs. GWAS, by scrutinizing extensive sets of genetic markers across diverse germplasm, unveils associations between specific genomic regions and trait variations. This not only elucidates the intricate genetic factors governing desired traits but also facilitates the identification of valuable alleles for breeding programs. Complementing GWAS, Genomic Selection leverages sophisticated statistical models and high-throughput genotyping to predict the breeding value of individuals based on their entire genomic information. This predictive power enables breeders to identify superior individuals for traits of interest at an early stage, significantly expediting the breeding cycle. The synergy between GWAS and GS holds immense promise for sustainable agriculture by expediting the development of improved crop varieties, enhancing yields, stress resilience, and resource-use efficiency–crucial components of sustainable agricultural practices. Moreover, the integration of genomics into breeding programs facilitates the adaptation of crops to specific agroecological contexts and evolving climatic conditions, ensuring the resilience of crops in the face of changing climates. This editorial embarks on a nuanced journey through a selection of pioneering studies, each offering a unique lens into the multifaceted world of plant genomics through GWAS and Genomic Selection and its profound implications for the future of agriculture.

The study of *StTGA* genes in potatoes emerges as a beacon of hope in the battle against bacterial wilt, a formidable adversary threatening potato crops worldwide. *StTGA* genes,

identified as a subgroup of bZIP transcription factors, play a crucial role in orchestrating the plant's defense mechanisms against *Ralstonia solanacearum*, the causative agent of bacterial wilt. Tian et al. meticulously analyzed the expression patterns of specific *StTGA* genes during infection, unveiling a sophisticated regulatory network governing bacterial wilt tolerance. The genome-wide studies not only identified key players in the plant's defense response but also laid the groundwork for targeted breeding programs aimed at enhancing the resilience of potato varieties to this devastating pathogen. The implications of this study are far-reaching, holding the potential to revolutionize potato breeding strategies. By understanding the genomic intricacies of bacterial wilt tolerance, researchers and breeders can now work hand in hand to develop potato varieties with enhanced resistance, ensuring global food security in the face of agricultural challenges.

Basmati rice, celebrated for its aromatic grains, stands at the center of an ambitious molecular mapping endeavor. The quest is to unravel the genetic underpinnings of grain dimension traits, an essential aspect of rice quality and market preference. Malik et al. deployed state-of-the-art molecular mapping techniques to pinpoint quantitative trait loci (QTLs) responsible for variations in grain length, width, and thickness. The results unveiled a complex genomic landscape, with multiple QTLs exerting influence over distinct grain dimensions. This newfound knowledge provides a valuable resource for rice breeders seeking to craft varieties with tailored grain characteristics. The implications of this study extend beyond the laboratories and into rice paddies globally. Armed with the genetic roadmap of grain dimension traits, breeders can now accelerate the development of Basmati rice varieties that not only meet but exceed consumer expectations, securing the prominence of this prized rice variety in international markets.

The flag leaf of bread wheat emerges as a focal point in the quest for optimizing grain yield. As the primary site for nutrient remobilization and grain filling, understanding the genetic drivers of flag leaf development is paramount for sustainable wheat production. This study done by Mehla et al. dives deep into the structural and functional nuances of candidate genes associated with different developmental stages of the flag leaf. Transcriptomic analyses coupled with functional genomics shed light on the orchestration of key physiological processes, including photosynthesis, nutrient transport, and senescence. The implications of this research reverberate across wheat fields, offering a nuanced understanding of how genetic factors influence grain yield. By unraveling the intricacies of flag leaf development, researchers pave the way for targeted genetic interventions, enhancing nutrient remobilization and ultimately boosting the yield potential of bread wheat.

Pepper, a staple in culinary landscapes worldwide, takes center stage in a genome-wide exploration of the OVATE FAMILY PROTEIN (OFP) gene family. This family of genes has emerged as a key regulator of plant development and stress responses. Luo et al. conducted an exhaustive analysis, unveiling multiple members of the *OFP* gene family in pepper. Each member exhibited unique expression patterns across different tissues and responded differentially to various stress conditions. Structural analyses provided insights into conserved motifs and domains, offering tantalizing clues about the functional roles of these genes. The implications of this study extend to the realms of stress-tolerant crop development. Understanding the diversity and functions of the *OFP* gene family in pepper provides researchers and breeders with molecular tools to fortify pepper varieties against environmental stresses, ensuring a resilient and productive crop.

Fusarium head blight (FHB), a scourge of wheat crops globally, prompts an intensive investigation into genetic sources of resistance. This study seeks to identify specific loci associated with resistance to FHB, offering a glimmer of hope to wheat farmers battling this devastating disease. Wu et al. undertook a comprehensive analysis, uncovering genetic sources of resistance and specific loci linked to Fusarium head blight resistance in wheat. Molecular markers tightly linked to resistance genes provide a powerful tool for breeders striving to develop FHB-resistant wheat varieties. The implications of this study resonate in wheat fields beset by FHB, providing a roadmap for developing resilient varieties. Armed with molecular markers, breeders can implement targeted selection strategies, fortifying wheat crops against the destructive impact of Fusarium head blight.

The architecture of maize leaves, particularly leaf angle, emerges as a critical determinant of light interception and photosynthetic efficiency. This study employs a combination of genome-wide association studies (GWAS) and meta-analysis to identify candidate loci influencing leaf angle in maize. Wu et al. successfully unearthed specific genomic loci associated with leaf angle variation in maize. These loci, acting as pivotal regulators of leaf architecture, hold the key to optimizing planting density and light capture in maize fields. The implications of this research unfold in maize fields worldwide, offering actionable insights for optimizing plant architecture. Crop management practices and breeding strategies can now be fine-tuned, harnessing the genetic determinants of leaf angle to enhance the overall productivity of maize crops.

Micronutrient deficiencies, particularly in iron (Fe) and zinc (Zn), pose a global health challenge. This study delves into the genetic basis of grain Fe and Zn content in wheat, particularly under drought and heat stress conditions. The genome-wide association study (GWAS) done by Devate et al. uncovered specific genomic regions linked to variations in grain Fe and Zn content under drought and heat stress. Understanding the genetic factors influencing nutrient content provides insights for biofortification strategies to enhance the nutritional value of wheat grains. The implications of this study transcend the field of agriculture, holding promise for addressing global malnutrition. The identification of genomic regions associated with grain Fe and Zn content offers a foundation for breeding wheat varieties with improved nutritional profiles, contributing to efforts to combat nutrient deficiencies.

Sugar beet, a vital sugar-producing crop, faces the challenge of drought stress impacting both yield and sugar content. This study employs a genome-wide association study (GWAS) to uncover the genetic factors contributing to drought tolerance traits during the seedling stage in sugar beet. Li et al. identified specific genomic regions associated with drought tolerance traits in sugar beet. Candidate genes within these regions provide valuable insights into the molecular mechanisms underlying adaptive responses to drought stress, laying the foundation for breeding resilient sugar beet varieties. The implications of this study extend to sugar beet fields grappling with water scarcity. By unraveling the genetic basis of drought tolerance, researchers empower breeders with the

knowledge needed to develop sugar beet varieties capable of thriving in water-limited environments.

Drought stress poses a severe threat to global wheat production. This study explores the application of marker-assisted selection (MAS) to expedite the transfer of drought tolerance quantitative trait loci (QTLs) to elite wheat lines, offering a targeted approach to developing resilient varieties. Sunilkumar et al. successfully employed MAS to transfer identified QTLs for drought tolerance to a promising wheat line. This approach not only accelerates the breeding process but also showcases the potential of molecular markers in developing drought-tolerant wheat varieties. The implications of this research reverberate across wheat-growing regions facing water scarcity. The successful application of MAS provides a blueprint for breeders, offering a precision tool to fortify wheat varieties against the challenges posed by drought stress.

Subtropical regions grapple with the dual challenges of drought and heat stress, significantly impacting wheat yield. This study investigates the genotypic capacity of post-anthesis stem reserve mobilization in wheat, focusing on its role in sustaining yield under stress conditions. The research done by Gurumurthy et al. provides insights into genotypic variations in post-anthesis stem reserve mobilization among different wheat varieties. Understanding the capacity of stem reserves to contribute to grain yield under stress conditions offers valuable information for selecting resilient genotypes for subtropical environments. The implications of this research are particularly relevant for wheat cultivation in subtropical regions facing the brunt of climate change. By identifying wheat varieties with enhanced post-anthesis stem reserve mobilization, researchers and farmers can work towards sustaining grain yield even under challenging climatic conditions.

In conclusion, these studies collectively weave a narrative of profound advancements in plant genomics. Each study unravels layers of genetic complexity, offering insights that range from bolstering crop resilience against diseases to optimizing nutrient content and enhancing yield sustainability. The intricate tapestry of genetic codes governing diverse traits emerges as a dynamic landscape, ripe with possibilities for crop improvement. As researchers decipher the genetic mysteries embedded in plant genomes, the promise of translating these insights into resilient, high-yielding crops becomes increasingly tangible. These advancements not only pave the way for a more sustainable and secure global food supply but also exemplify the collaborative and interdisciplinary nature of contemporary plant science. The journey into plant genomics is an ongoing exploration, promising a future where crops are not just resilient but finely tuned to meet the evolving demands of a changing world.

## Author contributions

DM: Supervision, Writing–original draft. NB: Supervision, Writing–review and editing. PJ: Writing–review and editing. SK: Supervision, Writing–review and 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

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 26 July 2022
doi: 10.3389/fgene.2022.894844

# A Genome-Wide Analysis of *StTGA* Genes Reveals the Critical Role in Enhanced Bacterial Wilt Tolerance in Potato During *Ralstonia solanacearum* Infection

*Tian Tian, Ruimin Yu, Yanyun Suo, Lixiang Cheng, Guizhi Li, Dan Yao, Yanjie Song, Huanjun Wang, Xinyu Li and Gang Gao**

*College of Life Sciences, Shanxi Normal University, Taiyuan, China*

OPEN ACCESS

**Edited by:**
*Dwijesh Chandra Mishra, Indian Council of Agricultural Research, India*

**Reviewed by:**
*Yordan Muhovski, Walloon Agricultural Research Centre, Belgium*
*Paola Gaiero, Universidad de la República, Uruguay*

***Correspondence:**
*Gang Gao*
*ggsxnu@126.com*

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

**Received:** *12 March 2022*
**Accepted:** *06 June 2022*
**Published:** *26 July 2022*

**Citation:**
*Tian T, Yu R, Suo Y, Cheng L, Li G, Yao D, Song Y, Wang H, Li X and Gao G (2022) A Genome-Wide Analysis of StTGA Genes Reveals the Critical Role in Enhanced Bacterial Wilt Tolerance in Potato During Ralstonia solanacearum Infection. Front. Genet. 13:894844. doi: 10.3389/fgene.2022.894844*

*TGA* is one of the members of TGACG sequence-specific binding protein family, which plays a crucial role in the regulated course of hormone synthesis as a stress-responsive transcription factor (TF). Little is known, however, about its implication in response to bacterial wilt disease in potato (*Solanum tuberosum*) caused by *Ralstonia solanacearum*. Here, we performed an *in silico* identification and analysis of the members of the *TGA* family based on the whole genome data of potato. In total, 42 *StTGAs* were predicted to be distributed on four chromosomes in potato genome. Phylogenetic analysis showed that the proteins of *StTGAs* could be divided into six sub-families. We found that many of these genes have more than one exon according to the conserved motif and gene structure analysis. The heat map inferred that *StTGAs* are generally expressed in different tissues which are at different stages of development. Genomic collinear analysis showed that there are homologous relationships among potato, tomato, pepper, Arabidopsis, and tobacco *TGA* genes. Cis-element *in silico* analysis predicted that there may be many cis-acting elements related to abiotic and biotic stress upstream of *StTGA* promoter including plant hormone response elements. A representative member *StTGA39* was selected to investigate the potential function of the *StTGA* genes for further analysis. Quantitative real-time polymerase chain reaction (qRT-PCR) assays indicated that the expression of the *StTGAs* was significantly induced by *R. solanacearum* infection and upregulated by exogenous salicylic acid (SA), abscisic acid (ABA), gibberellin 3 ($GA_3$), and methyl jasmonate (MeJA). The results of yeast one-hybrid (Y1H) assay showed that *StTGA39* regulates *S. tuberosum* BRI1-associated receptor kinase 1 (*StBAK1*) expression. Thus, our study provides a theoretical basis for further research of the molecular mechanism of the *StTGA* gene of potato tolerance to bacterial wilt.

**Keywords: potato, *StTGA*, qRT-PCR, MDA, Y1H**

## INTRODUCTION

Plants face a variety of challenges and stresses during growth and development. Specific transcription factors (TF) in plants can specifically regulate the expression of plant genes to enhance the ability of plants to adapt to stress environment. In plants, the gene regulatory network which TFs participate in plays a crucial role in the stress response pathway (Li B. et al., 2019). The members of the *basic Region-Leucine Zipper* (*bZIP*) gene family are crucial regulators, which play an important role in development, stress adaptation, and hormone synthesis of plants (Azeem et al., 2020; Kumar et al., 2021). The TGACG sequence-specific binding protein family (TGA) TFs belong to a subfamily of bZIP and play an important role in abiotic stress responses during the period of plant growth (Ke et al., 2022). The TGA TFs include large regulatory regions, which may affect DNA-binding by allosteric or electrostatic interaction as for the bZIP (Salladini et al., 2020).

Some *TGA* members play crucial roles in stress mitigation. Evidence revealed that the role of *TGA* class II is important during the response of Arabidopsis to control reactive oxygen species (ROS) levels (Ariel et al., 2021). In addition, there was evidence that $Cr^{6+}$ can cause the binding of TGA3 to the L-cysteine desulfhydrase (LCD) promoter in plants, thus increasing the expression of LCD (Fang et al., 2017). Furthermore, glutaredoxins (GRXs) may play a crucial role in the formation of floral organs with *TGA* (Rouhier et al., 2015). ROXYs (GRX,CC-type glutaredoxin, named ROXYs in *Arabidopsis thaliana*) and CC-type GRXs from rice and corn interact with the TGA family to control developmental processes (Gutsche et al., 2017; Uhrig et al., 2017).

An important role of *TGAs* during plant development has been covered in a variety of plants. It can be seen that the *TGA* gene family participates in many signaling pathways in plants and plays a crucial role during plant growth. In soybean, the *TGA* plays a crucial role in response to nitrogen availability (Ullah et al., 2019). Biosynthesis of salicylic acid (SA) is indirectly modulated by TGA1 and TGA4 TF (Budimir, et al., 2020). Nonexpresser of pathogenesis-related genes (NPR1) recruits TGA TFs in the presence of SA and facilitates gene expression to establish plant immunity (Chen et al., 2019). Multiple TGA members have been shown to play a key role in plant immunity. *Oryza sativa* TGA2 (OsTGA2) can directly regulate defense-related genes (Moon et al., 2018). The key points in the regulation of sunflower resistance by *TGA* may be involved in the resistance of sunflower to *Verticillium dahliae* (Guo et al., 2017). In addition, ROXY can be used as TGA-dependent promoter to control the negative regulators of detoxification genes in *A. thaliana* (Huang et al., 2016). Brassinoidsteroid (BR) induces apoplastic ROS, which activates the TGA2 factor and triggers the metabolism of pesticide residues in tomato (Hou et al., 2019). Taken together, studies of TGA TFs and the TGA family as a whole demonstrate that TGA proteins play important roles in regulating multiple biological processes in plants.

Although there are many studies on *TGA*, no study has been reported in potato (*Solanum tuberosum*), which is an important food crop and widely cultivated and consumed all over the world (Gebhardt, 2016). There are many kinds of major diseases affecting potato production worldwide, and bacterial wilt caused by *Ralstonia solanacearum* is the second one in importance (Barchenger et al., 2022). *R. solanacearum* is a soil-borne, devastating plant pathogen and uses type III effectors to inhibit the plant immune system (Salanoubat et al., 2002; Qi et al., 2022). This pathogen has a remarkably wide host range and global distribution and is involved in the invasion of plant root to the vessel of xylem and may eventually lead to plant death (Gutsche and Zachgo, 2016; Ferreira et al., 2017).

In the current study, a comprehensive investigation of TGA TFs in potato was conducted, and the analysis of the member distribution, evolutionary model, gene structure, and expression patterns was performed. In addition, we found a representative member *S. tuberosum TGA39 (StTGA39)* that could be induced by abiotic/biotic stress and has a critical role in enhanced bacterial wilt tolerance in potato during *R. solanacearum* infection. This will lay a foundation for further research on the function of *StTGA* genes.

## MATERIALS AND METHODS

### Identification and Classification of *StTGA* Gene Family

*StTGA* gene identification in the potato genome was performed using BLAST and hidden Markov models (HMM) search methods (Wang et al., 2018). We downloaded the potato genome from the Phytozome database. Briefly, the HMM seed file of delay of germination 1 (DOG1) domain (accession number: PF14144), which belongs to TGA, was downloaded from the Pfam database (http://pfam.xfam.org/). A round of HMM scan was performed for all the obtained hits against the Pfam database. We used the HMMER program to search all putative TGA protein sequences and extracted the corresponding sequence IDs, with expectation value (E-value) set to 1.0. The sequences IDs were submitted to the Spud DB (http://solanaceae.plantbiology.msu.edu/). The results from the two methods were compared, and common sequence IDs were selected. In total, 42 members of the *StTGA* gene family in potato were predicted and named *StTGA01–StTGA42*. The information about the renaming *StTGA* genes in potato is provided in **Supplementary Additional File S1**. The physicochemical parameters, such as molecular weight and isoelectric points (pIs), and subcellular localization of each StTGA protein were predicted *via* ExPASy (http://web.expasy.org/protparam/) and PSORT (https://wolfpsort.hgc.jp/) online tools (Liang et al., 2017).

### Phylogenetic Analysis, Conserved Motif Analysis, *In Silico* Chromosome Mapping, and Gene Structure Analysis of *StTGA*

We constructed the phylogenetic tree by MEGA 7.0 *via* the neighbor-joining (NJ) method, using 1000 bootstrap iterations and default parameters (Gao et al., 2018). The potato gff3. annotation file was parsed to extract the genome locations of

the identified *TGA* genes. According to this information, we visualized the *in silico* predicted chromosomal distribution of *StTGA* genes by TBtools software (Chen et al., 2020). The *in silico* exon–intron distribution of *StTGA* genes was predicted through the Gene Structure Display Server (GSDS) website (http://gsds.gao-lab.org/index.php). Conserved motifs of *StTGA* genes were predicted by the online service Multiple Em for Motif Elicitation (MEME) (Fan et al., 2021a). The information of *in silico* chromosome location of *StTGA* was retrieved from potato genome data downloaded from the phytozome database (Chen et al., 2020). The secondary and 3D structure prediction of StTGA proteins was performed using SOPMA and I-TASSER web servers, and the data were captured using Discovery Studio 4.5 (Geourjon and Deléage, 1995; Ambrish et al., 2010). The upstream 2000 bp region of the initiation codon ATG of each *StTGA* gene was used to search the promoter region in the downloaded sequence and predicted potential cis-acting elements of each binding site on PlantCARE website (Lescot et al., 2002; Zhu et al., 2020).

## Collinearity Analysis of *StTGAs*

Pair-wise all-against-all BLAST was performed for potato, tomato, tobacco, Arabidopsis, and pepper protein sequences. The obtained results and the gff3. annotation file were then determined by the Multiple Collinearity Scan toolkit (MCScanX) for determination of the gene duplication type (Wang et al., 2012; Fan et al., 2021b). Microsynteny relationships between potato and the other four species were analyzed to show the gene homology relationship (Gutsche and Zachgo, 2016), and visualized by TBtools (Chen et al., 2020).

## RNA-Seq Analysis of *StTGAs*

Based on the digital expression RNA-Seq data retrieved by previous methods (Cao et al., 2021), the expression level of *StTGAs* in different tissues and developmental stages under abiotic/biotic stress was indicated using Fragments Per Kilobase of transcript sequence per Millions base pairs (FPKM), which were retrieved through the website Spud DB (http://solanaceae.plantbiology.msu.edu/) (Cao et al., 2021). The heatmaps of *StTGA* expression and hierarchical clustering analysis were conducted using the TBtools software.

Furthermore, the co-expression pattern of *StTGA* genes was analyzed based on the Pearson correlation coefficient (PCC) and graphically presented using the Cytoscape package. The PCC and mutual rank (MR) were calculated according to previous reports (Da et al., 2019). Forty-two potato StTGA differently expressed genes and 24 related genes were analyzed by co-expression network analysis (Renamed in **Supplementary Additional File S2**). The interaction network of differentially expressed TFs was built by STRING: functional protein association networks (https://string-db.org/) and determined by Cytoscape (Kohl et al., 2011; He et al., 2022). When the PCC was greater than 0.8, it was assumed that they were co-expressed.

## Plant Material and Growth Conditions

Potato plants (ZHONG 3,2 $n = 48$, tetraploid cultivation) used in this study were provided by the Chinese Academy of Agricultural Sciences (CAAS). The pots which were used to culture these plants were 10 cm in diameter and 15 cm in height, and they contained 450 ml peat and vermiculite (3:1, volume ratio). In the process of culture, *S. tuberosum* (ZHONG 3) wild-type plants were grown under long-day conditions (16 h, 26°C, day/8 h, 18°C, night) on the abovementioned mixed soil. The light intensity was about 3000 lx, and the relative humidity was 75% (Gao et al., 2010; Yu et al., 2021a).

## Bacterial Strains and Inoculum

*R. solanacearum* belongs to PO41 strain (seed type II, race 3 biotype 2) (He et al., 1983), which was cultured in a CPG medium, and then the solution was diluted to $10^8$ cfu/ml (Optical Density, OD600 = 0.2). Potatoes were inoculated by the root injury irrigation method when they grew to the stage of two-week-old potato plants with 7–8 leaves (Yu et al., 2021b). The control group was inoculated with the same amount of water. At 12, 24, 36, 48, 60, 72, 84, 96, 108, and 120 hpi after inoculation, we sampled the leaves of the experimental group and control group. All samples were frozen in liquid nitrogen immediately and then stored at −80°C (Kong et al., 2016). Three replicates were performed for both *R. solanacearum* and mock inoculation.

## Hormone Treatment

When potatoes grew to the stage of two weeks with 7–8 leaves, the leaves were sprayed with 100 μM abscisic acid (ABA), 350 μM, gibberellin 3 ($GA_3$), 50 μM SA, and 50 μM methyl jasmonate (MeJA) (Yu et al., 2021a). At 1, 2, 3, 4, and 5 dpi, we sampled the leaves of each hormone treatment. All samples were frozen in liquid nitrogen immediately and then stored at −80°C until these samples were used to extract RNA (Fan et al., 2021a). Each treatment was repeated three times.

## Verification of RNA-Seq Data

The RNA-seq data of the *StTGA39* gene were verified by qRT-PCR according to the methods of Schmittgen and Livak, 2008. The leaf samples for RNA extraction were previously stored at −80°C, and the total RNA was isolated using the TRIzol reagent (Invitrogen, Carlsbad, CA, United States) according to the manufacturer's protocol. The mRNA was re-transcribed into cDNA using the Prime Script cDNA Synthesis Kit (TransGen, Beijing, China). Then, cDNA was used as a template for qRT-PCR, which was performed on the Quant Studio 3 qRT-PCR System (Thermo Fisher Scientific, Shanghai, China) (Kong et al., 2016). The following primers were used in the process of qRT-PCR with three replicates for each gene as described previously: StTGA-FP: TCC AGCACATCCAACACC; StTGA-RP: TTCACCAAGATTTCC CAC. ACTIN of potato served as the internal control (GenBank Accession: X55747), and the following primers were used in the process of qRT-PCR with three replicates for each gene as described previously: Actin-F: TATAACGAGCTTCGTGTTGCAC; Actin-R: ACTGGCATACAGCGAAAGAACA.

## Staining Treatment and Quantifying Levels of Malondialdehyde

After the potato seedlings were inoculated with *R. solanacearum* or water for 72 hpi, the leaves were collected for observing the

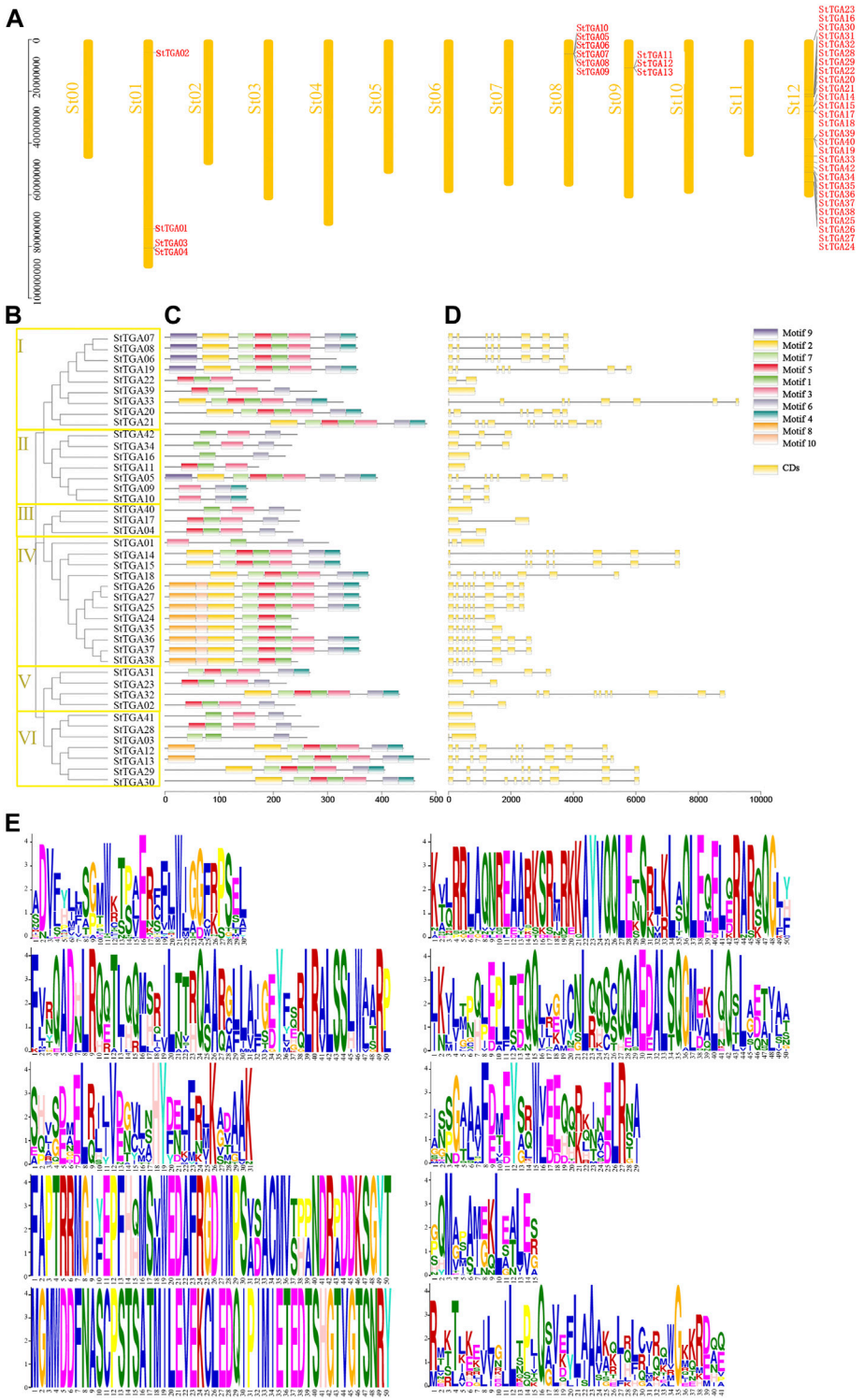



**FIGURE 1 |** Chromosomal location, phylogenetic relationship, exon–intron structure, and conserved motif analysis of StTGA protein. **(A)** Chromosomal locations of StTGA genes. The gene names of StTGA are shown in red to the right of the chromosome. **(B)** Phylogenetic tree of potato StTGA based on its amino acid sequence. Branch lines with different colors indicated different subgroups. The proteins on the tree can be divided into six distinct subfamilies, which are indicated by different colored backgrounds. **(C)** Conserved motif of StTGA proteins analyzed by online program MEME server. Different colored boxes indicated different motifs. **(D)** Gene structure of *StTGA*. The exon–intron structure of *StTGA* genes visualized by online tool GSDS 2.0, yellow boxes indicated exons, and black lines indicated introns. **(E)** Amino acid composition of 10 conserved motifs.

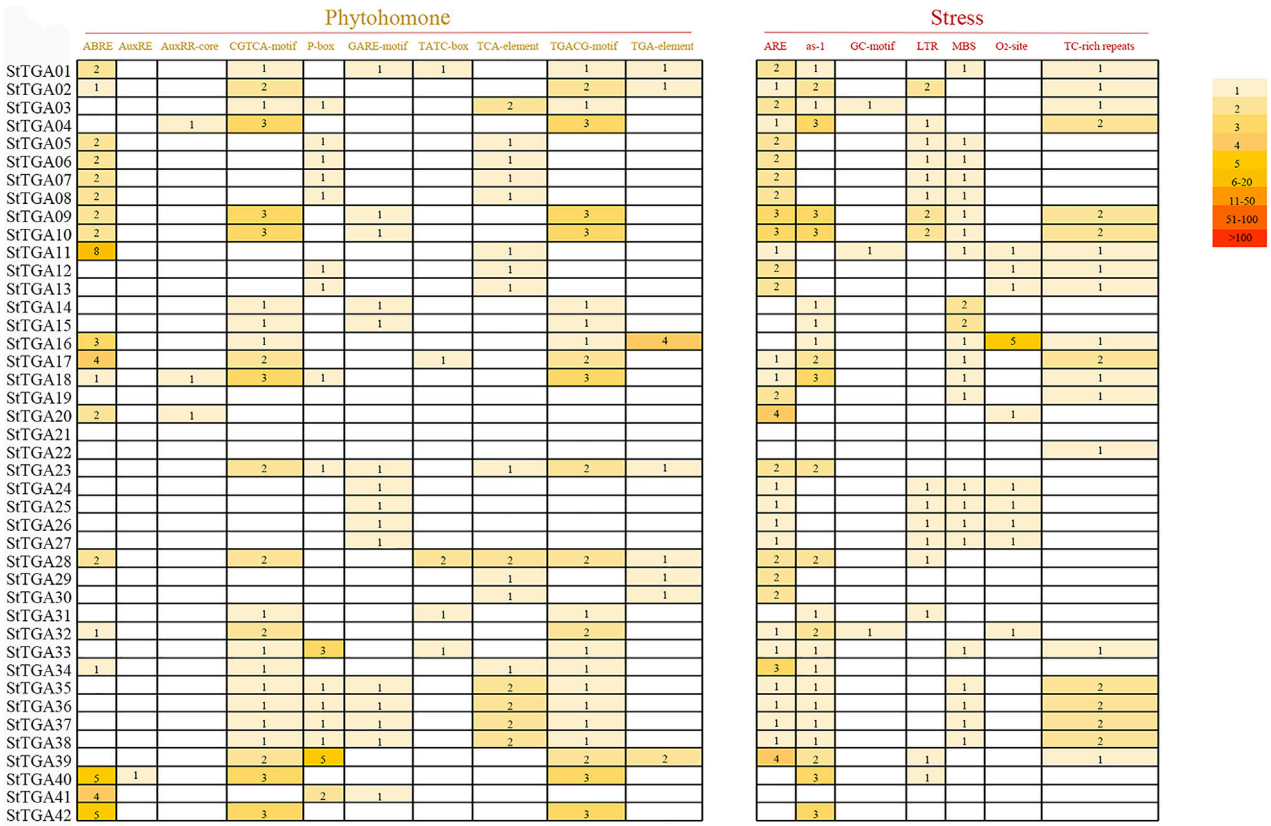



**FIGURE 2 |** Cis-acting elements analysis in the promoter regions of *StTGA*. The promoter cis-elements were analyzed by PlantCARE, and the number of ciselements was searched in the 2,000 bp region upstream of the translation initiation site of the *StTGA* gene. Cis-elements were divided into two different types: phytohomone and stress. Expressed by different colors according to their number. The darker the color, the higher the frequency of appearance, and the number indicates the number of cis-elements.

infection of *R. solanacearum*. The leaves were left in trypan blue solution for 24 h (Yu et al., 2021b). Levels of $H_2O_2$ were detected using the DAB staining method by the DAB Substrate Kit (Solarbio, Beijing, China). Leaves were soaked in the solution to stain for 3–10 min under dark conditions (Christensen et al., 1997; Song et al., 2021). The leaves were sampled from inoculated plants by *R. solanacearum* at 12, 24, 36, 48, 60, 72, 84, 96, 108, and 120 hpi for quantifying levels of MDA. The content of MDA of these leaves was tested by using an MDA Assay Kit (Abbkine, Wuhan, China) according to the previous method (Islamoglu et al., 2018). OD at 532 and 600 nm was determined by a microplate reader.

## Y1H Assays

In order to analyze whether there is a TGACG cis-element in the upstream promoter sequence of the gene homologous to *S. tuberosum Brassinosteroid insensitive 1 (BRI1)-associated receptor kinase 1 (StBAK1)*, we analyzed the relationship between StTGA1 and *StBAKl* by the Y1H assay Kit (Clontech, State of California, United States) in this study. Briefly, the sequence of *StBAK1* with 1512 bp was amplified from *S. tuberosum* cDNA by gene-specific primers and digested by EcoRI (CAATTC) and MluI (ACGCGT) so that the oligonucleotide sequences containing TGA cis-elements and other corresponding promoter sequences were cloned into the pHIS2 to generate the reporter vector. All of the constructs were co-transformed into yeast strain Y187 according to the manufacturer's instructions (Liu et al., 2016). The full-length open-reading frame (ORF) of StTGA (855bp) was fused to the GAL4 activation domain in the vector pGADT7 digested with EcoRI (CAATTC) and BamHI (GGATCC) to get a fusion protein TGA-pGADT7 (effector vector) (Huang et al., 2010). The others were carried out following the manufacturer's instructions.

## Statistical Analysis

Standard values and standard error of experimental data were calculated using Microsoft Excel 2016 and $n = 3$ for independent experiments according to a *t*-test. Analyses of the significance of differences were conducted *via* the data processing System. The PCC is calculated by the SPSS online services tool (https://spssau.com/indexs.html). PCR efficiency was estimated from the data obtained from the exponential phase of each individual amplification plot. The expression level of each gene of interest (GOI) is presented as $2^{-\Delta\Delta Ct}$; where $\Delta\Delta Ct = \Delta Ct_{GOI}$-$\Delta Ct_{Contorl}$; $\Delta Ct_{GOI} = Ct_{GOI}$-$Ct_{Actin}$; $\Delta Ct_{Contorl} = Ct_{Control}$-$Ct_{Actin}$.

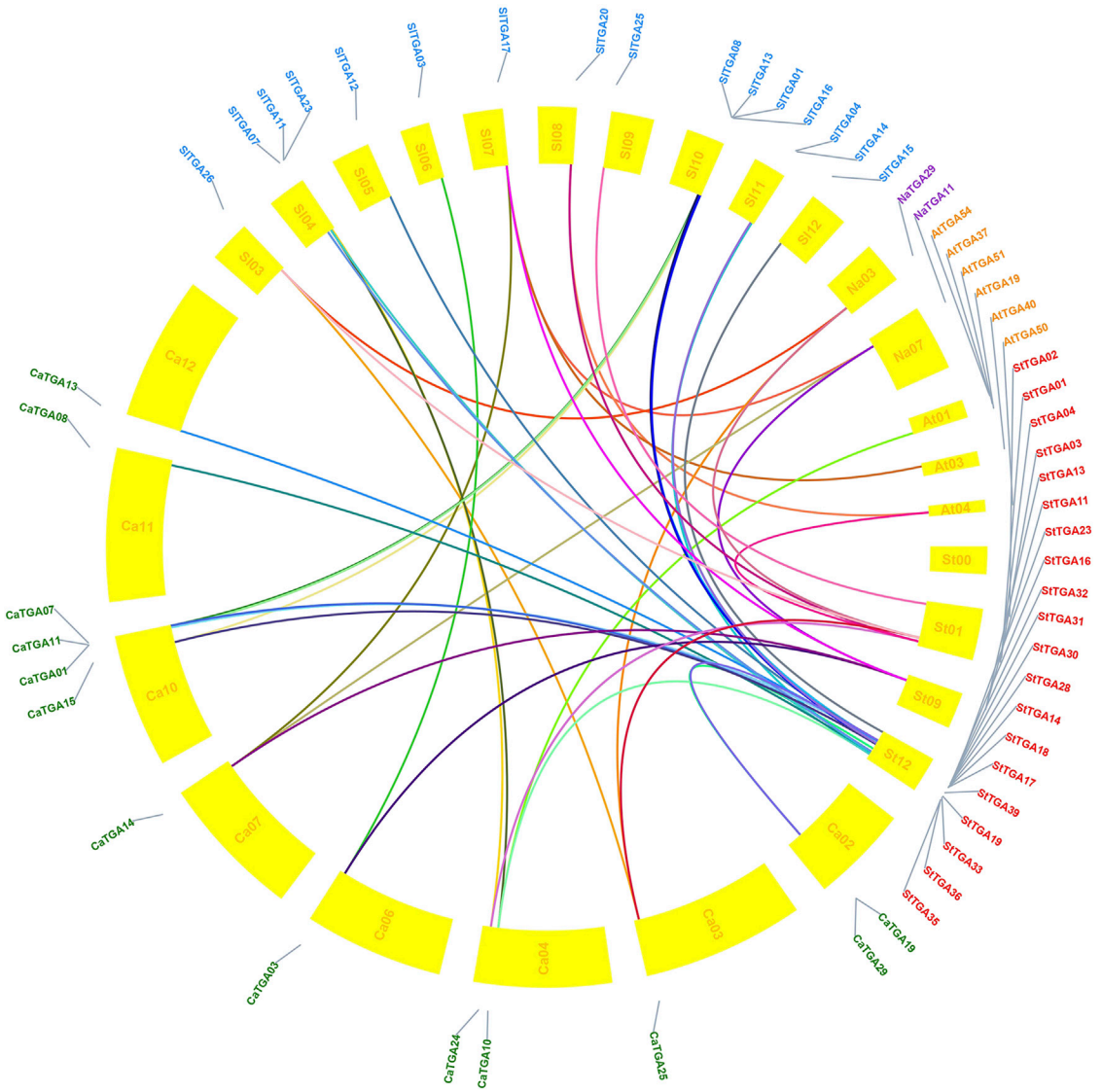



**FIGURE 3 |** Comparative orthologous relationships of *TGA* from five species. TBtools was used to analyze the gene homology relationship between potato, tomato, pepper, Arabidopsis, and tobacco *TGA* gene families and visualize the gene homology relationship. *TGA* genes connecting five species genome are shown in colored links.

# RESULTS

## Genome-Wide Identification of *StTGA* Genes

In total, 42 *StTGA* genes were identified in *S. tuberosum*, and they were named *StTGA01–StTGA42*. This study revealed that the largest number of amino acids is 488 in the *StTGA* gene family, and the relative molecular weight of these genes is between 17,290.85 and 54,371.83 Da. In addition, we predicted the pIs of StTGA protein. The results showed that the pI of StTGA17 is 4.96, which is the lowest in the StTGA protein, while the pI of StTGA01 is 9.1, which is the highest. Furthermore, we speculated the StTGA may be distributed in many parts of the cell and most of StTGA proteins are located in cytoplasm and nucleus (**Supplementary Table S1**).

## *In Silico* Chromosomal Location, Phylogenetic Relationship, Exon–Intron Structure, and Conserved Motif Analysis

Using TBtools, we drew the location map of *StTGA* and found they are distributed on four chromosomes of potato (**Figure 1A**). Chromosome St12 had 29 *StTGAs*, chromosome St08 contains six, St01 with four, and St09 contain three *StTGAs*. In order to study the StTGA family, we constructed the phylogenetic tree by MEGA 7 *via* the NJ method, which is shown in **Figure 1B**. The StTGA family was divided into six

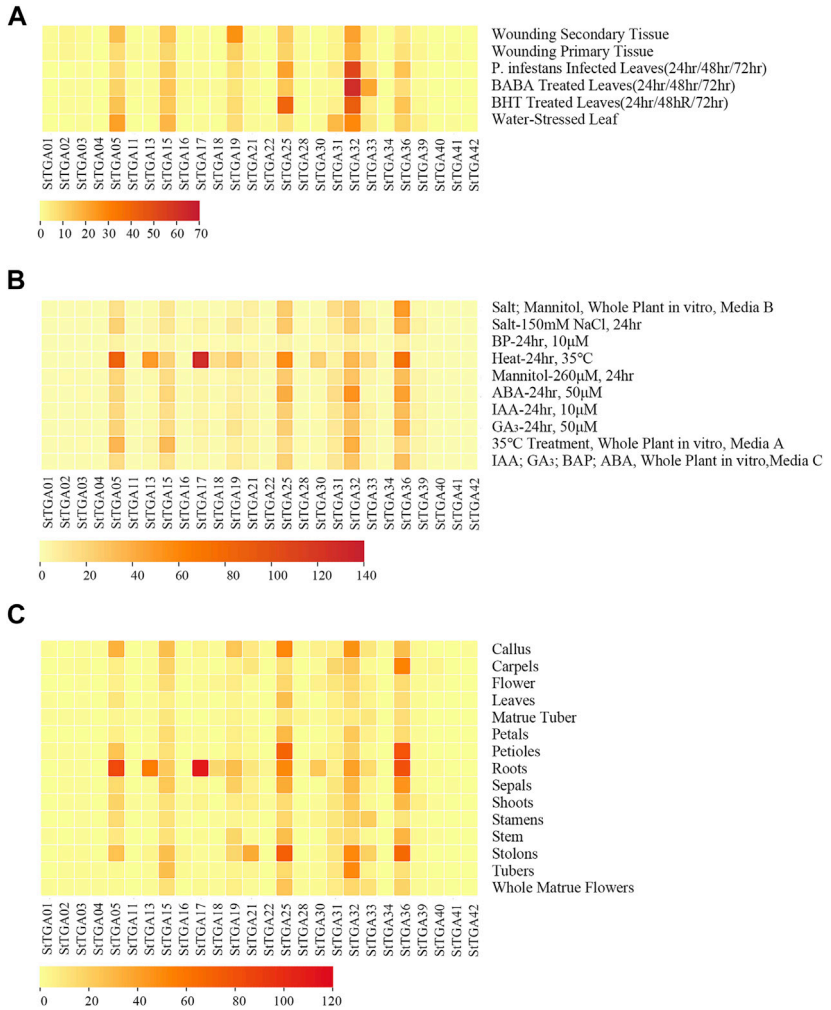



**FIGURE 4 |** Heat map of *StTGA* genes under different stresses and hormone treatments in potato. The heat map displayed the *StTGA* expression patterns of treatment in leaves, whole plant, or in different tissues and organs. **(A)** In leaves, distance values range from 0.00 to 70.00, **(B)** in the whole plant, distance values range from 0.00 to 140.00, **(C)** in different tissues or organs indicated to the right of the heat map, distance values range from 0.00 to 120.00. The gene names are indicated at the bottom. Distance value range is depicted by the gradient of colors ranging from light yellow (lowest distance value indicating high similarity between genomes) to red (highest distance value indicating low similarity between genomes). Stress treatment in leaves is indicated to the right of the heat map.

groups (designated classes I to VI) with 9, 7, 3, 12, 4, and 7 StTGAs. Then, the exon–intron structure of *StTGAs* was further analyzed. The number of exons of *StTGA* genes was between 1 and 11, of which six genes contained only one exon, and 42.86% of genes had five exons or less **Figure 1D**. Furthermore, 10 high conserved motifs were predicted in **Figure 1C** and **Figure 1E**. Compared to other classes, motifs 10 were exclusively found in classes IV, motifs 9 in classes I and II, and motifs 8 in classes IV and VI, separately **Figure 1C**.

## Cis-Acting Element Prediction Analysis of *StTGAs*

In order to study the transcriptional regulation, the 2000 bp promoter region upstream of the *StTGA* genes was extensively analyzed by PlantCARE. The cis-elements responding to abiotic/biotic stresses and hormone treatments with the prediction score being greater than or equal to 5 were considered for further analysis (**Figure 2**). The results showed that the numbers and distribution patterns of the cis-elements also greatly varied among the promoters. This means that the *StTGA* genes might have different regulatory mechanisms in expression, implying their functional divergence in responses to abiotic/biotic stress and phytohormone treatments.

## Collinear Analysis of *StTGAs*

To study the evolutionary relationship of *StTGA*, we selected *S. tuberosum* as the core and identified collinearities between tomato, tobacco, Arabidopsis, pepper, and potato *TGA* genes. There are 46 pairs of one-to-one microsynteny relations (**Figure 3**). Some of the genes could not be assigned to any

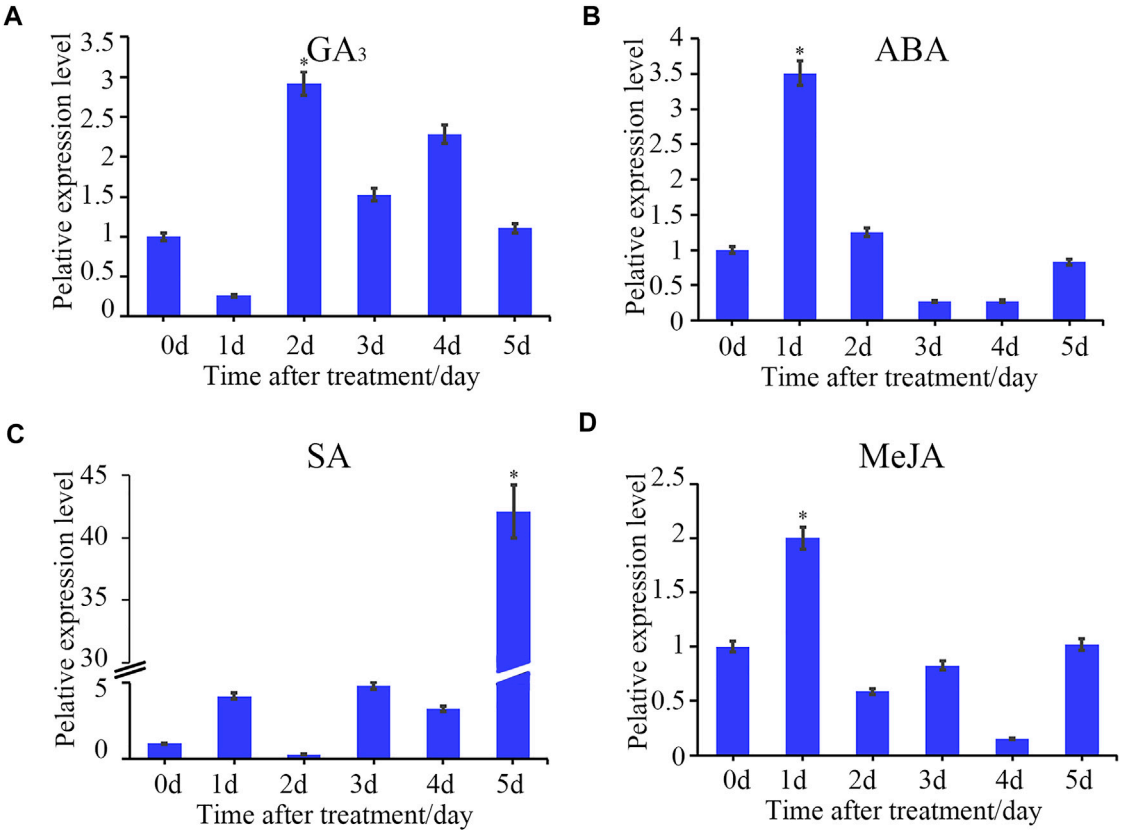



**FIGURE 5 |** Time-course gene expression pattern of *StTGA39* induced by exogenous phytohormones. The relative expression level of *StTGA39* at different time points in potato inoculated with plant hormones **(A)** $GA_3$, **(B)** ABA, **(C)** SA and **(D)** MeJA was analyzed by qRT-PCR. mRNA levels were normalized to actin (Student's *t*-test, $n = 3$ independent experiments, data shown are mean ± standard deviation).

of the chromosomes. In addition, 21 orthologous gene pairs were detected between potato and tomato, 20 orthologous genes were found between potato and pepper, 10 orthologous gene pairs were identified between potato and Arabidopsis, and two orthologous gene pairs were detected between potato and tobacco. The collinear relationship among these five species is shown in **Supplementary Additional File S3**.

## Expression Profile of *StTGAs*

We have retrieved the expression data of StTGA *via* the RNA-Seq Expression Browser and the heat map of gene FPKM value constructed using the TBtools software. We found that 11 of the StTGAs (StTGA 01, StTGA 02, StTGA03, StTGA 04, StTGA 11, StTGA 16, StTGA 23, StTGA 28, StTGA 34, StTGA 41, and StTGA42) were not expressed in almost all organs. On the contrary, StTGA05, StTGA15, StTGA25, StTGA32, and StTGA36 played roles in different developmental stages and tissues (**Figure 4**). For verification of RNA-seq data, the potato seedlings were treated with the exogenous hormones ABA, $GA_3$, SA, and MeJA, and the results showed the inducible upregulated expression of the *StTGA39* gene (**Figure 5**). By comparison, the effect of SA was relatively strong, while that of the MeJA, ABA, and $GA_3$ was light for inducing the upregulation.

## Co-Expression Network Construction

To determine the relationship of StTGA proteins and their interactors further, we performed a gene co-expression network analysis based on the STRING database. Twenty-four other potato proteins, such as 60S ribosomal protein L18 (RPLs), DNA-binding proteins (DBPs), and Glutaredoxin (GRXs), were identified **Figure 6**, which exhibited a co-expression relationship with StTGA (**Figure 6A**). The co-expression network comprised 36 pairs, and each pair of nodes was directly or indirectly connected to each component (**Figure 6A**). StTGA and StTGA members, StTGA and other potato proteins presented co-expression relationships, albeit with different weight values (**Figure 6B**). In particular, StTGA01 (StTGA28), StTGA02 (StTGA40), StTGA05 (StTGA13), etc., presented co-expression relationships (**Figure 6A**).

## Prediction of StTGA Structure

The secondary structure of StTGA39 was predicted by SOPMA. It was found that the a-helix of StTGA protein consists of 188 amino acids, accounting for 67.14%; the random coil of StTGA protein consists of 78 amino acids, accounting for 27.86%; the extended strand of StTGA protein consists of nine amino acids, accounting for 3.21%; and the β-turn of StTGA protein consists of five amino acids, accounting for 1.79% (**Figure 7A**). A

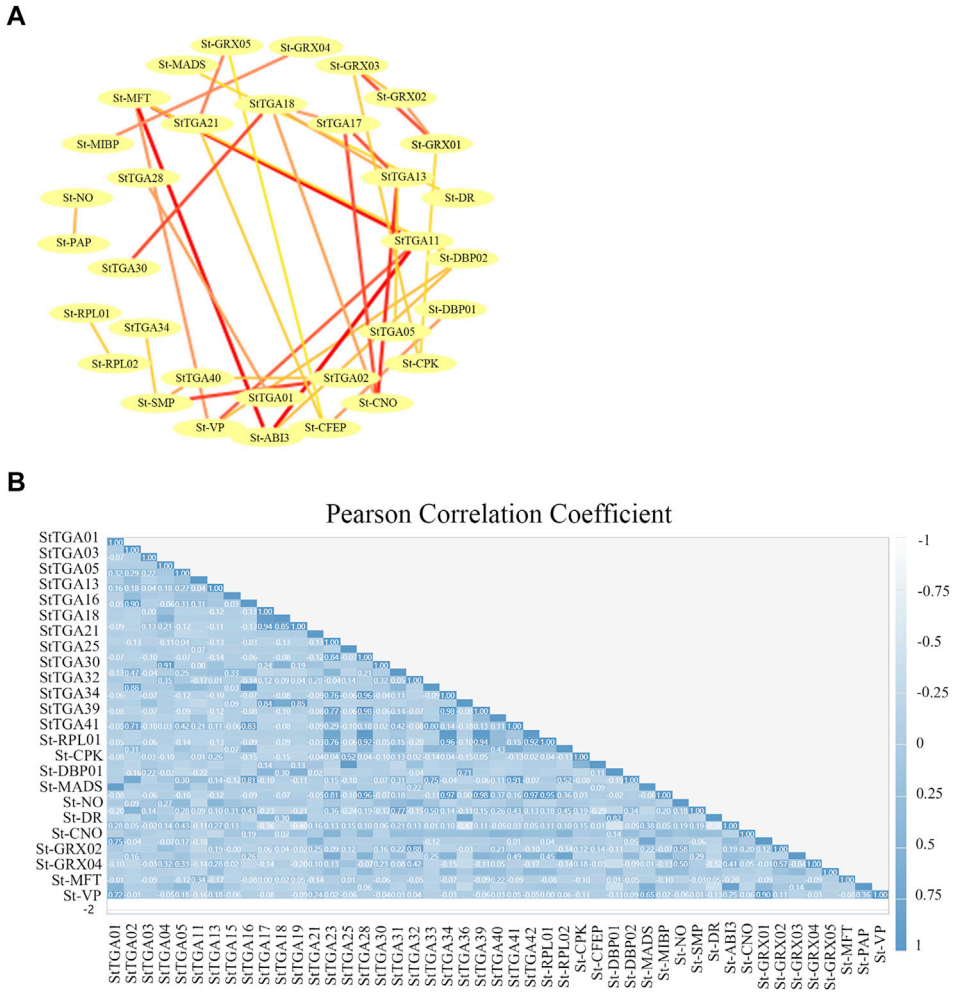



**FIGURE 6 |** Construction of the StTGA co-expression network. **(A)** Co-expression network analyses are performed to characterize the 36 differentially expressed StTGA. In the network, nodes represented RNAs, while lines represented co-expression and prediction relationship. **(B)** PCC of all genes in each gene show in this figure. Pearson correlation matrix between all StTGAs. (positive correlations = dark blue with value, negative correlations = light blue, non-significant correlations = white).

three-dimensional structure of StTGA39 was built with the I-TASSER web server. Five possible models of StTGA protein were predicted on the 3D structure, according to the website modeling with Discovery Studio 4.5 Client. The conservative domains of these models were basically the same (**Figure 7B**).

## Effect of *R. solanacearum* Infection on Potato

In order to confirm whether *StTGA39* was induced by *R. solanacearum*, the induced expression of *StTGA39* mRNA in potato seedlings inoculated with *R. solanacearum* was compared and analyzed by the qRT-PCR method (**Figure 8A**). We found that the expression level of *StTGA39* was low when plants were inoculated with water. In comparison with the control group, the expression increased at 60 hpi and reached the highest level at 72 hpi after inoculating with the highly pathogenic strain PO41. That is to say, the expression of the *StTGA39* gene was up-regulated when stressed by *R. solanacearum* and may play an important role in potato resistance to bacterial wilt. Similar results have been confirmed by other family members of *StTGA* (**Supplementary Additional File S4**).

In order to determine the defense response, various histochemical analyses were carried out on potato leaves (**Figure 8**). Under the stress of *R. solanacearum* solution of $10^8$ cfu/L, the effect of MDA content in potato seedling leaves is shown in **Figure 8B**. It could be seen from **Figure 8B** that with the prolongation of infection time of *R. solanacearum*, the change of MDA content in potato seedling leaves showed a trend, which was upregulated once at 24 h and then decreased at intervals of 24 h. Callose strengthens the plant cell walls and prevents cells from being invaded by pathogens. Aniline staining analysis and trypan blue analysis showed that the induction of callose deposition and the HR-like cell death was significantly increased in the treated potato leaves (**Figure 8C** and **Figure 8D**).

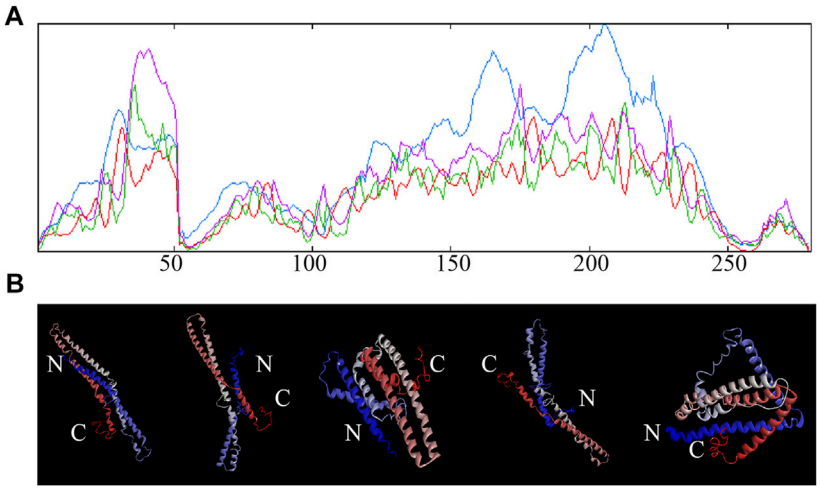



**FIGURE 7 |** Secondary and tertiary structure of potato StTGA39 protein. **(A)** Secondary structure of StTGA39 was predicted by SOPMA in potato. Blue lines represent alpha-helix, purple lines represent random coil, red lines represent extended strand, and green lines represent $\beta$-turn. **(B)** Three-dimensional structure of StTGA39 was built with the I-TASSER web server.

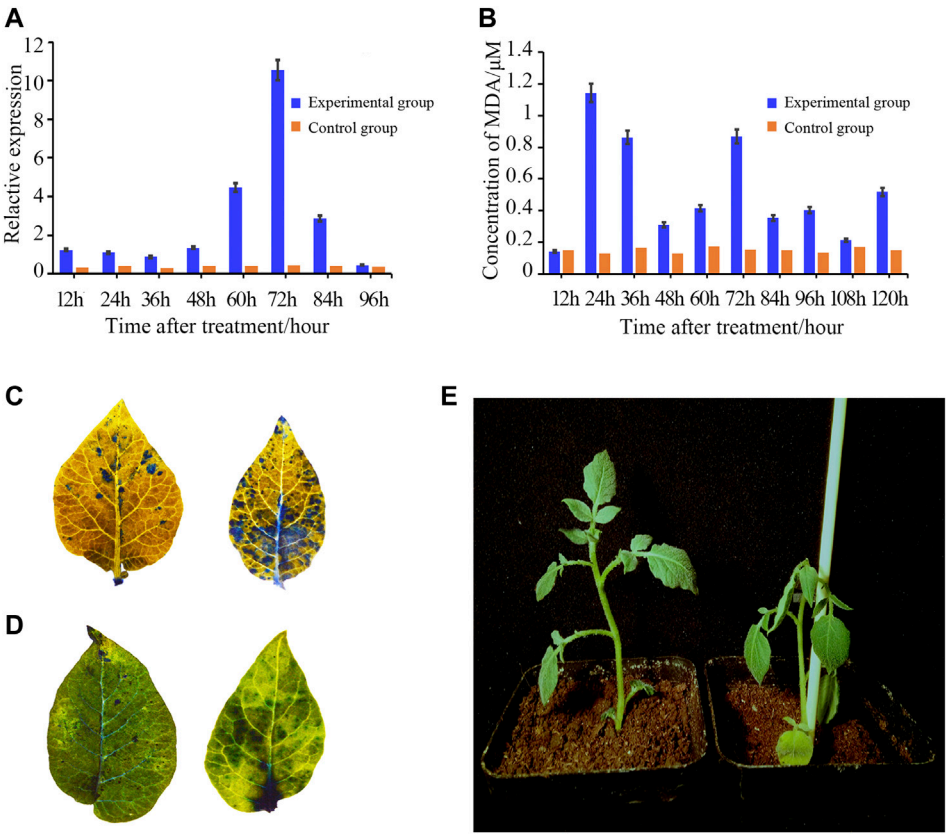



**FIGURE 8 |** Upregulated expression pattern of *StTGA39* and concentration of MDA during R. solanacearum infection on potato. **(A)** Expression pattern of *StTGA39* was induced by R. solanacearum. **(B)** Levels of MDA in potato plants were assessed at 12, 24, 36, 48, 60, 72, 84, 96, 108, and 120 hpi inoculated with R. solanacearum. **(C)** Phenotype of two-week-old potato leaves after 72 hpi with flg22 treatment was compared with WT. **(D)** Cell death in potato was detected by histochemical staining with trypan blue staining. **(E)** Photograph showed the phenotype of the WT and was taken 3 days after R. solanacearum infection. mRNA levels were normalized to actin (Student's t-test, $n$ = 3 independent experiments, data shown are mean ± standard deviation).

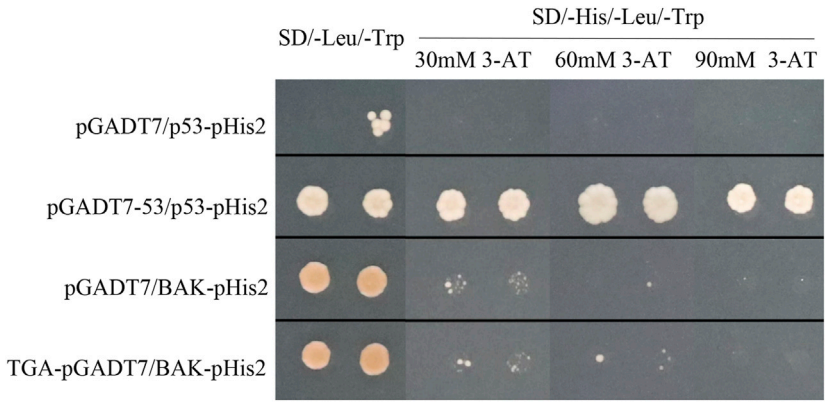



**FIGURE 9 |** Y1H showing the association of StTGA39 with the promoter of the StBAK1-20 gene. Y1H assay for determination of StTGA39-StABAK1-20 interaction. Growth of yeast cells transformed with the effector vector and the reporter vector on SD (-His, -Leu, -Trp) supplemented different concentrations of 3-AT. The positive control is pGADT7-53/p53-pHis2 vector and negative control is pGADT7/p53-pHis2 vector.

## Y1H Assay

*BAK1* is a well-established receptor for defense-related genes in plants. To identify if StTGA1 could interact directly with *StBAK1*, Y1H analysis was used (**Figure 9**). The yeast expression vector TGA-pGADT7 and BAK-pHis2 were constructed. The results of Y1H showed that all the Y1H gold strains co-transformed with pHIS2 vectors and clones grew on SD (-Leu, -Trp) medium, indicating that the co-transformation was successful. The growth was different at different concentrations of 3-AT-deficient SD (-His, -Leu, -Trp) medium. On SD (-His, -Leu, -Trp) medium, the growth inhibition of the pGADT7/BAK-pHis2 self-activated group deepened with the increase of 3-AT concentration. The growth condition of the TGA-pGADT7/BAK-pHis2 group was slightly better than that of the control group, although the experimental group could not grow on SD (-His, -Leu, -Trp; 90 mM) (**Figure 9**). The results showed that StTGA1 might bind to *StBAK1*.

## DISCUSSION

TGA TFs have been reported to function in various biological processes in plants (Jiang et al., 2021; Martins et al., 2022) and members of the TGA TFs play crucial roles in response to microbial pathogens in plants, such as Arabidopsis, soybean, rice, and tobacco (Pontier et al., 2002; Van et al., 2011; Moon et al., 2018; Ullah et al., 2019). There is a lack of reports on the potential link between TGA TFs and resistance to bacterial wilt caused by R. solanacearum in potato. In the present study, 42 StTG*A* genes were identified based upon the entire genome sequences of the potato, and the total number of TGA in potato was slightly expanded compared to that in tomato, tobacco, and pepper identified in this study but lower than that in Arabidopsis (**Supplementary Additional File S1**), which was different from the earlier reports (Ullah et al., 2019). This was mainly due to the continuous updates in the database and the different approaches used by authors. In addition, the phylogenetic analysis and collinearity was carried out. The results indicated that the StTGA family is conserved in the evolutionary history with the TGA TFs orthologs of other plants. This result corresponds well with previous observations in many plant species (Moon et al., 2018; Ullah et al., 2019; Jiang et al., 2021).

In order to investigate the expression patterns in response to different abiotic and biotic stresses, 10 members were selected for further study. The expression heat map demonstrated that the TGA TFs may have played significant and complex roles in potato (**Figure 4**), which was reflected in at least two aspects. First, the RNA-seq data-based heat map showed that the expression of most of the StTGAs was significantly different under abiotic/biotic stresses, which was consistent with previous studies (Jiang et al., 2021). Second, the same gene showed different or even opposite expression patterns under different stresses or in different tissues. For instance, the StTGA17 gene was upregulated in leaves under heat stress but downregulated under salt stress (**Figure 4B**). These results illustrate that these genes are widely involved in the response to various stresses and participate in a complex cross-regulatory network and signaling pathways (Duan et al., 2022).

To uncover the possible signal-associated functions of *StTGAs* in response to hormone stress, we conducted qRT-PCR to analyze their relative expressions under SA, MeJA, ABA, and $GA_3$ treatments (**Figure 5**). This induction may be related to the upstream cis-element elements in the promoters of the genes. According to the analysis of cis-acting elements (**Figure 2**) and the results of qRT-PCR (**Figure 5**), we speculated StTGA could be induced by multiple phytohormones and might play a crucial role in multiple hormone signal pathways. Consistent with our results, some cis-elements play an important role in responses to ABA (Baker et al., 1994; Busk and Pagés, 1998; Li J. et al., 2019). There is an interaction between NPR1 and TGA TFs, which eventually leads to activation of SA-dependent response (Rahman et al., 2012). Studies have shown that a *R. solanacearum* effector targets TGA TFs to subvert SA signaling (Qi et al., 2022). Furthermore, the JA signal pathway is related to the function of *Tripterygium wilfordii hook*. F. *TGA1* (*TwTGA1*) (Han et al., 2020). All of these research studies further support our idea that *TGA* may play a key role in the regulation of these hormone signal pathways.

An important goal in this study was to speculate the function of the *StTGA* gene in the response mechanism of potato to *R. solanacearum*. In potato, the invasion of *R. solanacearum* could cause up-regulation of *StTGA* gene expression according to **Figure 8A**. The result showed that *StTGA* was involved in resisting the invasion of *R. solanacearum*. Although the detailed mechanism needs further study, to prove this view, our study showed some other data related to resistance. Plants could resolve pathogens such as bacteria by causing ROS burst (Li et al., 2016; Ferreira et al., 2017). Evidence revealed that TGA class II plays an important role in the tolerance response to control ROS levels in Arabidopsis (Ariel et al., 2021). In addition, ROS induces gene expression and stress response and has regulatory roles in a wide range of important plant biological processes (Xia et al., 2010). Consistent with this possibility, R. solanacearum infection can result in enhanced production of ROS and lead to ROS-related oxidative damage (Ferreira et al., 2017). The results of DAB staining confirmed there is a mechanism in S. tuberosum to resist this adverse effect (**Figure 8D**). Furthermore, MDA can indicate the degree of membrane per-oxidation and further support the results of DAB staining (**Figure 8B**). The abovementioned views further support StTGA TFs are crucial regulatory factors in potato resistance to bacterial wilt, and our further data on identification of the target gene of this TF seem to support this view.

TGA TFs recognize the TGACG-motif (TGA-binding site) within the promoters of their target genes (Hou et al., 2019). To detect whether StTGA39 binds to the TGACG-motif present in the *StBAK1-20* promoter (**Supplementary Additional File S5**) in yeast, the Y1H assays were performed, and the result suggested that StTGA39 specifically binds to the TGACG-motif in the promoter of the *BAK1* gene which plays an essential role in regulated plant immunity (Beg et al., 2013). The results showed that the *StBAK1-20* gene involved in the PAMP-triggered immunity (PTI) signal pathway (Beg et al., 2013) and interacted with StTGA in our materials. The abovementioned results indicated that the StTGA can function by combining with the promoter region of *StBAK1-20* (**Supplementary Additional File S6**). To our knowledge, there are few reports on the TGA-promoted transcription of BAKs despite recent reports that TGAs directly activate respiratory burst oxidase homolog D (RBOHD) and pathogenesis-related protein 1 (PR1) expression in Arabidopsis (Qi et al., 2022; Shimizu et al., 2022).

## CONCLUSION

To sum up, we conducted a genome-wide analysis of the *StTGA* gene family in potato and primarily explored the role of them. A total of 42 *StTGA* genes were identified from the potato genome. The structure diversity, chromosomal distribution, and evolutionary history of *StTGA* were comprehensively analyzed. These results extended the understanding on the abundance and diversity of *StTGA* genes in this important crop, which may serve as a fundamental resource for the molecular breeding of potato. However, the detailed molecular mechanism needs to be further studied. Taken together, this study lays the foundation for further investigation of *StTGA* in potato.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**supplementary materials**; further inquiries can be directed to the corresponding author.

## AUTHOR CONTRIBUTIONS

TT planned and designed the research, analyzed the data, and wrote the manuscript. TT and YS studied gene expression by qRT-PCR. TT, LC, and GL studied the effect of MDA content in potato seedling leaves. GG supervised the research. RY, DY, YS, HW, XL, and GG revised the manuscript. All authors read and approved the final manuscript.

## FUNDING

This study is supported by the National Natural Science Foundation of China (31771858).

## ACKNOWLEDGMENTS

This work was funded by the National Natural Science Foundation of China (31771858). We thank all of the others in our laboratory for providing useful discussions and technical assistance and all the laboratories whose data were used in our analyses. We are very grateful to the editor and reviewers for critically evaluating the manuscript and providing constructive comments for its improvement.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.894844/full#supplementary-material

## REFERENCES

Azeem, F., Tahir, H., Ijaz, U., and Shaheen, T. (2020). A Genome-wide Comparative Analysis of bZIP Transcription Factors in *G. Arboreum* and *G. Raimondii* (Diploid Ancestors of Present-Day Cotton). *Physiol. Mol. Biol. Plants* 26 (3), 433–444. doi:10.1007/s12298-020-00771-9

Baker, S. S., Wilhelm, K. S., and Thomashow, M. F. (1994). The 5?-region of *Arabidopsis thaliana* Cor15a Has Cis-Acting Elements that Confer Cold-, Drought- and ABA-Regulated Gene Expression. *Plant Mol. Biol.* 24, 701–713. doi:10.1007/BF00029852

Barchenger, D. W., Hsu, Y.-m., Ou, J.-y., Lin, Y.-p., Lin, Y.-c., Balendres, M. A. O., et al. (2022). Whole Genome Resequencing and Complementation Tests Reveal Candidate Loci Contributing to Bacterial Wilt (Ralstonia sp.) Resistance in Tomato. *Sci. Rep.* 12 (1), 8374. doi:10.1038/s41598-022-12326-x

Budimir, J., Treffon, K., Nair, A., Thurow, C., and Gatz, C. (2020). Redox-active Cysteines in TGACG-BINDING FACTOR 1 (TGA1) Do Not Play a Role in Salicylic Acid- or Pathogen-Induced Expression of TGA1-Regulated Target Genes in *Arabidopsis thaliana*. *BioRxiv* 230 (6), 2420–2432. doi:10.1101/2020.01.30.926758

Busk, P. K., and Pagès, M. (1998). Regulation of Abscisic Acid-Induced Transcription. *Plant Mol. Biol.* 37, 425–435. doi:10.1023/A:1006058700720

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, 691384. 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, J., Mohan, R., Mohan, R., Zhang, Y., Li, M., Chen, H., et al. (2019). NPR1 Promotes its Own and Target Gene Expression in Plant Defense by Recruiting CDK8. *Plant Physiol.* 181, 289–304. doi:10.1104/pp.19.00124

Da, L., Liu, Y., Yang, J., Tian, T., She, J., Ma, X., et al. (2019). AppleMDO: A Multi-Dimensional Omics Database for Apple Co-expression Networks and Chromatin States. *Front. Plant Sci.* 10, 1333. doi:10.3389/fpls.2019.01333

Duan, L., Mo, Z., Fan, Y., Li, K., Yang, M., Li, D., et al. (2022). Genome-wide Identification and Expression Analysis of the bZIP Transcription Factor Family Genes in Response to Abiotic Stress in *Nicotiana Tabacum* L. *BMC Genomics* 23 (1), 318. doi:10.1186/s12864-022-08547-z

Fan, Y., Yan, J., Lai, D., Yang, H., Xue, G., He, A., et al. (2021a). Genome-wide Identification, Expression Analysis, and Functional Study of the GRAS Transcription Factor Family and its Response to Abiotic Stress in Sorghum [*Sorghum Bicolor* (L.) Moench]. *BMC Genomics* 22, 509. doi:10.1186/s12864-021-07848-z

Fan, Y., Yang, H., Lai, D., He, A., Xue, G., Feng, L., et al. (2021b). Genome-wide Identification and Expression Analysis of the bHLH Transcription Factor Family and its Response to Abiotic Stress in Sorghum [*Sorghum Bicolor* (L.) Moench]. *BMC Genomics* 22, 415. doi:10.1186/s12864-021-07652-9

Fang, H., Liu, Z., Long, Y., Liang, Y., Jin, Z., Zhang, L., et al. (2017). The Ca2+/calmodulin2-Binding Transcription Factor TGA3 elevatesLCDexpression and H2S Production to Bolster Cr6+tolerance in Arabidopsis. *Plant J.* 91 (6), 1038–1050. doi:10.1111/tpj.13627

Ferreira, V., Pianzzola, M. J., Vilaró, F. L., Galván, G. A., Tondo, M. L., Rodriguez, M. V., et al. (2017). Interspecific Potato Breeding Lines Display Differential Colonization Patterns and Induced Defense Responses after Ralstonia Solanacearum Infection. *Front. Plant Sci.* 8, 1424. doi:10.3389/fpls.2017.01424

Gao, G., Jin, L. P., Xie, K. Y., and Qu, D. Y. (2009). The Potato StLTPa7 Gene Displays a Complex Ca 2+ -associated Pattern of Expression during the Early Stage of Potato- Ralstonia Solanacearum Interaction. *Mol. Plant Pathol.* 10, 15–27. doi:10.1111/j.1364-3703.2008.00508.x

Gao, H., Wang, Z., Li, S., Hou, M., Zhou, Y., Zhao, Y., et al. (2018). Genome-wide Survey of Potato *MADS-Box* Genes Reveals that *StMADS1* and *StMADS13* Are Putative Downstream Targets of Tuberigen *StSP6A*. *BMC Genomics* 19 (1), 726. doi:10.1186/s12864-018-5113-z

Gebhardt, C. (2016). The Historical Role of Species from the Solanaceae Plant Family in Genetic Research. *Theor. Appl. Genet.* 129, 2281–2294. doi:10.1007/s00122-016-2804-1

Geourjon, C., and Deléage, G. (1995). SOPMA: Significant Improvements in Protein Secondary Structure Prediction by Consensus Prediction from Multiple Alignments. *Bioinformatics* 11 (6), 681–684. doi:10.1093/bioinformatics/11.6.681

Guo, S., Zuo, Y., Zhang, Y., Wu, C., Su, W., Jin, W., et al. (2017). Large-scale Transcriptome Comparison of Sunflower Genes Responsive to *Verticillium dahliae*. *BMC Genomics* 18 (1), 42. doi:10.1186/s12864-016-3386-7

Gutsche, N., Holtmannspötter, M., Maß, L., O'Donoghue, M., Busch, A., Lauri, A., et al. (2017). Conserved Redox-dependent DNA Binding of ROXY Glutaredoxins with TGA Transcription Factors. *Plant Direct* 1 (6), e00030. doi:10.1002/pld3.30

Gutsche, N., and Zachgo, S. (2016). The N-Terminus of the Floral Arabidopsis TGA Transcription Factor PERIANTHIA Mediates Redox-Sensitive DNA-Binding. *Plos One* 11 (4), e0153810. doi:10.1371/journal.pone.0153810

Han, J., Liu, H.-t., Wang, S.-c., Wang, C.-r., and Miao, G.-p. (2020). A Class I TGA Transcription Factor from Tripterygium Wilfordii Hook.F. Modulates the Biosynthesis of Secondary Metabolites in Both Native and Heterologous Hosts. *Plant Sci.* 290, 110293. doi:10.1016/j.plantsci.2019.110293

He, L. Y., Sequeira, L., and Kelman, A. (1983). Characteristics of Strains of Pseudomonas Solanacearumfrom China. *Plant Dis.* 67 (12), 1357–1361. doi:10.1094/PD-67-1357

He, Z. Q., Li, Y. H., Feng, G. H., Yuan, X. W., Lu, Z. B., Dai, M., et al. (2022). Pharmacological Perturbation of Mechanical Contractility Enables Robust Transdifferentiation of Human Fibroblasts into Neurons. *Adv. Sci.* 9, 2104682. doi:10.1002/advs.202104682

Herrera-Vásquez, A., Fonseca, A., Ugalde, J. M., Lamig, L., Seguel, A., Moyano, T. C., et al. (2021). TGA Class II Transcription Factors Are Essential to Restrict Oxidative Stress in Response to UV-B Stress in Arabidopsis. *J. Exp. Bot.* 72 (5), 1891–1905. doi:10.1093/jxb/eraa534

Hou, J., Sun, Q., Li, J., Ahammed, G. J., Yu, J., Fang, H., et al. (2019). Glutaredoxin S25 and its Interacting TGACG Motif-Binding Factor TGA2 Mediate Brassinosteroid-Induced Chlorothalonil Metabolism in Tomato Plants. *Environ. Pollut.* 255 (2), 113256. doi:10.1016/j.envpol.2019.113256

Huang, L.-J., Li, N., Thurow, C., Wirtz, M., Hell, R., and Gatz, C. (2016). Ectopically Expressed Glutaredoxin ROXY19 Negatively Regulates the Detoxification Pathway in *Arabidopsis thaliana*. *BMC Plant Biol.* 16, 200. doi:10.1186/s12870-016-0886-1

Huang, X.-S., Liu, J.-H., and Chen, X.-J. (2010). Overexpression of PtrABF Gene, a bZIP Transcription Factor Isolated from *Poncirus Trifoliata*, Enhances Dehydration and Drought Tolerance in Tobacco via Scavenging ROS and Modulating Expression of Stress-Responsive Genes. *BMC Plant Biol.* 10 (1), 230. doi:10.1186/1471-2229-10-230

Islamoglu, H., Cao, R., Teskey, G., Gyurjian, K., Lucar, S., Fraix, M., et al. (2018). Effects of readiSorb L-GSH in Altering Granulomatous Responses against mycobacterium Tuberculosis Infection. *Jcm* 7, 40. doi:10.3390/jcm7030040

Jiang, H., Gu, S., Li, K., and Gai, J. (2021). Two TGA Transcription Factor Members from Hyper-Susceptible Soybean Exhibiting Significant Basal Resistance to Soybean Mosaic Virus. *Ijms* 22 (21), 11329. doi:10.3390/ijms222111329

Ke, D., He, Y., Fan, L., Niu, R., Cheng, L., Wang, L., et al. (2022). The Soybean TGA Transcription Factor GmTGA13 Plays Important Roles in the Response to Salinity Stress. *Plant Biol. J.* 24 (2), 313–322. doi:10.1111/plb.13360

Kim, B. H., Kim, S. Y., and Nam, K. H. (2013). Assessing the Diverse Functions of BAK1 and its Homologs in Arabidopsis, beyond BR Signaling and PTI Responses. *Mol. Cells* 35 (1), 7–16. doi:10.1007/s10059-013-2255-3

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, C.-Y., Luo, Y.-p., Duan, T.-T., Xue, Z., Gao, X.-D., Zhao, X., et al. (2016). Potato Remorin Gene *StREMa4* Cloning and its Spatiotemporal Expression Pattern under *Ralstonia Solanacearum* and Plant Hormones Treatment. *Phytoparasitica* 44 (4), 575–584. doi:10.1007/s12600-016-0536-z

Kumar, P., Kumar, P., Sharma, D., Verma, S. K., Halterman, D., and Kumar, A. (2021). Genome-wide Identification and Expression Profiling of Basic Leucine Zipper Transcription Factors Following Abiotic Stresses in Potato (*Solanum tuberosum* L.). *Plos One* 16 (3), e0247864. doi:10.1371/journal.pone.0247864

Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Peer, Y. V. D., et al. (2002). PlantCARE, a Database of Plant Cis-Acting Regulatory Elements and a Portal to Tools for In Silico Analysis of Promoter Sequences. *Nucleic Acids Res.* 30 (1), 325–327. doi:10.1093/nar/30.1.325

Li, B., Liu, Y., Cui, X.-Y., Fu, J.-D., Zhou, Y.-B., Zheng, W.-J., et al. (2019). Genome-wide Characterization and Expression Analysis of Soybean TGA Transcription Factors Identified a Novel *TGA* Gene Involved in Drought and Salt Tolerance. *Front. Plant Sci.* 10, 549. doi:10.3389/fpls.2019.00549

Li, J., Han, G., Sun, C., and Sui, N. (2019). Research Advances of MYB Transcription Factors in Plant Stress Resistance and Breeding. *Plant Signal. Behav.* 14 (8), 1613131. doi:10.1080/15592324.2019.1613131

Li, K., Du, Y., and Miao, Y. (2016). Future Challenges in Understanding ROS in Plant Responses to Abiotic Stress. *Sci. China Life Sci.* 59 (12), 1343–1344. doi:10.1007/s11427-016-0362-7

Liang, Y., Wan, N., Cheng, Z., Mo, Y., Liu, B., Liu, H., et al. (2017). Whole-genome Identification and Expression Pattern of the Vicinal Oxygen Chelate Family in Rapeseed (*Brassica Napus* L.). *Front. Plant Sci.* 8, 745. doi:10.3389/fpls.2017.00745

Liu, C., Long, J., Zhu, K., Liu, L., Yang, W., Zhang, H., et al. (2016). Characterization of a Citrus R2R3-MYB Transcription Factor that Regulates the Flavonol and Hydroxycinnamic Acid Biosynthesis. *Sci. Rep.* 6, 25352. doi:10.1038/srep25352

Martins, D. C., Rubiales, D., and VazPatto, M. C. (2022). Association Mapping of *Lathyrus Sativus* Disease Response to Uromyces Pisi Reveals Novel Loci Underlying Partial Resistance. *Front. Plant Sci.* 13 (13), 842545. doi:10.3389/fpls.2022.842545

Moon, S.-J., Park, H. J., Kim, T.-H., Kang, J.-W., Lee, J.-Y., Cho, J.-H., et al. (2018). *OsTGA2* Confers Disease Resistance to Rice against Leaf Blight by Regulating Expression Levels of Disease Related Genes via Interaction with NH1. *Plos One* 13 (11), e0206910. doi:10.1371/journal.pone.0206910

Pontier, D., Privat, I., Trifa, Y., Zhou, J.-M., Klessig, D. F., and Lam, E. (2002). Differential Regulation of TGA Transcription Factors by Post-transcriptional Control. *Plant J.* 32, 641–653. doi:10.1046/j.1365-313X.2002.01461.x

Qi, P., Huang, M., Hu, X., Zhang, Y., Wang, Y., Li, P., et al. (2022). A *Ralstonia Solanacearum* Effector Targets TGA Transcription Factors to Subvert Salicylic Acid Signaling. *Plant Cell* 34, 1666–1683. doi:10.1093/plcell/koac015

Rahman, T. A. E., Oirdi, M. E., Gonzalez-Lamothe, R., and Bouarab, K. (2012). Necrotrophic Pathogens Use the Salicylic Acid Signaling Pathway to Promote Disease Development in Tomato. *Mpmi* 25 (12), 1584–1593. doi:10.1094/MPMI-07-12-0187-R

Rouhier, N., Cerveau, D., Couturier, J., Reichheld, J.-P., and Rey, P. (2015). Involvement of Thiol-Based Mechanisms in Plant Development. *Biochimica Biophysica Acta (BBA) - General Subj.* 1850 (8), 1479–1496. doi:10.1016/j.bbagen.2015.01.023

Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a Unified Platform for Automated Protein Structure and Function Prediction. *Nat. Protoc.* 5 (4), 725–738. doi:10.1038/nprot.2010.5

Salanoubat, M., Genin, S., Artiguenave, F., Gouzy, J., Mangenot, S., Arlat, M., et al. (2002). Genome Sequence of the Plant Pathogen *Ralstonia Solanacearum*. *Nature* 415, 497–502. doi:10.1038/415497a

Salladini, E., Jørgensen, M. L. M., Theisen, F. F., and Skriver, K. (2020). Intrinsic Disorder in Plant Transcription Factor Systems: Functional Implications. *Ijms* 21 (24), 9755. doi:10.3390/ijms21249755

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

Shimizu, K., Suzuki, H., Uemura, T., Nozawa, A., Desaki, Y., Hoshino, R., et al. (2022). Immune Gene Activation by NPR and TGA Transcriptional Regulators in the Model Monocot *Brachypodium Distachyon*. *Plant J.* 110 (2), 470–481. doi:10.1111/tpj.15681

Song, Y., Zhai, Y., Li, L., Yang, Z., Ge, X., Yang, Z., et al. (2021). BIN2 Negatively Regulates Plant Defence against *Verticillium dahliae* in Arabidopsis and Cotton. *Plant Biotechnol. J.* 19, 2097–2112. doi:10.1111/pbi.13640

Thordal-Christensen, H., Zhang, Z., Wei, Y., and Collinge, D. B. (1997). Subcellular Localization of H2O2 in Plants. H2O2 Accumulation in Papillae and Hypersensitive Response during the Barley-Powdery Mildew Interaction. *Plant J.* 11 (6), 1187–1194. doi:10.1046/j.1365-313X.1997.11061187.x

Uhrig, J. F., Huang, L.-J., Barghahn, S., Willmer, M., Thurow, C., and Gatz, C. (2017). CC-type Glutaredoxins Recruit the Transcriptional Co-repressor TOPLESS to TGA-dependent Target Promoters in *Arabidopsis thaliana*. *Biochimica Biophysica Acta (BBA) - Gene Regul. Mech.* 1860 (2), 218–226. doi:10.1016/j.bbagrm.2016.11.001

Ullah, I., Magdy, M., Wang, L., Liu, M., and Li, X. (2019). Genome-wide Identification and Evolutionary Analysis of TGA Transcription Factors in Soybean. *Sci. Rep.* 9 (1), 11186. doi:10.1038/s41598-019-47316-z

van Verk, M. C., Neeleman, L., Bol, J. F., and Linthorst, H. J. M. (2011). Tobacco Transcription Factor NtWRKY12 Interacts with TGA2.2 *In Vitro* and *In Vivo*. *Front. Plant Sci.* 2 (32), 1–10. doi:10.3389/fpls.2011.00032

Wang, Y., Tang, H., Debarry, J. D., Tan, X., Li, J., Wang, X., et al. (2012). MCScanX: a Toolkit for Detection and Evolutionary Analysis of Gene Synteny and Collinearity. *Nucleic Acids Res.* 40, e49. doi:10.1093/nar/gkr1293

Wang, Y., Zhang, N., Li, T., Yang, J., Zhu, X., Fang, C., et al. (2019). Genome-wide Identification and Expression Analysis of StTCP Transcription Factors of Potato (*Solanum tuberosum* L.). *Comput. Biol. Chem.* 78, 53–63. doi:10.1016/j.compbiolchem.2018.11.009

Xia, X.-J., Chen, Z., and Yu, J.-Q. (2010). ROS Mediate Brassinosteroids-Induced Plant Stress Responses. *Plant Signal. Behav.* 5 (5), 532–534. doi:10.4161/psb.10989

Yu, R. M., Chang, Y. N., Song, Y. J., Tian, T., Wang, H. J., and Gao, G. (2021a). Genome-Wide Identification and Expression Analysis of the Zinc Transporter Protein ZIP Family in Potato (Solanum tuberosum). *Intl J. Agric. Biol.* 24, 1533–1542. doi:10.17957/IJAB/15.1592

Yu, R.-M., Suo, Y.-Y., Yang, R., Chang, Y.-N., Tian, T., Song, Y.-J., et al. (2021b). StMBF1c Positively Regulates Disease Resistance to Ralstonia Solanacearum via It's Primary and Secondary Upregulation Combining Expression of StTPS5 and Resistance Marker Genes in Potato. *Plant Sci.* 307, 110877. doi:10.1016/j.plantsci.2021.110877

Zhu, C., Zhang, S., Zhou, C., Chen, L., Fu, H., Li, X., et al. (2020). Genome-wide Investigation and Transcriptional Analysis of Cytosine-5 DNA Methyltransferase and DNA Demethylase Gene Families in Tea Plant (*Camellia Sinensis*) under Abiotic Stress and Withering Processing. *Peer J.* 8, e8432. doi:10.7717/peerj.8432

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

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

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 02 August 2022
DOI 10.3389/fgene.2022.932166



OPEN ACCESS

EDITED BY
Sundeep Kumar,
National Bureau of Plant Genetic Resources (ICAR), India

REVIEWED BY
Pallavi Sinha,
IRRI South Asia Hub, India
Abdul Fiyaz R.,
Indian Institute of Rice Research (ICAR), India
Galal Bakr Anis,
China National Rice Research Institute, Chinese Academy of Agricultural Sciences, China

*CORRESPONDENCE
AK Singh,
aks_gene@yahoo.com

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

RECEIVED 29 April 2022
ACCEPTED 28 June 2022
PUBLISHED 02 August 2022

CITATION
Malik A, Kumar A, Ellur RK, Krishnan S G, Dixit D, Bollinedi H, Vinod KK, Nagarajan M, Bhowmick PK, Singh NK and Singh AK (2022), Molecular mapping of QTLs for grain dimension traits in Basmati rice.
*Front. Genet.* 13:932166.
doi: 10.3389/fgene.2022.932166



# Molecular mapping of QTLs for grain dimension traits in Basmati rice

Ankit Malik[1,2], Aruna Kumar[2], Ranjith Kumar Ellur[1], Gopala Krishnan S[1], Deepshikha Dixit[1], Haritha Bollinedi[1], KK Vinod[1], M Nagarajan[3], PK Bhowmick[1], NK Singh[4] and AK Singh[1]*

[1]Division of Genetics, ICAR-Indian Agricultural Research Institute (ICAR-IARI), New Delhi, India, [2]Amity Institute of Biotechnology, Amity University, Noida, India, [3]Rice Breeding and Genetics Research Centre, ICAR-IARI, Aduthurai, India, [4]ICAR-National Institute for Plant Biotechnology, IARI, New Delhi, India

Basmati rice is known for its extra-long slender grains, exceptional kernel dimensions after cooking, high volume expansion, and strong aroma. Developing high yielding Basmati rice varieties with good cooking quality is a gigantic task. Therefore, identifying the genomic regions governing the grain and cooked kernel dimension traits is of utmost importance for its use in marker-assisted breeding. Although several QTLs governing grain dimension traits have been reported, limited attempts have been made to map QTLs for grain and cooked kernel dimension traits of Basmati rice. In the current study, a population of recombinant inbred lines (RIL) was generated from a cross of Sonasal and Pusa Basmati 1121 (PB1121). In the RIL population, there was a significant positive correlation among the length (RRL: rough rice length, MRL: milled rice length, CKL: cooked kernel length) and breadth (RRB: rough rice breadth, MRB: milled rice breadth and CKB: cooked kernel breadth) of the related traits, while there was significant negative correlation between them. QTL mapping has led to the identification of four major genomic regions governing MRL and CKL. Two QTLs co-localize with the earlier reported major gene *GS3* and a QTL *qGRL7.1,* while the remaining two QTLs *viz., qCKL3.2* (*qMRL3.2*) and *qCKL4.1* (*qMRL4.1*) were novel. The QTL *qCKL3.2* has been bracketed to a genomic region of 0.78 Mb between the markers RM15247 and RM15281. Annotation of this region identified 18 gene models, of which the genes predicted to encode pentatricopeptides and brassinosteroid insensitive 1-associated receptor kinase 1 precursor may be the putative candidate genes. Furthermore, we identified a novel QTL *qKER2.1* governing kernel elongation ratio (KER) in Basmati rice.

KEYWORDS

**quantitative trait loci, basmati rice, grain quality, SSR markers, mapping**

## Introduction

Basmati rice attracts consumers worldwide due to its pleasant aroma and unique grain and cooking quality parameters. Among the Basmati rice varieties developed, PB1121 stands out due to its exclusive grain qualities, making it one of the most traded crop varieties in the world. It possesses extra-long slender grains, very long cooked kernel (upto 22 mm), high kernel elongation ratio, volume expansion of more than four times upon cooking, and pleasant aroma. The grain dimension traits that determine rice grain quality are the length and breadth of rough, milled, and cooked rice grain as well as the kernel elongation ratio (KER). The genetics of rice grain size and weight has been extensively studied and >300 QTLs have been mapped (www.gramene.org) (Bhatia et al., 2018). However, only few of them *viz., GS3* (Fan et al., 2006), *GS5* (Li et al., 2011), *GS7* (Shao et al., 2012), *qGL3* (Zhang et al., 2012b), *GW2* (Song et al., 2007), *GW8* (Wang et al., 2012), *TGW6* (Ishimaru et al., 2013), *GW5* (Duan et al., 2017), *qTGW3* (Zejun et al., 2018), and *qGS10* (Zhu et al., 2019) have been cloned and functionally validated. Among these, *GS3* has been identified as a major gene governing the rice grain length and weight (Mao et al., 2010). Several polymorphic loci of *GS3 viz,* $SR_{17}$ (2nd intron), $RGS_1$ (4th intron), $RGS_2$ (5th exon), *aksGS3-12* (5th exon) etc., have been reported (Wang et al., 2011; Anand et al., 2013a). The gene *qGL3* has been identified to encode a putative serine/threonine protein phosphatase that contains a Kelch-like repeat domain and acts on its cell cycle-related substrate, Cyclin-T1; 3, to increase the rice grain length (Zhang et al., 2012b).

Several independent studies have identified QTLs for the grain width and weight. In a large-grained genotype WY3, a major gene *GW2* encoding a RING-type E3 ubiquitin ligase was identified wherein, a single base pair deletion results into premature stop codon, thereby reducing its expression while increasing the grain width and weight (Song et al., 2007). *GW5* was mapped on chromosome 5 which encodes a putative serine carboxypeptidase, wherein a deletion increases the cell number in the outer glume of the rice flower thereby regulating the grain width, filling, and weight (Li et al., 2011). *TGW6* was mapped on chromosome 6 which encodes a novel protein with an indole-3-acetic acid (IAA)-glucose hydrolase activity and its loss of function enhances the rice grain weight (Ishimaru et al., 2013).

Identifying genomic regions governing grain dimension traits in Basmati rice genotypes helps in their precise manifestation in Basmati improvement programs. Earlier, in a RIL population generated from a cross Pusa 1342/PB1121, several novel QTLs governing grain length, elongation ratio, and aroma in Basmati rice were identified (Amarawathi et al., 2008). Subsequently, a major gene *GW8* promoting cell division and proliferation thereby enhancing the grain width and yield was identified in a traditional Basmati rice variety, Basmati 385 (Wang et al., 2012). Furthermore, three QTLs for grain elongation on chromosomes 4 (*qGE4.1*) and 6 (*qGE6.1* and *qGE6.2*) were identified in a Basmati rice genotype (Arikit et al., 2019).

Although, Basmati rice possesses unique grain dimension traits such as extra-long slender grains and exceptional cooked kernel dimensions, there are limited efforts made toward identifying the genomic regions governing these traits. In the present study, we attempted to map QTLs governing the grain and cooked kernel dimension traits using a RIL population generated from the cross of Sonasal and PB1121.

## Materials and methods

### Development of the mapping population

A RIL population comprising of 173 lines was developed by crossing an aromatic short grain genotype Sonasal with an extra-long grained Basmati rice variety PB1121. Sonasal has extremely shorter grains, which are at least 50% shorter than that of PB1121, with slightly wider grains. Both parents were aromatic. The population was advanced till the $F_{13}$ generation through the single seed descent method (Figure 1).

### Evaluation of grain dimension traits

The RIL population along with the parental lines and checks were grown in an augmented randomized block design following all the standard agronomic practices. The seed harvested from each of the lines was shade-dried to a uniform moisture content of 14% and stored at room temperature. Observations on rough, milled, and cooked grain dimension traits were recorded after six months. The grain dimension traits namely, RRL, RRB, MRL, MRB, CKL, and CKB were recorded in three replications with ten grains in each replication. For CKL and CKB, ten milled rice grains were soaked in 20 ml of distilled water for 20 min in a test tube. The samples were placed in a water bath at 100°C for 7 min followed by cooling them to room temperature. A photo analyzer was used (C-DAC) to record the data on the length and breadth of the grains.

### Genotyping and data analysis

DNA was extracted using a standard CTAB (Cetyl- Tri Methyl Ammonium Bromide) method (Murray and Thompson, 1980). The PCR reaction was carried out using 25 ng of genomic DNA, 1X PCR assay buffer with 1.5 mM MgCl2, 5 pmol of forward and reverse primers, 0.05 mM dNTPs, and 1U of Taq polymerase. The final reaction volume was made up to 10 µl using nuclease-free water. Amplified PCR products were resolved by 3.5% agarose gel electrophoresis and documented on a gel documentation system (BIO-RAD Gel Doc XR+, California, United States). A total of 1,054 markers comprising of sequence-tagged microsatellite site (STMS) and gene-based markers were used in a parental polymorphism survey. The polymorphic markers were used to

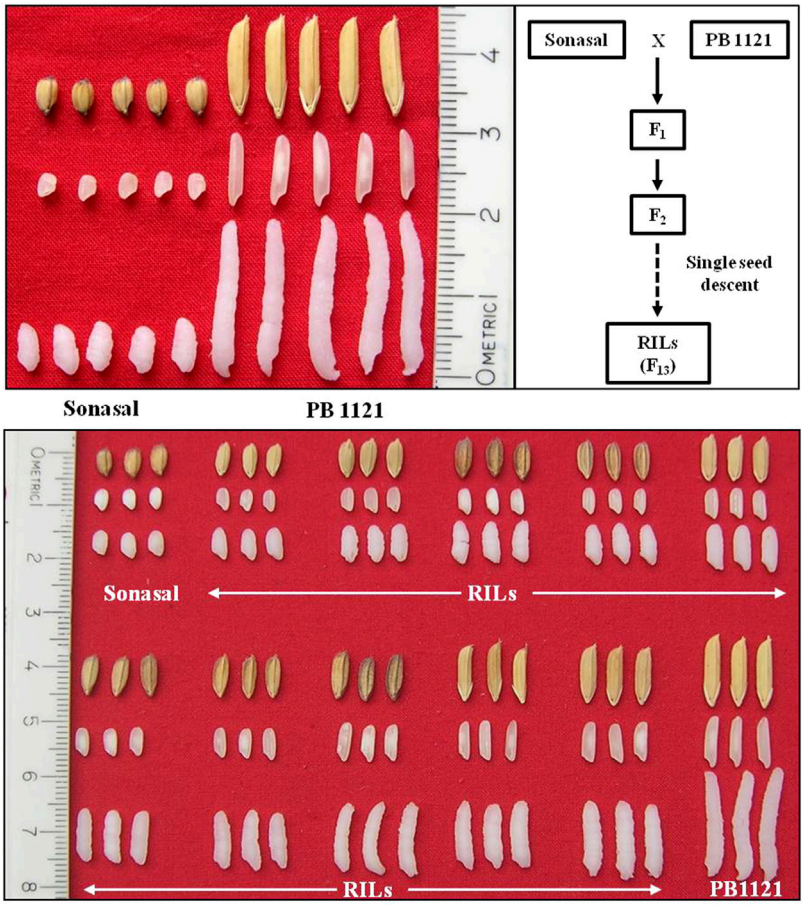



FIGURE 1
Variation for grain dimension traits between parental lines, Sonasal, and PB1121; and RIL population.

genotype the RIL population. The linkage map was constructed using the QTL IciMapping software (Li et al., 2007; Meng et al., 2015). Grouping and ordering of polymorphic markers were carried out using the regression mapping algorithm using recombination frequency (REC) with a default value set to 0.3. Furthermore, rippling was carried out for fine-tuning the ordered markers on their respective chromosomes using the REC algorithm with a default window size of 5. QTLs were identified using the inclusive composite interval mapping (ICIM) algorithm for additive gene effects.

# Results

## Phenotyping of the RIL population

The RIL population comprising of 173 lines were evaluated during three subsequent crop growing seasons, *viz.*, *Kharif* 2019, 2020, and 2021 (Table 1). The population showed significant variations for all the grain dimension-related traits. The RRL ranged from 4.67 to 12.33 mm while, RRB ranged from 1.40 to 3.00 mm. The corresponding measures for RRL in Pusa Basmati 1121 were 12.43 and 6.37 mm in Sonasal. Similarly, for RRB, the parents showed average values of 2.30 and 2.60 mm for Pusa Basmati 1121 and Sonasal, respectively. The MRL among the RILs varied from 3.37 to 8.07 mm, while Pusa Basmati 1121 measured 8.07 mm and Sonasal 4.07 mm. For MRB, RILs ranged between 1.33 and 2.47 mm, while the parents recorded values of 1.80 and 2.20 mm, respectively for Pusa Basmati 1121 and Sonasal. There was a huge variation for CKL which ranged from 6.27 to 17.13 mm among the RILs, while CKB ranged between 1.78 and 3.10 mm. CKL and CKB for Pusa Basmati 1121 were 16.60 and 2.33 mm, respectively, while that of Sonasal was 6.25 and 2.47 mm, in that order. Among the RILs, KER varied from 1.23 to 2.59, compared to 2.06 and 1.54 recorded in Pusa Basmati 1121 and Sonasal, respectively. The phenotypic coefficient of variation (PCV) among the traits ranged from a minimum of 7.12% for CKB to a maximum of

TABLE 1 Descriptive statistics of Sonasal, PB1121, and RIL population.

| Traits | Sonasal | PB1121 | Variability in RILs *Kharif* 2019 | | | Variability in RILs *Kharif* 2020 | | | Variability in RILs *Kharif* 2021 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Mean ± SE | Range (mm) | CV | Mean ± SE | Range (mm) | CV | Mean ± SE | Range (mm) | CV |
| RRL | 6.37 | 12.43 | 8.35 ± 0.12 | 5.67–12.33 | 19.29 | 8.35 ± 0.11 | 5.80–12.07 | 17.17 | 7.64 ± 0.12 | 4.67–11.37 | 19.87 |
| RRB | 2.60 | 2.30 | 2.40 ± 0.02 | 1.87–3.00 | 8.49 | 2.39 ± 0.02 | 1.73–2.93 | 10.69 | 2.02 ± 0.01 | 1.40–2.67 | 9.13 |
| RRLB | 2.45 | 5.40 | 3.52 ± 0.06 | 2.28–6.54 | 23.44 | 3.56 ± 0.07 | 2.28–6.85 | 24.83 | 3.82 ± 0.07 | 2.03–6.24 | 23.67 |
| MRL | 4.07 | 8.07 | 5.31 ± 0.08 | 3.47–7.73 | 20.94 | 5.33 ± 0.08 | 3.53–7.73 | 19.68 | 5.42 ± 0.09 | 3.37–8.07 | 21.57 |
| MRB | 2.20 | 1.80 | 1.88 ± 0.01 | 1.40–2.47 | 9.19 | 1.90 ± 0.01 | 1.47–2.33 | 7.43 | 1.66 ± 0.01 | 1.33–2.00 | 5.14 |
| MRLB | 1.85 | 4.48 | 2.86 ± 0.06 | 1.80–4.76 | 25.43 | 2.83 ± 0.05 | 1.76–4.78 | 22.62 | 3.27 ± 0.06 | 1.94–5.63 | 22.42 |
| CKL | 6.25 | 16.60 | 10.23 ± 0.18 | 6.27–17.13 | 23.41 | 10.18 ± 0.16 | 6.80–16.20 | 20.27 | 10.15 ± 0.16 | 6.83–15.90 | 20.87 |
| CKB | 2.47 | 2.33 | 2.47 ± 0.01 | 2.07–2.87 | 7.12 | 2.48 ± 0.01 | 2.00–2.93 | 6.40 | 2.17 ± 0.02 | 1.78–3.10 | 9.35 |
| CKLB | 2.53 | 7.11 | 4.17 ± 0.08 | 2.64–7.56 | 25.86 | 4.13 ± 0.07 | 2.52–7.55 | 22.44 | 4.72 ± 0.08 | 2.64–7.81 | 23.26 |
| KER | 1.54 | 2.06 | 1.93 ± 0.01 | 1.29–2.25 | 9.33 | 1.92 ± 0.02 | 1.23–2.59 | 10.83 | 1.89 ± 0.02 | 1.37–2.58 | 10.84 |

SE, standard error; CV, coefficient of variance; RRL, rough rice length; RRB, rough rice breadth; RRLB, rough rice length to breadth ratio; MRL, milled rice length; MRB, milled rice breadth; MRLB, milled rice length to breadth ratio; CKL, cooked kernel length; CKB, cooked kernel breadth; CKLB, cooked kernel length to breadth ratio; KER, kernel elongation ratio.

25.86% for MRLB during *Kharif* 2019, while during *Kharif* 2020, it ranged from 6.40% for CKB to 24.83% in RRLB. During *Kharif* 2021, PCV was minimum for MRB (5.14%) and maximum for RRLB (23.67%). The PCV for KER ratio was 9.33%, 10.83%, and 10.84% during *Kharif* 2019, *Kharif* 2020, and *Kharif* 2021, respectively (Table 1; Figure 1). The transgressive segregation was observed for the traits RRL, MRL, RRB, MRB, CKB, and KER (Figure 2).

## Correlation between the grain dimension traits in the RIL population

The significant positive correlation was observed among the grain length and length/breadth (rough rice, milled, and cooked) traits. However, there was a significant negative correlation between the grain length and breadth traits. Furthermore, there was a significant positive correlation between CKL and KER; while negative correlation was observed between CKB and CKL, MRLB, and KER; and no correlation between KER and rough rice dimension traits (Figure 3).

## Construction of a genetic linkage map

A total of 116 markers polymorphic between Sonasal and PB1121 were used to construct the linkage map. A total of 12 linkage groups were constructed with a cumulative map distance of 2,469.38 cM. The number of markers on different chromosomes ranged from 4 on chromosome 10 to 16 on chromosome 3 (Table 2). The average genetic distance between two adjacent markers was 21.29 cM.

## QTLs mapped

QTLs were identified for ten grain dimension traits using ICIM. A total of twenty-three QTLs were detected on six different chromosomes (chromosomes 2, 3, 4, 6, 7, and 8) for grain dimension-related traits (Figure 4). For each of these grain dimension traits, the proportion of phenotypic variance explained (PVE), its additive effect, nearest left marker, and nearest right marker were determined (Table 3).

## QTL for rough rice length and breadth

A total of eight QTLs associated with rough rice dimension traits were identified on chromosomes 3, 4, and 7 across 3 years of testing. A novel QTL *qRRL4.1* flanking the markers RM551 and HvSSR04-15 was identified on chromosome 4 with PVE of 10.86% and 18.11% during the years 2019 and 2021, respectively. The PVE of *qRRL3.1* ranged from 51.37% to 65.92% and the PVE of *qRRL7.1* ranged from 19.19% to 21.95% during 3 years of testing. The QTLs *qRRL3.1* and *qRRL7.1* correspond to already cloned genes *GS3* and *qGRL7.1*, respectively.

The trait RRB was found to mainly governed by two major QTLs *qRRB3.1* (PVE: 3.37%) and *qRRB7.1* (PVE: 12.49% and 23.48%) and for the trait RRLB, two robust QTLs *qRRLB3.1* (PVE: 30.11–46.15%) and *qRRLB7.1* (PVE: 22.41–29.21%) were mapped on chromosome 3 and chromosome 7 which correspond to the major genes *GS3* and *qGRL7.1*, respectively (Table 3). A major novel QTL *qRRLB4.1* flanking the region RM551 and HvSSR04-15 was identified for the trait RRLB on chromosome 4 during *Kharif* 2020 and *Kharif* 2021 which explained the phenotypic variance of 10.03%–20.51%, respectively.

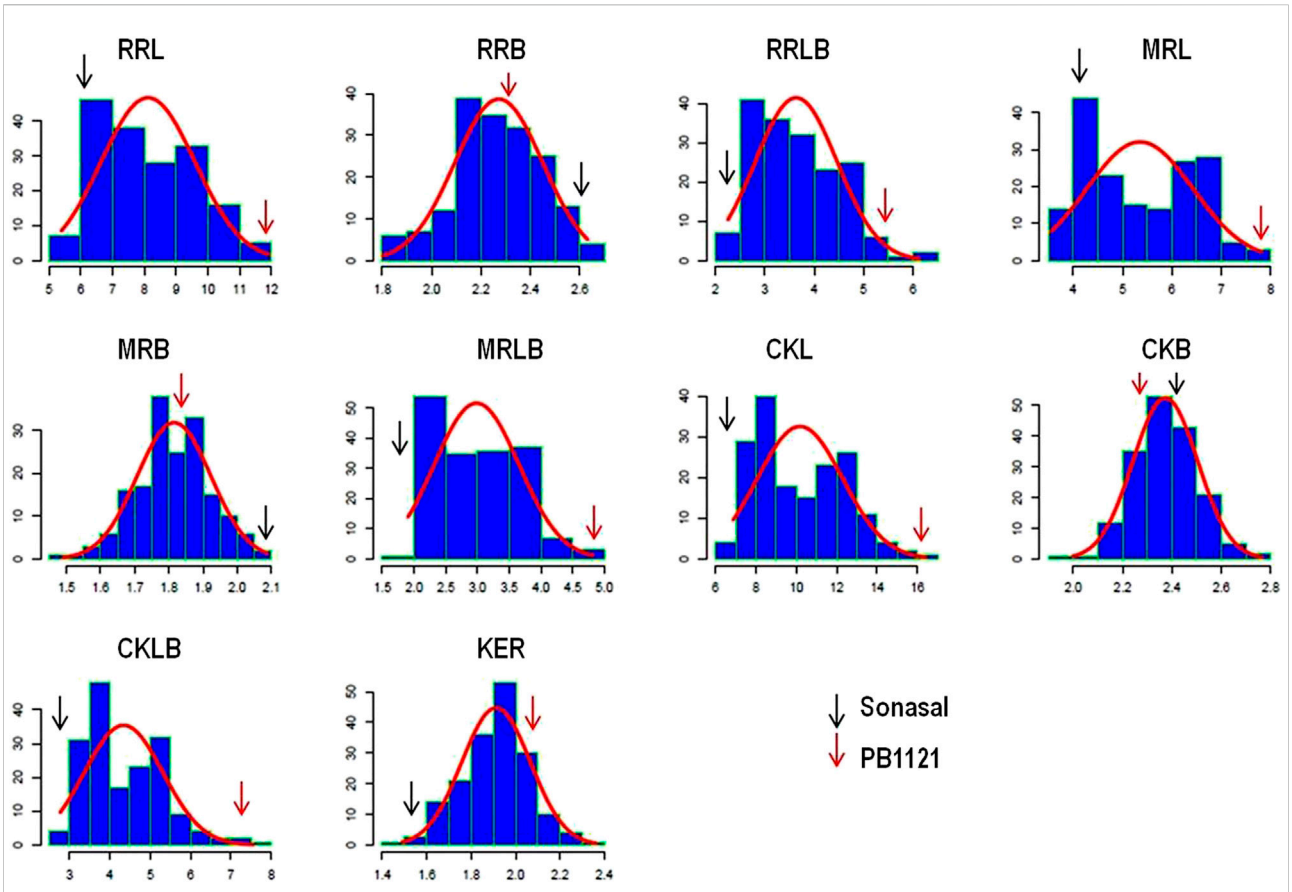



FIGURE 2
Distribution pattern of grain dimension traits in an RIL population. rough rice length (RRL), rough rice breadth (RRB), rough rice length to breadth ratio (RRLB), milled rice length (MRL), milled rice breadth (MRB), milled rice length to breadth ratio (MRLB), cooked kernel length (CKL), cooked kernel breadth (CKB), cooked kernel length to breadth ratio (CKLB), and kernel elongation ratio (KER).

## QTLs for milled rice length and breadth

A total of four QTLs were found to govern the trait MRL. The major QTL *qMRL3.1* was identified on chromosome 3 with PVE of 64.59%–76.14% during 3 years of testing. The QTL *qMRL7.1* was mapped on chromosome 7 which accounted for the PVE of 7.36 and 8.72% during *Kharif* 2019 and *Kharif* 2020, respectively. Two novel QTLs, namely, *qMRL3.2* flanked by the markers RM15247 and RM15281; and *qMRL4.1* flanked by the markers RM551 and HvSSR04-15 located on chromosomes 3 and 4, respectively, were mapped. The PVE of QTL *qMRL3.1* was 12.92% during *Kharif* 2020 while, the PVE of *qMRL4.1* was 14.34% and 13.65% during *Kharif 2019* and *Kharif 2021,* respectively.

For the trait MRB, a robust QTL *qMRB7.1* was identified on chromosome 7 with PVE of 15.54% and 10.86% during *Kharif* 2019 and *Kharif* 2020, respectively. For the trait MRLB, two robust QTLs, namely, *qMRLB3.1* and *qMRLB7.1* were mapped on chromosomes 3 and 7, respectively. Another novel QTL *qMRLB4.1* flanked by the markers RM551 and HvSSR04-15 with PVEs of 9.83% and 14.30% during *kharif* 2019 and *kharif* 2020, respectively.

## QTLs for cooked kernel length and breadth

CKL is one of the key Basmati traits which determine the cooking quality. In the current population, four QTLs governing the trait CKL were identified of which three QTLs namely; *qCKL3.2, qCKL4.1,* and *qCKL6.1* were identified to be novel. A major QTL *qCKL3.1* with PVE ranging from 29.81% to 58.26% and an additive effect ranging from 1.12 to 1.62 mm was identified during all the 3 years of testing. The QTL *qCKL3.2* (PVE: 13.01%–13.96%) flanked by RM15247 and RM15281 and the QTL *qCKL4.1* (PVE: 11.26%–11.89%) flanked by the markers RM551 and HvSSR04-15 were identified during *Kharif* 2019 and

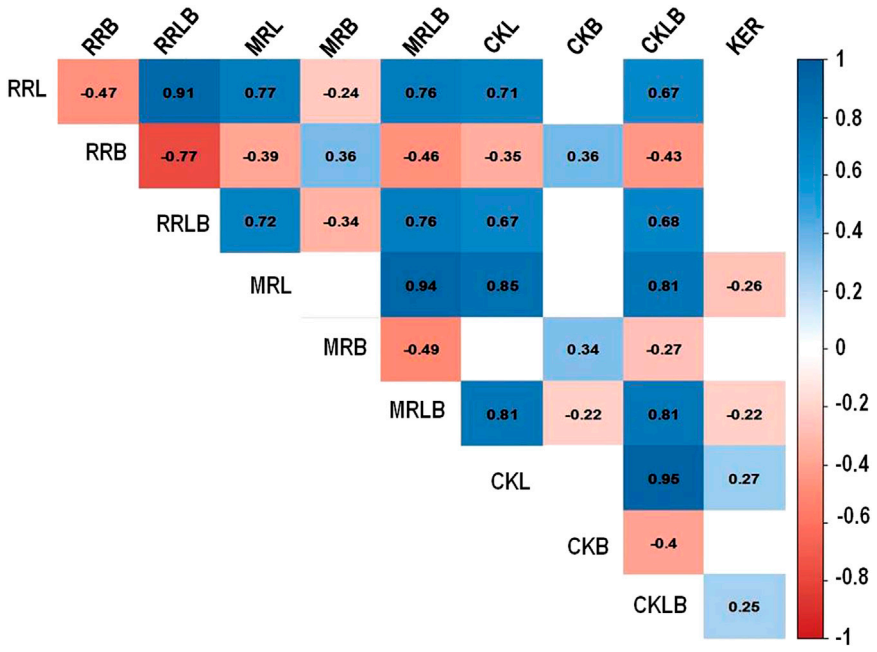



FIGURE 3
Correlation among grain dimension traits in rice. Rough rice length (RRL), rough rice breadth (RRB), rough rice length to breadth ratio (RRLB), milled rice length (MRL), milled rice breadth (MRB), milled rice length to breadth ratio (MRLB), cooked kernel length (CKL), cooked kernel breadth (CKB), cooked kernel length to breadth ratio (CKLB), and kernel elongation ratio (KER).

TABLE 2 Chromosome-wise number of the markers used for construction of linkage map, map length, and average marker interval.

| Chromosome | No. of markers | Map length (cM) |
|---|---|---|
| 1 | 15 | 325.04 |
| 2 | 11 | 287.46 |
| 3 | 16 | 349.74 |
| 4 | 8 | 183.98 |
| 5 | 11 | 192.93 |
| 6 | 8 | 189.63 |
| 7 | 11 | 165.21 |
| 8 | 10 | 160.65 |
| 9 | 6 | 159.70 |
| 10 | 4 | 136.33 |
| 11 | 9 | 164.54 |
| 12 | 7 | 154.17 |
| **TOTAL** | **116** | **2469.38** |

cM, centi Morgan.

*Kharif* 2020. Another novel QTL, *qCKL6.1* was mapped on chromosome 6 with PVE of 5.60% during *Kharif* 2020. The promising QTL *qCKL3.2* was bracketed between the markers RM15247 and RM15281 which spans a region of 0.78 Mb (17.90 Mb–18.68 Mb) on chromosome 3. The region comprises of 114 annotated gene models, of which 18 genes were putative candidates which may play a role in determining the grain length (Table 4).

## QTLs for kernel elongation ratio

KER is one of the key traits of Basmati rice which determines the extent of the linear elongation of the grain upon cooking. A novel QTL *qKER2.1* was mapped on chromosome 2 flanked by markers RM424 and RM452 with PVE of 9.56% during *Kharif* 2020.

# Discussion

Rice grain shape is mainly determined by the grain length and breadth. These traits are under the influence of a few large effects and multiple small effect genes, thus possessing complex patterns of inheritance. All the grain dimension traits showed a high degree of phenotypic variation in the RIL population during all the 3 years of testing. Furthermore, the grain dimension traits showed a near-normal distribution with slight skewness toward either of the parental lines which indicated the influence of the combination of few oligo-genes

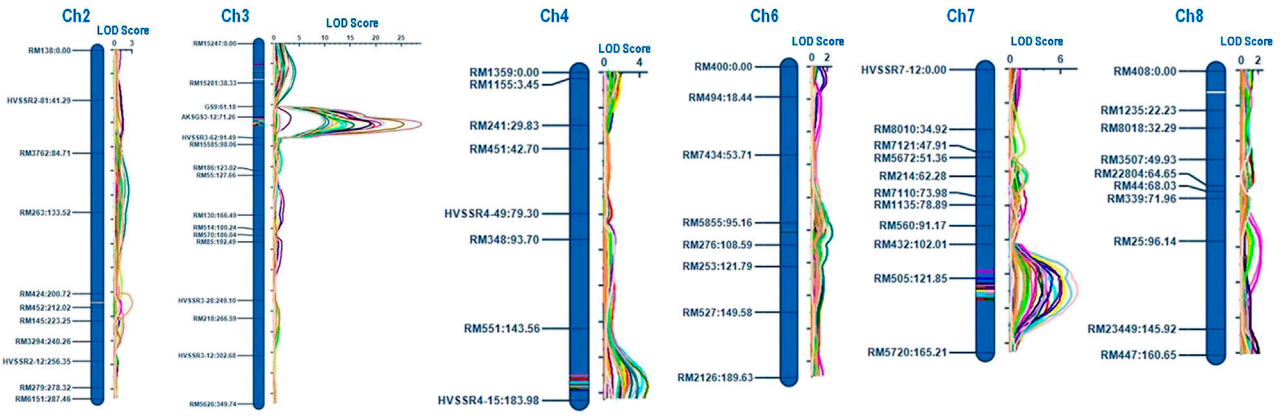



FIGURE 4
Graphical representation of identified QTLs on chromosomes.

and many minor genes in determining these traits. The significant positive correlation was observed among length-related traits while a significant negative correlation was observed between the length- and breadth-related traits. The observed correlation among the traits in the present study was consistent with the previous studies (Amarawathi et al., 2008; Bai et al., 2010).

The linkage map comprising of 12 linkage groups with a cumulative map distance of 2,469.38 cM was constructed using 116 polymorphic markers with an average marker interval distance of 21.29 cM. Although the average marker interval distance was 21.29 cM, the genetic distances between neighboring markers varied across the linkage groups due to the spatial variation in recombination frequencies. Sonasal is a short grain aromatic rice variety and PB1121 is an extra-long grain Basmati rice variety with sufficient genomic and phenotypic diversity among the parental lines. It was demonstrated that there exists tremendous diversity between the aromatic short grain and Basmati varieties (Garris et al., 2005), indicating the existence of tremendous structural changes, which may lead to the occurrence of segregation distortion at a certain loci in the mapping population generated. There are several possible reasons for segregation distortion, such as gametic or zygotic selection, genetic incompatibility, chromosome rearrangement, pollen competition, preferential fertilization etc., (Xu et al., 1997).

The genetic and molecular dissection of grain dimension parameters in rice have received a lot of attention, and until now, several QTLs have been identified and few genes have been functionally characterized (www.gramene.org). The major focus has been to understand the genetics and the identification of genomic regions governing the grain length and grain width in non-Basmati rice genotypes. This has resulted in the identification of major effect QTLs/genes such as *GS3*, a major gene governing the grain length and weight which was identified in a population generated from a cross Minghui 63/Chaun 7; *GW2*, a major gene for the grain width and weight mapped in a population generated from a cross WY3/Fengaizhan 1; *GS7*, associated with grain length was mapped using a population generated from a cross D50/HB277; *qSW5*, associated with the grain breadth was mapped in a population Nipponbare/Kasalath etc.,

Basmati is a unique combination of grain and cooking quality parameters with a strong and pleasing aroma. However, limited attempts were made to understand the grain and cooked rice dimension-related traits in Basmati genotypes. Earlier, in a population generated from the cross Basmati 385 and HJX74, a major gene *GW8* was identified. In the present study, a RIL population was generated using highly contrasting parental lines for the target traits. The major effect QTLs *viz., qRRL3.1, qRRLB3.1, qMRL3.1, qMRLB3.1,* and *qCKL3.1* were found to be co-mapped to a genomic region on chromosome 3 which is known to possess the *GS3* gene. The functional allele of *GS3* reduces the grain size and weight (Fan et al., 2009). In *GS3* a C to A mutation in the exon 2 introduces a stop codon leading to long grains. Based on this mutation, a CAPS marker SF28 was developed for its use in molecular breeding (Fan et al., 2009). Furthermore, previously we identified a novel indel *aksGS3-12* in the short grain genotype Sonasal, which was predominant in short grain landraces (Anand et al., 2013b). The functional allele at SF28 reduces the grain size, and if coupled with the 342bp deletion at *aksGS3-12*, it may further reduce the grain size, as in Sonasal. In the current study, it has been demonstrated that *GS3* has a key role in governing rough, milled, and cooked rice length as well as MRLB. The other major effect QTLs identified were *qRRL7.1, qRRB7.1, qMRL7.1, qMRB7.1,* and *qMRLB7.1* which were co-mapped to the genomic region on chromosome 7. This region co-localizes with the earlier reported QTL *qGRL7.1* (Amarawathi et al., 2008).

The novel QTLs *viz., qMRL3.2*, *qCKL3.2,* and *qCKLB3.2* are co-mapped between the markers RM15247 and RM15281 on

10.3389/fgene.2022.932166

TABLE 3 QTLs identified for grain quality parameters by ICIM for three seasons.

| QTL Name | Trait name | Chr | Position | Left marker | Right marker | LOD | PVE (%) | Add |
|---|---|---|---|---|---|---|---|---|
| *qRRL3.1* | RRL_2019 | 3 | 79 | aksGS3-12 | HvSSR03-62 | 19.67 | 65.92 | 1.30 |
| | RRL_2020 | 3 | 79 | aksGS3-12 | HvSSR03-62 | 15.74 | 62.57 | 1.13 |
| | RRL_2021 | 3 | 79 | aksGS3-12 | HvSSR03-62 | 13.53 | 51.37 | 1.09 |
| *qRRL4.1* | RRL_2019 | 4 | 172 | RM551 | HvSSR04-15 | 2.59 | 10.86 | 0.53 |
| | RRL_2021 | 4 | 174 | RM551 | HvSSR04-15 | 4.35 | 18.11 | 0.65 |
| *qRRL7.1* | RRL_2019 | 7 | 134 | RM505 | RM5720 | 4.47 | 20.86 | 0.74 |
| | RRL_2020 | 7 | 134 | RM505 | RM5720 | 3.89 | 19.19 | 0.63 |
| | RRL_2021 | 7 | 132 | RM505 | RM5720 | 5.35 | 21.95 | 0.71 |
| *qRRB3.1* | RRB_2020 | 3 | 72 | aksGS3-12 | HvSSR03-62 | 3.37 | 7.96 | −0.07 |
| *qRRB7.1* | RRB_2019 | 7 | 122 | RM505 | RM5720 | 4.95 | 12.49 | −0.07 |
| | RRB_2020 | 7 | 128 | RM505 | RM5720 | 6.04 | 23.48 | −0.12 |
| *qRRLB3.1* | RRLB_2019 | 3 | 78 | aksGS3-12 | HvSSR03-62 | 12.67 | 46.15 | 0.56 |
| | RRLB_2020 | 3 | 77 | aksGS3-12 | HvSSR03-62 | 10.17 | 33.10 | 0.51 |
| | RRLB_2021 | 3 | 77 | aksGS3-12 | HvSSR03-62 | 9.79 | 30.11 | 0.50 |
| *qRRLB4.1* | RRLB_2020 | 4 | 176 | RM551 | HvSSR04-15 | 3.01 | 10.03 | 0.28 |
| | RRLB_2021 | 4 | 174 | RM551 | HvSSR04-15 | 4.88 | 20.51 | 0.41 |
| *qRRLB7.1* | RRLB_2019 | 7 | 129 | RM505 | RM5720 | 6.45 | 22.41 | 0.39 |
| | RRLB_2020 | 7 | 129 | RM505 | RM5720 | 7.92 | 26.35 | 0.45 |
| | RRLB_2021 | 7 | 131 | RM505 | RM5720 | 7.23 | 29.21 | 0.49 |
| *qMRL3.1* | MRL_2019 | 3 | 78 | aksGS3-12 | HvSSR03-62 | 25.46 | 64.59 | 0.89 |
| | MRL_2020 | 3 | 78 | aksGS3-12 | HvSSR03-62 | 20.91 | 76.14 | 0.91 |
| | MRL_2021 | 3 | 78 | aksGS3-12 | HvSSR03-62 | 28.91 | 75.67 | 1.01 |
| *qMRL3.2* | MRL_2020 | 3 | 20 | RM15247 | RM15281 | 2.95 | 12.92 | 0.38 |
| *qMRL4.1* | MRL_2019 | 4 | 177 | RM551 | HvSSR04-15 | 4.90 | 14.34 | 0.42 |
| | MRL_2021 | 4 | 170 | RM551 | HvSSR04-15 | 2.98 | 13.65 | 0.43 |
| *qMRL7.1* | MRL_2019 | 7 | 127 | RM505 | RM5720 | 3.13 | 7.36 | 0.30 |
| | MRL_2021 | 7 | 131 | RM505 | RM5720 | 2.58 | 8.72 | 0.34 |
| *qMRB7.1* | MRB_2019 | 7 | 126 | RM505 | RM5720 | 4.27 | 15.54 | −0.07 |
| | MRB_2020 | 7 | 118 | RM432 | RM505 | 3.14 | 10.86 | −0.05 |
| *qMRLB3.1* | MRLB_2019 | 3 | 78 | aksGS3-12 | HvSSR03-62 | 18.90 | 54.77 | 0.54 |
| | MRLB_2020 | 3 | 78 | aksGS3-12 | HvSSR03-62 | 19.20 | 66.60 | 0.52 |
| | MRLB_2021 | 3 | 77 | aksGS3-12 | HvSSR03-62 | 21.56 | 58.28 | 0.56 |
| *qMRLB4.1* | MRLB_2019 | 4 | 178 | RM551 | HvSSR04-15 | 3.72 | 9.83 | 0.23 |
| | MRLB_2021 | 4 | 174 | RM551 | HvSSR04-15 | 3.70 | 14.30 | 0.28 |
| *qMRLB7.1* | MRLB_2019 | 7 | 127 | RM505 | RM5720 | 5.86 | 14.52 | 0.28 |
| | MRLB_2020 | 7 | 125 | RM505 | RM5720 | 3.85 | 9.04 | 0.19 |
| | MRLB_2021 | 7 | 130 | RM505 | RM5720 | 4.07 | 12.11 | 0.25 |
| *qCKL3.1* | CKL_2019 | 3 | 75 | aksGS3-12 | HvSSR03-62 | 10.15 | 24.49 | 1.18 |
| | CKL_2020 | 3 | 77 | aksGS3-12 | HvSSR03-62 | 10.78 | 29.81 | 1.12 |
| | CKL_2021 | 3 | 79 | aksGS3-12 | HvSSR03-62 | 14.22 | 58.26 | 1.62 |
| *qCKL3.2* | CKL_2019 | 3 | 27 | RM15247 | RM15281 | 3.88 | 13.01 | 0.86 |
| | CKL_2020 | 3 | 28 | RM15247 | RM15281 | 4.27 | 13.96 | 0.77 |
| *qCKL4.1* | CKL_2019 | 4 | 177 | RM551 | HvSSR04-15 | 3.83 | 11.89 | 0.83 |
| | CKL_2020 | 4 | 175 | RM551 | HvSSR04-15 | 3.11 | 11.26 | 0.70 |
| *qCKL6.1* | CKL_2020 | 6 | 101 | RM5855 | RM276 | 2.62 | 5.60 | 0.49 |
| *qCKLB3.1* | CKLB_2019 | 3 | 73 | aksGS3-12 | HvSSR03-62 | 12.82 | 32.51 | 0.61 |
| | CKLB_2020 | 3 | 75 | aksGS3-12 | HvSSR03-62 | 9.41 | 23.59 | 0.45 |

(Continued on following page)

TABLE 3 (*Continued*) QTLs identified for grain quality parameters by ICIM for three seasons.

| QTL Name | Trait name | Chr | Position | Left marker | Right marker | LOD | PVE (%) | Add |
|---|---|---|---|---|---|---|---|---|
| | CKLB_2021 | 3 | 77 | aksGS3-12 | HvSSR03-62 | 11.51 | 37.27 | 0.67 |
| *qCKLB3.2* | CKLB_2020 | 3 | 35 | RM15247 | RM15281 | 3.32 | 7.55 | 0.25 |
| *qKER2.1* | KER_2020 | 2 | 208 | RM424 | RM452 | 3.05 | 9.56 | 0.07 |

Chr, chromosome; LOD, logarithm of the odds; PVE, phenotypic variance; Add, additive effect; RRL, rough rice length; RRB, rough rice breadth; RRLB, rough rice length to breadth ratio; MRL, milled rice length; MRB, milled rice breadth; MRLB, milled rice length to breadth ratio; CKL, cooked kernel length; CKB, cooked kernel breadth; CKLB, cooked kernel length to breadth ratio; KER, kernel elongation ratio.

TABLE 4 The predicted gene models in the novel QTLs qCKL3.2/qMRL3.2

| Gene models | Start position (bp) | Stop position (bp) | Putative function |
|---|---|---|---|
| LOC_Os03g31480 | 17934776 | 17935790 | expansin precursor, putative, expressed |
| LOC_Os03g31550 | 17985563 | 17998498 | aldehyde oxidase, putative, expressed |
| LOC_Os03g31594 | 18026740 | 18041873 | jmjC domain containing protein, expressed |
| LOC_Os03g31679 | 18085558 | 18087371 | annexin A7, putative, expressed |
| LOC_Os03g31690 | 18087460 | 18093916 | GCN5-related N-acetyltransferase, putative, expressed |
| LOC_Os03g32030 | 18318940 | 18319467 | DNA-directed RNA polymerase II subunit RPB1, putative, expressed |
| LOC_Os03g32050 | 18325230 | 18332522 | peroxidase precursor, putative, expressed |
| LOC_Os03g32090 | 18355904 | 18359839 | pentatricopeptide, putative, expressed |
| LOC_Os03g32160 | 18397036 | 18401598 | calmodulin binding protein, putative, expressed |
| LOC_Os03g32180 | 18410063 | 18411482 | polygalacturonase inhibitor 1 precursor, putative, expressed |
| LOC_Os03g32220 | 18430111 | 18431229 | ZOS3-11 - C2H2 zinc finger protein, expressed |
| LOC_Os03g32230 | 18435989 | 18437090 | ZOS3-12 - C2H2 zinc finger protein, expressed |
| LOC_Os03g32270 | 18450996 | 18456312 | hydrolase, alpha-beta fold family domain containing protein, expressed |
| LOC_Os03g32314 | 18485606 | 18488371 | allene oxide cyclase 4, chloroplast precursor, putative, expressed |
| LOC_Os03g32580 | 18635876 | 18638518 | BRASSINOSTEROID INSENSITIVE 1-associated receptor kinase 1 precursor, putative, expressed |
| LOC_Os03g32590 | 18638520 | 18647607 | transcription initiation factor, putative, expressed |
| LOC_Os03g32620 | 18677824 | 18682844 | pentatricopeptide, putative, expressed |
| LOC_Os03g32630 | 18686894 | 18689388 | ABC transporter, ATP-binding protein, putative, expressed |

chromosome 3. The annotation of the QTL-bracketed region of 0.78 Mb identified 18 putative candidate gene models which may have a significant role in determining the kernel length. This genomic region possessed gene models which code for pentatricopeptides, brassinosteroid insensitive 1-associated receptor kinase 1 precursor and transcription factors. A maize QTL *qKW9* has been reported to encode a pentatricopeptide repeat protein which affects the maternal photosynthates available during grain filling (Huang et al., 2020), while, brassinosteroids are known to play a key role in determining grain size and plant architecture (Zhang et al., 2021). Other gene models code for proteins such as expansin, aldehyde oxidase, N-acetyltransferase, ploygalacturonase, allene oxide cyclase, etc., which are known to play important roles in cell wall loosening, elongation, and cell division (Agrawal et al., 2003; Riemann et al., 2013; Byeon and Back, 2014; Marowa et al., 2016; Lee and Back, 2017; Cao et al., 2021; Shi et al., 2021), which may therefore be associated with an increased grain length. Furthermore, fine-mapping and functional validation is required to pin-point the gene underlying this important QTL.

The KER is one of the important traits of Basmati rice which determines the degree of linear elongation after cooking. This trait is proportional to other economically important cooking quality traits such as volume expansion and water uptake. In the current study, we have identified a novel QTL *viz., qKER2.1* which has a large effect on KER in Basmati rice.

In all, the Basmati rice grain length is mainly governed by three loci, namely *GS3, qGL7.1,* and *qMRL3.2*. However, the cooked grain length is mainly governed by four loci, *viz., GS3, qGL7.1, qCKL3.2,* and *qCKL4.1*. Therefore, targeting these loci through marker-

assisted breeding can help in developing Basmati rice genotypes with extra-long slender grains and exceptional cooking quality traits.

## Conclusion

The present study identified novel QTLs governing grain dimension traits in a RIL population generated by crossing a short grain aromatic rice landrace, Sonasal and Pusa Basmati 1121, a Basmati rice variety possessing extra-long slender grains. Two previously reported QTLs and two novel QTLs identified in the current study were identified to influence the exceptional milled rice length and cooked kernel length in Basmati rice. One of the novel QTLs was bracketed to a genomic region of 0.78 Mb which possesses putative candidate genes. However, further fine mapping and map-based cloning is required to precisely identify the candidate genes governing the trait. Furthermore, the markers flanking the QTLs identified can be utilized in the marker-assisted breeding to develop Basmati rice varieties with extra-long slender grains and high kernel elongation upon cooking

## 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

AS: conceptualization. GK, RE, NS, and AS: methodology. AM: genotyping and phenotyping. RE and AM: data curation, statistical analysis. AM and RE: original draft preparation. RE, KV, and DD: editing. RE, GS, NM, PB, HB, AS, AK: supervision. All authors have read and agreed to the final version of the manuscript.

## Funding

This research was funded by the ICAR- "Network Project on Functional Genomics and Genetic Modification in Crops (Rice)."

## Acknowledgments

The study is part of the Ph.D. thesis research of AM and AM wishes to acknowledge ICAR-Indian Agricultural Research Institute, New Delhi, and Indian Council of Agriculture Research (ICAR) for providing all the necessary facilities.

## 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

Agrawal, G. K., Jwa, N. S., Shibato, J., Han, O., Iwahashi, H., Rakwal, R., et al. (2003). Diverse environmental cues transiently regulate *OsOPR1* of the "octadecanoid pathway" revealing its importance in rice defense/stress and development. *Biochem. Biophys. Res. Commun.* 310, 1073–1082. doi:10.1016/j.bbrc.2003.09.123

Amarawathi, Y., Singh, R., Singh, A. K., Singh, V. P., Mohapatra, T., Sharma, T. R., et al. (2008). Mapping of quantitative trait loci for basmati quality traits in rice (*Oryza sativa* L.). *Mol. Breed.* 21, 49–65. doi:10.1007/s11032-007-9108-8

Anand, D., Baunthiyal, M., Gopala Krishnan, S., Singh, N. K., Prabhu, K. V., Singh, A. K., et al. (2013a). Novel InDel variation in *GS3* locus and development of InDel based marker for marker assisted breeding of short grain aromatic rices. *J. Plant Biochem. Biotechnol.* 24, 120–127. doi:10.1007/s13562-013-0243-5

Anand, D., Baunthiyal, M., Singh, A., Gopala Krishnan, S., Singh, N. K., Prabhu, K. V., et al. (2013b). Validation of gene based marker-QTL association for grain dimension traits in rice. *J. Plant Biochem. Biotechnol.* 22, 467–473. doi:10.1007/s13562-012-0176-4

Arikit, S., Wanchana, S., Khanthong, S., Saensuk, C., Thianthavon, T., Vanavichit, A., et al. (2019). QTL-seq identifies cooked grain elongation QTLs near soluble starch synthase and starch branching enzymes in rice (*Oryza sativa* L.). *Sci. Rep.* 9, 8328. doi:10.1038/s41598-019-44856-2

Bai, X., Luo, L., Yan, W., Kovi, M. R., Zhan, W., Xing, Y., et al. (2010). Genetic dissection of rice grain shape using a recombinant inbred line population derived from two contrasting parents and fine mapping a pleiotropic quantitative trait locus *qGL7*. *BMC Genet.* 11, 16. doi:10.1186/1471-2156-11-16

Balram, M., Smriti, G., Trilochan, M., Rajender, P., Nagarajan, M., Vinod, K. K., et al. (2012). QTL analysis of novel genomic regions associated with yield and yield related traits in new plant type based recombinant inbred lines of rice (*Oryza sativa* L.). *BMC Plant Biol.* 12, 137. doi:10.1186/1471-2229-12-137

Bhatia, D., Wing, R. A., Yu, Y., Chougule, K., Kudrna, D., Lee, S., et al. (2018). Genotyping by sequencing of rice interspecific backcross inbred lines identifies QTLs for grain weight and grain length. *Euphytica* 214, 41. doi:10.1007/s10681-018-2119-1

Byeon, Y., and Back, K. (2014). An increase in melatonin in transgenic rice causes pleiotropic phenotypes, including enhanced seedling growth, delayed flowering, and low grain yield. *J. pineal Res.* 56, 408–414. doi:10.1111/jpi.12129

Cao, Y., Zhang, Y., Chen, Y., Yu, N., Liaqat, S., Wu, W., et al. (2021). *OsPG1* encodes a polygalacturonase that determines cell wall architecture and affects

resistance to bacterial blight pathogen in rice. *Rice* 14, 36. doi:10.1186/s12284-021-00478-9

Duan, P., Xu, J., Zeng, D., Zhang, B., Geng, M., Zhang, G., et al. (2017). Natural variation in the promoter of *GSE5* contributes to grain size diversity in rice. *Mol. Plant* 10, 685–694. doi:10.1016/j.molp.2017.03.009

Fan, C., Xing, Y., Mao, H., Lu, T., Han, B., Xu, C., et al. (2006). *GS3*, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. *Theor. Appl. Genet.* 12, 1164–1171. doi:10.1007/s00122-006-0218-1

Fan, C., Yu, S., Wang, C., and Xin, Y. (2009). A causal C–A mutation in the second exon of *GS3* highly associated with rice grain length and validated as a functional marker. *Theor. Appl. Genet.* 118, 465–472. doi:10.1007/s00122-008-0913-1

Garris, A. J., Tai, T. H., Coburn, J., Kresovich, S., and McCouch, S. (2005). Genetic structure and diversity in *Oryza sativa* L. *Genetics* 169, 1631–1638. doi:10.1534/genetics.104.035642

Hu, Z., Lu, S. J., Wang, M. J., He, H., Sun, L., Wang, H., et al. (2018). A novel QTL *qTGW3* encodes the GSK3/SHAGGY-Like Kinase *OsGSK5/OsSK41* that interacts with *OsARF4* to negatively regulate grain size and weight in rice. *Mol. Plant* 11, 736–749. doi:10.1016/j.molp.2018.03.005

Huang, J., Lu, G., Liu, L., Raihan, M. S., Xu, J., Jian, L., et al. (2020). The kernel size-related quantitative trait locus *qkw9* encodes a pentatricopeptide repeat protein that affects photosynthesis and grain filling. *Plant Physiol.* 183, 1696–1709. doi:10.1104/pp.20.00374

Ishimaru, K., Hirotsu, N., Madoka, Y., Murakami, N., Hara, N., Onodera, H., et al. (2013). Loss of function of the IAA-glucose hydrolase gene *TGW6* enhances rice grain weight and increases yield. *Nat. Genet.* 45, 707–711. doi:10.1038/ng.2612

Jiang, G. H., Hong, X. Y., Xu, C. G., Li, X. H., and He, Y. Q. (2005). Identification of quantitative trait loci for grain appearance and milling quality using a doubled-haploid rice population. *J. Integr. Plant Biol.* 47, 1391–1403. doi:10.1111/j.1744-7909.2005.00089.x

Kosambi, D. D. (1944). The estimation of map distance from recombination values. *Ann. Eugen.* 12, 172–175. doi:10.1111/j.1469-1809.1943.tb02321.x

Lander, E. S., and Botstein, D. (1989). Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121, 185–199. doi:10.1093/genetics/121.1.185

Lee, K., and Back, K. (2017). Overexpression of rice serotonin N-acetyltransferase 1 in transgenic rice plants confers resistance to cadmium and senescence and increases grain yield. *J. pineal Res.* 62, e12392. doi:10.1111/jpi.12392

Li, H., Ye, G., and Wang, J. (2007). A modified algorithm for the improvement of composite interval mapping. *Genetics* 175, 361–374. doi:10.1534/genetics.106.066811

Li, Y., Fan, C., Xing, Y., Jiang, Y., Luo, L., Sun, L., et al. (2011). Natural variation in *GS5* plays an important role in regulating grain size and yield in rice. *Nat. Genet.* 43, 1266–1269. doi:10.1038/ng.977

Mao, H., Sun, S., Yao, J., Wang, C., Yu, S., Xu, C., et al. (2010). Linking differential domain functions of the *GS3* protein to natural variation of grain size in rice. *Proc. Natl. Acad. Sci. U. S. A.* 107, 19579–19584. doi:10.1073/pnas.1014419107

Marowa, P., Ding, A., and Kong, Y. (2016). Expansins: Roles in plant growth and potential applications in crop improvement. *Plant Cell. Rep.* 35, 949–965. doi:10.1007/s00299-016-1948-4

Meng, L., Li, H., Zhang, L., and Wang, J. (2015). QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *Crop J.* 3, 269–283. doi:10.1016/j.cj.2015.01.001

Murray, M. G., and Thompson, W. F. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* 8, 4321–4325. doi:10.1093/nar/8.19.4321

Riemann, M., Haga, K., Shimizu, T., Okada, K., Ando, S., Mochizuki, S., et al. (2013). Identification of rice Allene Oxide Cyclase mutants and the function of jasmonate for defence against *Magnaporthe oryzae*. *Plant J.* 74, 226–238. doi:10.1111/tpj.12115

Shao, G., Wei, X., Chen, M., Tang, S., Luo, J., Jiao, G., et al. (2012). Allelic variation for a candidate gene for *GS7*, responsible for grain shape in rice. *Theor. Appl. Genet.* 125, 1303–1312. doi:10.1007/s00122-012-1914-7

Shi, X., Tian, Q., Deng, P., Zhang, W., and Jing, W. (2021). The rice aldehyde oxidase *OsAO3* gene regulates plant growth, grain yield, and drought tolerance by participating in ABA biosynthesis. *Biochem. Biophys. Res. Commun.* 548, 189–195. doi:10.1016/j.bbrc.2021.02.047

Shomura, A., Izawa, T., Ebana, K., Ebitani, T., Kanegae, H., Konishi, S., et al. (2008). Deletion in a gene associated with grain size increased yields during rice domestication. *Nat. Genet.* 40, 1023–1028. doi:10.1038/ng.169

Song, X. J., Huang, W., Shi, M., Zhu, M. Z., and Lin, H. X. (2007). A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nat. Genet.* 39, 623–630. doi:10.1038/ng2014

Wang, C., Chen, S., and Yu, S. (2011). Functional markers developed from multiple loci in *GS3* for fine marker assisted selection of grain length in rice. *Theor. Appl. Genet.* 122, 905–913. doi:10.1007/s00122-0101497-0

Wang, S., Wu, K., Yuan, Q., Liu, X., Liu, Z., Lin, X., et al. (2012). Control of grain size, shape and quality by *OsSPL16* in rice. *Nat. Genet.* 44, 950–954. doi:10.1038/ng.2327

Xu, Y., Zhu, L., Xiao, J., Huang, N., and McCouch, S. R. (1997). Chromosomal regions associated with segregation distortion of molecular markers in F2, backcross, double haploid, and recombinant inbred populations in rice (*Oryza sativa* L.). *Mol. Gen. Genet.* 253, 535–545. doi:10.1007/s004380050355

Zhang, J., Gao, X., Cai, G., Wang, Y., Li, J., Du, H., et al. (2021). An adenylate kinase *OsAK3* involves brassinosteroid signaling and grain length in rice (*Oryza sativa* L.). *Rice* 14, 105. doi:10.1186/s12284-021-00546-0

Zhang, X., Wang, J., Huang, J., Lan, H., Wang, C., Yin, C., et al. (2012b). Rare allele of *OsPPKL1* associated with grain length causes extra-large grain and a significant yield increase in rice. *Proc. Natl. Acad. Sci. U. S. A.* 109, 21534–21539. doi:10.1073/pnas.1219776110

Zhu, Y. J., Sun, Z. C., Niu, X. J., Ying, J. Z., Fan, Y. Y., Mou, T. M., et al. (2019). Dissection of three quantitative trait loci for grain size on the long arm of chromosome 10 in rice (*Oryza sativa* L.). *PeerJ* 7, e6966. doi:10.7717/peerj.6966

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 24 August 2022
DOI 10.3389/fgene.2022.933560



OPEN ACCESS

EDITED BY
Philomin Juliana,
International Maize and Wheat Improvement Center, Mexico

REVIEWED BY
Muhammad Waheed Riaz,
Zhejiang Agriculture and Forestry University, China
Ashutosh Pandey,
National Institute of Plant Genome Research (NIPGR), India

*CORRESPONDENCE
Upendra Kumar,
baliyan.upendra@gmail.com

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

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

RECEIVED 01 May 2022
ACCEPTED 04 July 2022
PUBLISHED 24 August 2022

CITATION
Mehla S, Kumar U, Kapoor P, Singh Y, Sihag P, Sagwal V, Balyan P, Kumar A, Ahalawat N, Lakra N, Singh KP, Pesic V, Djalovic I, Mir RR and Dhankher OP (2022), Structural and functional insights into the candidate genes associated with different developmental stages of flag leaf in bread wheat (Triticum aestivum L.).
*Front. Genet.* 13:933560.
doi: 10.3389/fgene.2022.933560



# Structural and functional insights into the candidate genes associated with different developmental stages of flag leaf in bread wheat (*Triticum aestivum* L.)

Sheetal Mehla[1†], Upendra Kumar[1†*], Prexha Kapoor[1], Yogita Singh[1], Pooja Sihag[1], Vijeta Sagwal[1], Priyanka Balyan[2], Anuj Kumar[3,4], Navjeet Ahalawat[1], Nita Lakra[1], Krishna Pal Singh[5,6], Vladan Pesic[7], Ivica Djalovic[8], Reyazul Rouf Mir[9] and Om Parkash Dhankher[10]

[1]Department of Molecular Biology, Biotechnology and Bioinformatics, College of Biotechnology, CCS Haryana Agricultural University, Hisar, India, [2]Department of Botany, Deva Nagri P. G. College, CCS University, Meerut, India, [3]Shantou University Medical College, Shantou, China, [4]Dalhousie University, Halifax, NS, Canada, [5]Biophysics Unit, College of Basic Sciences and Humanities, GB Pant University of Agriculture and Technology, Pantnagar, India, [6]Vice-Chancellor's Secretariat, Mahatma Jyotiba Phule Rohilkhand University, Bareilly, India, [7]Department of Genetics and Plant Breeding, Faculty of Agriculture, University of Belgrade, Belgrade, Serbia, [8]Institute of Field and Vegetable Crops, National Institute of the Republic of Serbia, Maxim Gorki 30, Novi Sad, Serbia, [9]Division of Genetics and Plant Breeding, Sher-e-Kashmir University of Agricultural Sciences and Technology of Kashmir (SKUAST-Kashmir), Srinagar, India, [10]Stockbridge School of Agriculture, University of Massachusetts Amherst, Amherst, MA, United States

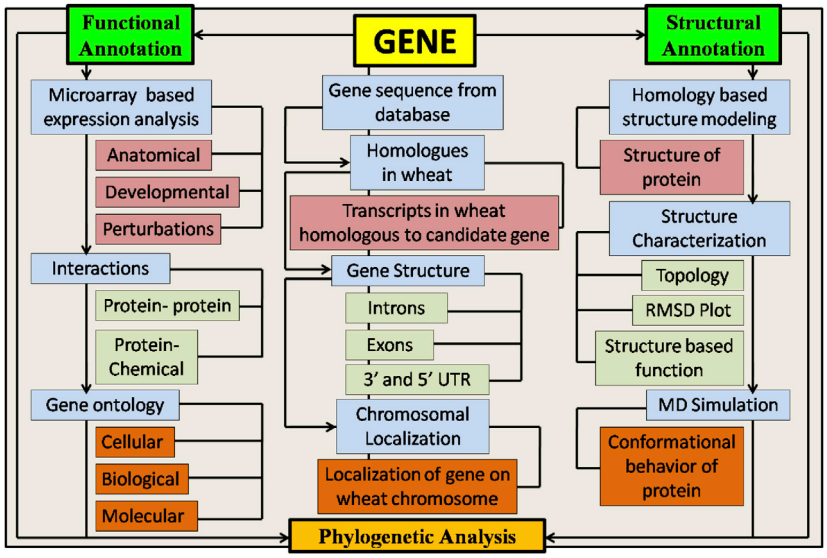
Grain yield is one of the most important aims for combating the needs of the growing world population. The role of development and nutrient transfer in flag leaf for higher yields at the grain level is well known. It is a great challenge to properly exploit this knowledge because all the processes, starting from the emergence of the flag leaf to the grain filling stages of wheat (*Triticum aestivum* L.), are very complex biochemical and physiological processes to address. This study was conducted with the primary goal of functionally and structurally annotating the candidate genes associated with different developmental stages of flag leaf in a comprehensive manner using a plethora of *in silico* tools. Flag leaf-associated genes were analyzed for their structural and functional impacts using a set of bioinformatics tools and algorithms. The results revealed the association of 17 candidate genes with different stages of flag leaf development in wheat crop. Of these 17 candidate genes, the expression analysis results revealed the upregulation of genes such as *TaSRT1-5D*, *TaPNH1-7B*, and *TaNfl1-2B* and the downregulation of genes such as *TaNAP1-7B*, *TaNOL-4D*, and *TaOsl2-2B* can be utilized for the generation of high-yielding wheat varieties. Through MD simulation and other *in silico* analyses, all these proteins were found to be stable. Based on the outcome of bioinformatics and molecular analysis, the identified candidate genes were found to play

10.3389/fgene.2022.933560

principal roles in the flag leaf development process and can be utilized for higher-yield wheat production.

KEYWORDS

flag leaf development, homology modeling, gene expression analysis, qRT-PCR (quantitative real-time polymerase chain reaction), MD simulation




GRAPHICAL ABSTRACT

## 1 Introduction

Bread wheat (*Triticum aestivum* L.) has been considered one of the initial food crops that have been domesticated as a staple food for thousands of years in the major civilizations of West Asia, North America, and Europe (Giraldo et al., 2019). The area under the cultivation of wheat is 239.63 million hectares in the world and 29.31 million hectares in India, with a total production of 899.37 million metric tonnes in the world and 103.59 million metric tonnes in India (FAO, 2020). The cereal utilization forecast for the years 2020–21 has been raised to 766 million tonnes on a global scale, 54 million tonnes (2%) above the 2019–20 level and 4.3 million tonnes higher than previous reports[1]. The grain yield in wheat is driven by the amount of energy harvested by upper leaves. Hence, it is said that wheat is a limited-source crop. Approximately 75% of the overall yield is contributed by the flag leaf and spike in wheat[2]. As evident from the previous reports, the supply of wheat needs to be increased by one billion metric tonnes to meet the needs of an increasing population. All parts of the plant contribute to the development of spikes in cereal crops (Blade and Baker, 1991). However, it was discovered that the uppermost three leaves determine cereal yield potential due to their importance in grain filling (Bĭrsĭn, 2005). It was found that the defoliation of flag leaf generated an 18–30% loss in grain yield in wheat (Youssef and Salem, 1976; Banitaba et al., 2007). Therefore, it is clear that studying the molecular pathways of flag leaf development is critical for gaining a better understanding of its function and meeting the needs of its exploration to combat the need for food supply.

Photosynthesis is the ultimate process for increasing grain yield in wheat, and the flag leaf is the part of the plant that receives the highest amount of sunlight for the preparation of food. A flag leaf is referred to as the last leaf to emerge from indeterminate plants of the Poaceae family (Palmer et al., 2015). In cereals, the flag leaf is known to provide the largest portion of photoassimilates for grain filling (Biswal and Kohli, 2013). Wheat is known to be a source-limited crop, as its yield is primarily driven through the amount of energy captured by the upper leaves (Koshkin and Tararina, 1989). Morphological characteristics such as the shape and area of the flag leaf contribute to wheat productivity (Liu K. et al., 2018). Yield-related traits were found to be positively associated with the size of flag leaf in wheat (Fan et al., 2015; Liu K. et al., 2018, Liu et al.,

1 www.agrochart.com

2 43bf642e28db4188a97b09ffb580409bd5e7a2a4 @www.fas.scot.

2018 Y.). Duwayri (1984) reported that the removal of flag leaves resulted in reduced grain yield and kernel numbers. Different stages occur during the development of a flag leaf, each with its specific function. Flag leaves provide the storage organs with photosynthetic products (Yan et al., 2020); hence, it becomes necessary to identify candidate genes associated with different stages of flag leaf development. The synergy of leaf development and environmental factors results in the complex trait of natural senescence (Liu et al., 2016).

Despite the significance of flag leaves in wheat production, a lesser amount of information about their molecular, morphological, and physiological characteristics is available. In the present study, the main objective was to initially identify and then characterize the genes associated with flag leaf development and perform a detailed analysis of proteins encoded in wheat. Through computational analysis, 17 candidate genes were found to be involved in flag leaf development from reference sequences of the wheat genome. These genes were further structurally and functionally annotated through various *in silico* tools and automated servers. Structural annotation involved the elucidation of gene structure and protein modeling followed by molecular dynamics simulation configuration at different nanoseconds and their physiochemical properties, whereas functional annotation involved gene ontology, phylogenetic relationships, identification of functional domains, molecular interaction networks, and microarray profiling.

# 2 Materials and methods

## 2.1 Structural annotation

### 2.1.1 Identification and chromosomal localization of candidate genes involved in wheat flag leaf development

The protein sequences of candidate genes associated with the development of flag leaf were retrieved from different model plants such as *Arabidopsis thaliana* and *Oryza sativa* using the NCBI database (www.ncbi.nlm.nih.gov) (Jenuth, 2000). The orthologs of these proteins responsible for flag leaf development in wheat were identified by performing BLASTp (http://plants.ensembl.org/Triticum_aestivum/Tools/Blast) against the Ensembl Plants database (Bolser et al., 2015) with a selection of *T. aestivum,* which contains data from different assemblies, namely, PGSBv2.0, Elv1.1, and WeebilV1. The genes associated with these predicted proteins were further annotated for their chromosomal localization in the wheat genome (Kumar et al., 2018b).

### 2.1.2 Identification of gene structure

The Gene Structure Display 2.0 (GSDV2) (http://gsds.gao-lab.org/) server was utilized to explore information regarding the detailed structure of genes. The coding and genomic sequences of the predicted proteins were taken into account for marking three major regions of a gene: 3′ and 5′ UTRs, exons, and introns.

### 2.1.3 Physiochemical properties of candidate proteins

Physiochemical properties are the basic properties of a particular protein, such as the theoretical isoelectric point, instability index, molecular weight, and aliphatic index. The ProtParam server (https://web.expasy.org/protparam/) (Garg et al., 2016) was used to analyze all these parameters of candidate proteins associated with flag leaf development in the wheat crop.

### 2.1.4 Homology modeling

3D structural modeling of all 17 candidate proteins associated with the development of flag leaf was performed through the SWISS-MODEL Server (https://swissmodel.expasy.org/) based on the homology approach. Protein sequences were placed in the target sequence dialog box and a template search was initiated. A template was selected from a set of templates ranging from 3 to 50 concerning PSI BLAST and HHblits. Based on the similarity, the most suitable template was selected for performing homology-based modeling of candidate genes. The structures of proteins after modeling were further analyzed with UCSF ChimeraX software (https://www.cgl.ucsf.edu/chimerax/) and the ProFunc server (http://www.ebi.ac.uk/thornton-srv/databases/ProFunc/) for proper characterization both at the surface and internal levels. The quality of the protein was checked through the PROCHECK server by dihedral analysis of the Ramachandran plots of predicted candidate proteins.

## 2.2 Annotating candidate genes at the functional level

Functional annotation is the process which includes the annotation of a particular biological function to biomolecules based on a computational approach by comparing the query data with the existing datasets selected in the database. This is one of the most prominent steps after structural annotation of biomolecules (Humann et al., 2019). Functional annotation was performed by analyzing the gene ontology analysis, expression profiling, and molecular simulation studies of the candidate proteins.

### 2.2.1 Gene ontology

Complete functional annotation of the 17 candidate genes was performed at both the nucleotide and protein levels. At the cellular, molecular, and biological levels, gene ontology at the nucleotide level of the candidate genes was performed using the Gene Ontology Resource (http://geneontology.org), Ensembl database (https://plants.ensembl.org/index.html), URGI (https://


FIGURE 1
Morphology of the flag leaf of DBW303 and WH147 at different stages of sample collection.

wheat-urgi.versailles.inra.fr/), and UniParc (https://www.uniprot.org/help/uniparc). For the gene ontology study at the structural level, the modeled protein structures were analyzed with the help of the automated ProFunc server (http://www.ebi.ac.uk/thornton-srv/databases/ProFunc/) for the proper depiction of functions at the cellular, biological, and biochemical levels.

### 2.2.2 Gene expression profiling

The expression profiling of all the candidate genes was performed with the help of GENEVESTIGATOR (https://genevestigator.com/) at different levels. First, at the anatomical level, depiction was performed on a quantitative basis in different tissues of wheat, taking into account the TA_mRNASeq_WHEAT_GL-0 dataset and selecting all 17 candidate genes in a heat map format. Second, at ten different developmental stages, a set of experiments was selected as the desired perturbation for the presence of 17 candidate genes in a heat map format.

### 2.2.3 Experimental validation of candidate genes through expression analysis

#### 2.2.3.1 Plant material and sampling stages

For experimental validation of candidate genes, two varieties DBW303 (high yielding) and WH147 (low yielding) were selected. The plants were grown at the research field of wheat and barley section, the Department of Genetics and Plant Breeding, Chaudhary Charan Singh Haryana Agricultural University, Hisar, India. The flag leaf samples of the varieties were collected at three different developmental stages: flag leaf emergence, fully developed flag leaf, and flag leaf senescence (Figure 1).

#### 2.2.3.2 Relative expression of candidate genes in high- and low-yielding wheat varieties

Total RNA was isolated from 50 to 100 mg flag leaf tissue of selected varieties using a Maxwell RSC Plant RNA Kit (Promega, United States) according to the manufacturer's recommendations. A RevertAid cDNA Synthesis Kit (Thermo Scientific, United States) was used for the synthesis of first-strand cDNA from the total RNA isolated. The primers of all the 17 genes along with actin as an endogenous control were designed with the help of Primer Express Software v3.0.1 (Applied Biosystems). A detailed list of the primers used is given in Supplementary Table S1. Quantitative real-time polymerase chain reaction was performed using a QuantaStudio™ 6 Flex Real-Time PCR Detection System (Applied Biosystems) with the PowerUp™ SYBR™ Green PCR Master Mix (Applied Biosystem) for 2 min at 50°C for UDG activation, 10 min at 95°C for initial denaturation followed by 40 cycles of 15 s of denaturation at 95°C, 15 s of annealing 53°C and extension at 72°C for 1 min). Fold changes were calculated by the $2^{-\Delta\Delta Ct}$ method.

### 2.2.4 Network analysis of candidate genes with native protein and chemical interactors

Protein–protein and protein–chemical interaction network analyses were performed with the help of servers such as GeneMANIA (https://genemania.org/) and the STITCH v 5.0 (http://stitch.embl.de/), respectively. Network analysis was performed to identify the biomolecules interacting with the candidate genes for proper functional annotation of genes encoding proteins.

### 2.2.5 Molecular dynamics simulation

To evaluate the stability of the predicted 3D structures, MD simulations were performed. Modeled 3D structures from the SWISS-MODEL server were used as the initial configuration of all simulations. The MD simulations were conducted as per the protocol previously described in Gajula et al. (2016); Kumar et al. (2016, 2018a, and 2018b); Mathpal et al. (2018); and Gautam et al. (2019). A minimum of 10 Å distance was used between the protein surfaces and the simulation box edges. All the systems

TABLE 1 List of identified flag leaf development–associated genes along with their basic characteristics elucidated with the help of the Ensembl Plants database.

| S No. | Gene | Ensembl ID | Location | Cellular location | CDS | AA |
|---|---|---|---|---|---|---|
| 1 | *TaAct1-4B* | TraesCS4B02G050600 | 4B | Cytoskeleton | 1982 | 385 |
| 2 | *TaBri1-3D* | TraesCS3D02G246500 | 3D | Endosome, membrane | 4,445 | 1,124 |
| 3 | *TaGATA12-3D* | TraesCS3D02G274600 | 3D | Nucleus | 1747 | 386 |
| 4 | *TaNAP1-7B* | TraesCS7B02G167200 | 7B | Cytoplasm | 4,074 | 1,357 |
| 5 | *TaNfl1-2B* | TraesCS2B02G464200 | 2B | Nucleus | 1,428 | 392 |
| 6 | *TaNOL-4D* | TraesCS4D02G012100 | 4D | Chloroplast thylakoid | 1,432 | 345 |
| 7 | *TaNyc1-3D* | TraesCS3D02G159800 | 3D | Plastid, integral part of membrane | 2,121 | 499 |
| 8 | *TaNyc3-7A* | TraesCS7A02G338900 | 7A | Chloroplast | 2,258 | 547 |
| 9 | *TaOsh1-4A* | TraesCS4A02G256700 | 4A | Nucleus | 1,616 | 362 |
| 10 | *TaOsl2-2B* | TraesCS2B02G440400 | 2B | Mitochondria | 1,536 | 511 |
| 11 | *TaPME1-1B* | TraesCS1B02G274500 | 1B | Cytoplasm | 1981 | 555 |
| 12 | *TaPNH1-7B* | TraesCS7B02G256500 | 7B | Cytoplasm | 3,410 | 956 |
| 13 | *TaRCCR1-7D* | TraesCS7D02G492300 | 7D | Chloroplast | 1,565 | 455 |
| 14 | *TaSCR-5B* | TraesCS5B02G143100 | 5B | Nucleus | 2,339 | 638 |
| 15 | *TaSGR-5D* | TraesCS5D02G325900 | 5D | Chloroplast mitochondria | 1798 | 280 |
| 16 | *TaSRT1-5D* | TraesCS5D02G066000 | 5D | Nucleus | 2,870 | 755 |
| 17 | *TaTSD2-6B* | TraesCS6B02G348700 | 6B | Golgi apparatus | 2,961 | 660 |

were solvated with the TIP3P water model. $Cl^-$ and $Na^+$ were further added to the simulation boxes to neutralize the systems. All unbiased MD simulations were carried out using GROMACS-2020 (https://www.gromacs.org) with an all-atom CHARMM36m force field. All the systems were energy minimized and equilibrated with different initial atomic velocities using the following steps: 1) minimization of energy to steepest descent level, 2) restrainment of position (all heavy atoms of protein) NVT [moles (N), volume (V), and temperature (T)] simulation with a restraining force constant of 1,000 kJ/mol, and 3) a 500 ps position-restrained NPT [moles (N), pressure (P), and temperature (T)] simulation with a restraining force constant of 1,000 kJ/mol. The NPT ensemble was utilized for the simulation production at 20 ns with an average temperature of 300 K *via* the v-rescale thermostat and Parrinello–Rahman barostat with 2.0 ps as the coupling constant. For the cutoff of Lenard-Jones and other short-range electrostatic interactions, a Verlet cutoff scheme was employed with a cutoff of 1.0 nm throughout the simulation. The particle mesh Ewald summation method was employed to treat electrostatic interactions over a long range. The LINCS algorithm was used for the constraint of all hydrogen atoms, and the SETTLE algorithm was used for constraining all the bonds and angles of TIP3P water molecules. The root means square deviation (RMSD) of Cα atoms of protein concerning a reference structure was calculated with the GROMACS program "gmxrms" by least-square fitting the structure to the reference structure. Principal component analysis (PCA) is a widely used technique to highlight the slowest functional motions of proteins (Bahar et al., 1998). The principal components (PCs) representing the collective motion of the proteins were obtained after diagonalization of the covariance matrix. The eigenvectors and eigenvalues of the covariance matrix represent the principal direction of the motion and the magnitude of motion along with it, respectively. The GROMACS tools "gmxcovar" and "gmxanaeig" were used for the calculation of the covariance matrix and PC projections.

## 2.3 Phylogenetic analysis

Phylogenetic analysis was conducted to obtain information regarding the relatedness among the 17 candidate proteins, and the phylogenetic tree was constructed with the help of Mega X (https://www.megasoftware.net/) software using the neighbor-end joining method with 1,000 bootstraps.

# 3 Results

## 3.1 Identification and chromosomal localization of candidate genes in the wheat genome associated with flag leaf development

The protein sequences of candidate genes associated with the development of flag leaf were retrieved from different plants such as *Oryza sativa* and *Arabidopsis thaliana,* and the orthologs of these proteins responsible for flag leaf development in wheat were identified by BLASTp. All 17 candidate genes that were identified from the Ensembl Plants database were annotated with the help of various bioinformatics tools. The BLASTp results

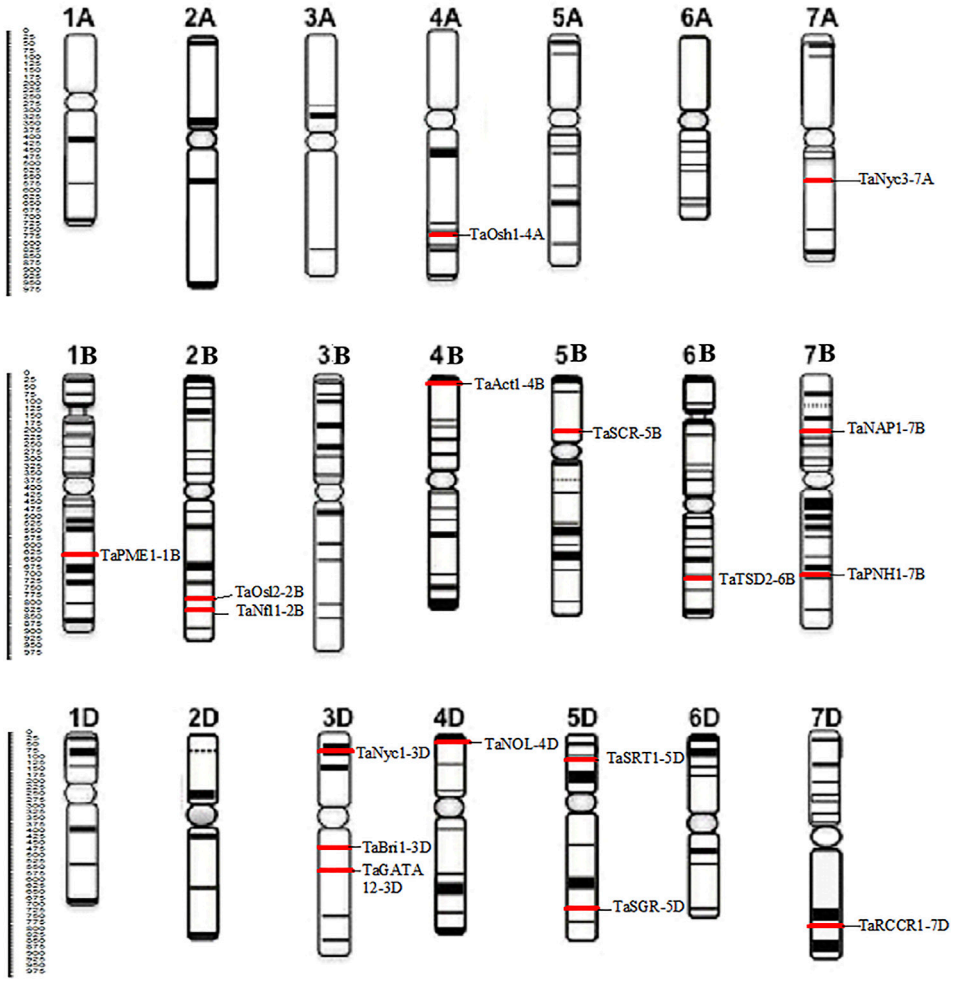



FIGURE 2
Distribution of 17 candidate genes associated with the development of the flag leaf on different chromosomes of wheat.

yielded more than 50 wheat scaffolds, of which the one with the highest sequence identity with the respective genes was taken as an ortholog for further analysis (Dhaliwal et al., 2014).With each gene queried, numerous associated transcripts were found. The full-length transcripts of candidate genes related to flag leaf genes were found to have 67.6–99.4% identity with the query sequence. The size of the CDS ranged from 1,428 to 4,445 bp, and the subsequent protein length ranged from 280 to 1,124 amino acids (Table 1). In total, 17 flag leaf-associated genes were found to be scattered on chromosomes 1B, 2B, 3D, 4A, 4B, 4D, 5B, 5D, 6B, 7A, 7B, and 7D of wheat (Figure 2).

## 3.2 Identification of gene structure

The Gene Structure Display Server (GSDS) v 2.0 was used to elucidate the gene structures of candidate genes associated with the development of the flag leaf. The gene structure comprises exons, introns, and 3′ and 5′ untranslated regions. The number of introns ranged from one in TaGATA12-3D to 22 in TaNAP1-7B (Figure 3). The upstream and downstream regions were also analyzed for their detailed localization (Hu et al., 2015). However, the common ancestral origin of different genes was confirmed, as the exon–intron composition of flag leaf development genes was not different from their homologs.

## 3.3 Physiochemical properties of candidate proteins

The physiochemical properties of candidate proteins associated with the development of the flag leaf were elucidated using the ProtParam server. The molecular weight of translated protein ranged from 31,188.61 g/mol (TaSGR-5D)

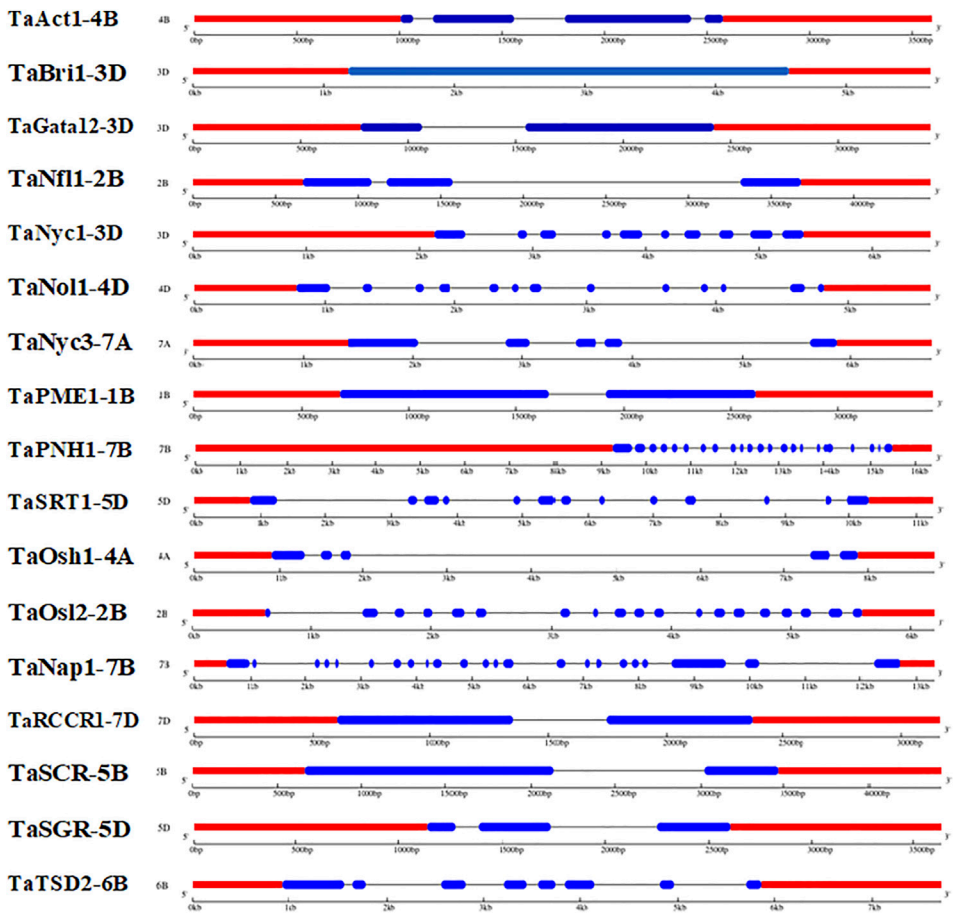



FIGURE 3
Distribution of exons and introns along with upstream and downstream regions in wheat candidate genes associated with flag leaf development (predicted by the GSDV v2.0 server).

to 151,621.01 g/mol (TaNAP1-7B), and the isoelectric point (pI) ranged from 5.37 (TaAct1-4B) to 9.6 (TaNOL-4D). Out of 17 proteins, five were found to be stable in nature, while 12 were unstable, with their instability index ranging from 40.68 to 67.91 during preliminary analysis. The higher aliphatic index of proteins ranging from 67.10 (TaGATA12-3D) to 97.54 (TaBri1-3D) suggests high stability of proteins over a wider range of temperatures. The GRAVY score ranged from −0.579 (TaOsh1-4A) to 0.009 (TaPME1-1B), indicating the hydrophilic nature of the proteins (Table 2).

## 3.4 Homology modeling

The modeling was performed based on a homology-based approach through the SWISS-MODEL server (Schwede et al., 2003; Kumar et al., 2018b; Waterhouse et al., 2018). Many templates were found ranging from 3 to 50, concerning PSI-BLAST and HHblits having a % identity ranging from 17.07 to 90.67. The state of proteins was from monomer to homodimer and homotetramer. Qualitative Model Energy Analysis (QMEAN) values of predicted proteins ranged from −0.28 in TaRCCR1-7D to −5.24 in TaNyc3-7A (Table 3). The derived structures were analyzed further with the help of UCSF ChimeraX1.1 software (Pettersen et al., 2021) for the further depiction of secondary structures such as coils, helices, sheets, and surface features (Figure 4).

The topological architecture of proteins was predicted through the ProFunc server, which elucidates the detailed characteristics of the modeled protein structure and the ligands attached to it. Ligands such as 4′-deoxy-4′-aminopyridoxal-5′phosphate (PMP), cacodylate (CAC) ions, and magnesium (Mg) ions were found to be associated with TaOsl2-2B, TaPME1-1B, and TaSRT1-5D, respectively (Figure 5).

## 3.5 Dihedral analysis

Ramachandran dihedral statistics for modeled 3D structures of proteins associated with the development of the flag leaf in

TABLE 2 Physiochemical properties of candidate proteins associated with the development of flag leaf through the ProtParam server.

| S No. | Protein | Molecular weight | Theoretical isoelectric point | Instability index | Aliphatic index | GRAVY | Stability |
|---|---|---|---|---|---|---|---|
| 1 | TaAct1-4B | 42,740.95 | 5.37 | 35.52 | 85.87 | −0.151 | Stable |
| 2 | TaBri1-3D | 120,078.06 | 5.8 | 33.29 | 97.54 | −0.014 | Stable |
| 3 | TaGATA12-3D | 40,961.89 | 5.96 | 67.28 | 67.1 | −0.445 | Unstable |
| 4 | TaNAP1-7B | 151,621.01 | 6.35 | 48.48 | 94.48 | −0.147 | Unstable |
| 5 | TaNfl1-2B | 43,014.87 | 8.96 | 49.08 | 71.17 | −0.493 | Unstable |
| 6 | TaNOL-4D | 37,714.23 | 9.6 | 42.73 | 80.09 | −0.26 | Unstable |
| 7 | TaNyc1-3D | 41,890.06 | 8.12 | 37.96 | 89.87 | 0.005 | Stable |
| 8 | TaNyc3-7A | 42,721.67 | 6.02 | 54.32 | 82.27 | −0.212 | Unstable |
| 9 | TaOsh1-4A | 40,079.08 | 6.28 | 55.69 | 74.75 | −0.579 | Unstable |
| 10 | TaOsl2-2B | 48,653.93 | 6.81 | 42.22 | 86.15 | −0.066 | Unstable |
| 11 | TaPME1-1B | 58,739.1 | 7.53 | 34.11 | 79.17 | 0.009 | Stable |
| 12 | TaPNH1-7B | 107,038.12 | 9.32 | 47.07 | 81.5 | −0.387 | Unstable |
| 13 | TaRCCR1-7D | 35,590.81 | 6.41 | 46.7 | 94.68 | 0 | Unstable |
| 14 | TaSCR-5B | 44,271.15 | 7.82 | 67.91 | 82.17 | −0.32 | Unstable |
| 15 | TaSGR-5D | 31,188.61 | 8.69 | 52.31 | 81.61 | −0.296 | Unstable |
| 16 | TaSRT1-5D | 82,979.01 | 5.62 | 39.97 | 84.82 | −0.131 | Stable |
| 17 | TaTSD2-6B | 71,552.22 | 6.61 | 40.68 | 75.09 | −0.372 | Unstable |

TABLE 3 Enumeration of dihedral properties of proteins after dihedral analysis through SWISS-MODEL server.

| S No. | Protein | MolProbity score | QMEAN | Stability | Solvation | Torsion | R favor (%) | State |
|---|---|---|---|---|---|---|---|---|
| 1 | TaAct1-4B | 0.77 | −0.48 | Stable | 0.01 | −0.41 | 97.88 | Monomer |
| 2 | TaBri1-3D | 1.41 | −1.97 | Stable | −2.25 | −1.35 | 92.05 | Monomer |
| 3 | TaGATA12-3D | 1.74 | −2.20 | Stable | 0.03 | −2.82 | 86.11 | Monomer |
| 4 | TaNfl1-2B | 0.87 | −1.00 | Stable | 0.60 | −1.21 | 97.52 | Monomer |
| 5 | TaNOL-4D | 2.66 | −4.10 | Unstable | −0.74 | −3.31 | 89.24 | Homotetramer |
| 6 | TaNyc1-3D | 1.53 | −3.14 | Stable | −0.84 | −2.61 | 91.88 | Homotetramer |
| 7 | TaNyc3-7A | 2.14 | −5.24 | Unstable | −0.41 | −4.84 | 89.47 | Monomer |
| 8 | TaPME1-1B | 1.43 | −2.88 | Stable | −1.32 | −2.10 | 93.71 | Monomer |
| 9 | TaPNH1-7B | 1.69 | −2.02 | Stable | 0.08 | −1.67 | 92.64 | Monomer |
| 10 | TaSRT1-5D | 1.79 | −4.55 | Unstable | −0.17 | −4.11 | 89.71 | Monomer |
| 11 | TaOsh1-4A | 1.28 | −0.90 | Stable | 0.61 | −1.29 | 96.67 | Monomer |
| 12 | TaOsl2-2B | 1.76 | −1.80 | Stable | 0.71 | −1.87 | 93.79 | Homodimer |
| 13 | TaNAP1-7B | 1.67 | −4.43 | Unstable | 0.26 | −3.98 | 91.18 | Monomer |
| 14 | TaRCCR1-7D | 1.91 | −0.28 | Stable | 0.68 | −0.29 | 94.88 | Homodimer |
| 15 | TaSCR-5B | 1.05 | −0.48 | Stable | 0.64 | −0.87 | 98.04 | Monomer |
| 16 | TaSGR-5D | 1.45 | −3.01 | Stable | −3.47 | −1.74 | 89.74 | Monomer |
| 17 | TaTSD2-6B | 2.41 | −3.28 | Stable | −2.04 | −2.30 | 94.7 | Monomer |

wheat deciphered with MolProbity scores ranging from 0.77 (TaAct1-4B) to 2.66 (TaNOL-4D). TaAct1-4B, TaBri1-3D, TaGATA12-3D, TaNfl-2B, TaNyc1-3D, TaPME1-1B, TaPNH1-7B, TaOsh1-4A, TaOsl2-2B, TaRCCR1-7D, TaSCR-5, TaSGR-5D, and TaTSD2-6B were all approximately zero, indicating a good agreement between the modeled structure and an experimental structure of similar size (Table 3). The PROCHECK server analysis of the modeled protein revealed a varied percentage of residues under the most favored, generously allowed, additionally allowed, and disallowed regions (Laskowski et al., 1993). The G-score that provides a measure of how normal a structure works with different proteins ranged from −0.32

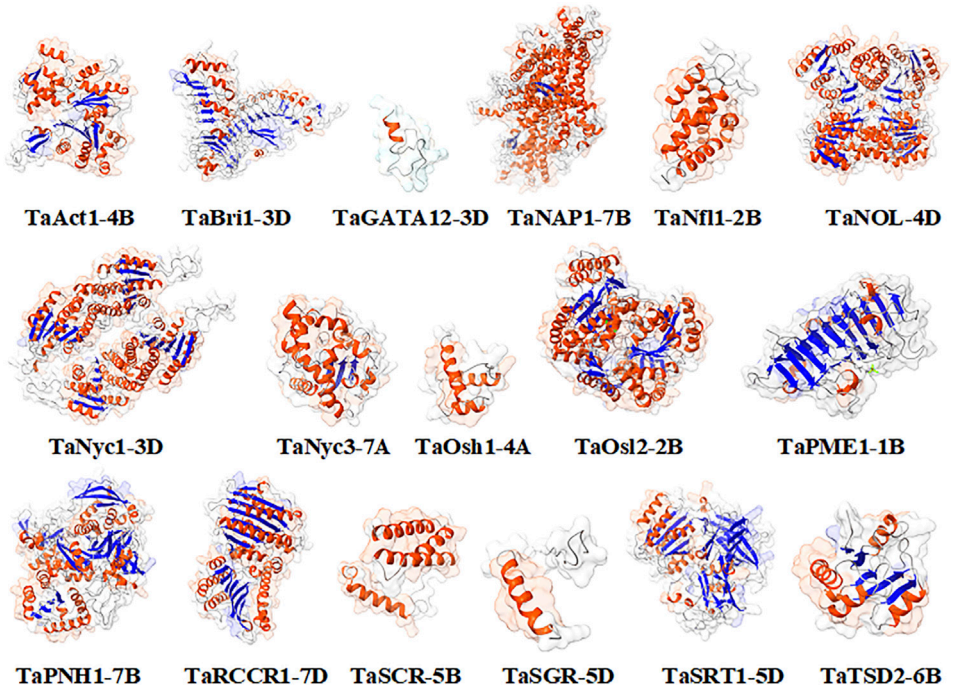



FIGURE 4
Homology-based models (molecular graphical images) produced using UCSF Chimera.

(TaNyc3-7B) to 0.07 (TaNfl1-2B), indicating that the predicted models were of excellent geometry and were accepted for further analysis (Table 4; Supplementary Figure S1).

## 3.6 Functional enrichment analysis

Gene ontology enrichment analysis of genes associated with flag leaf development revealed their association with various cellular and metabolic processes. It was found that the candidate genes were involved in several biological and molecular functions, including population maintenance, leaf development, protein phosphorylation, regulation of transcription of various catabolic and anabolic processes that are required for the proper functioning of photosynthesis and nutrient transfer from flag leaf to developing grains, cellular component organization and localization, cell wall modification, and macromolecular metabolism.

## 3.7 Functional elucidation based on protein structure

The ProFunc server (Laskowski et al., 2005) was used to explore the functional annotation of genes based on protein structures derived from a homology modeling approach. It revealed the association of these modeled proteins with various cellular, biological, and biochemical processes. Cellular processes include cellular organization and localization of the cytoskeleton, periplasmic spaces, membrane, cytoplasm, and intracellular and extracellular regions. Biological processes include metabolic processes such as phosphate and pectin catabolism and carbohydrate and organic acid metabolism. Biochemical processes include the activities ofnucleotide, ATP, metal ion, and protein binding, and catalytic activities include the activities of hydrolase, oxidoreductase, nuclease, carboxylesterase, aspartyl esterase, and transaminase.

## 3.8 Profiling of gene expression

A microarray TA_mRNASeq_WHEAT_GL-0 dataset of all 17 genes was found to be available on the GENEVESTIGATOR (Grennan, 2006) platform, which was further utilized for gene expression analysis at different levels.

### 3.8.1 Differential expression in tissues

Expression profiling of all 17 genes revealed their presence in 44 different tissues in wheat, which were analyzed, and it was found that the expression of these genes was higher in the root tip and radicle, shoot apical meristem, seedling, and in the zone undergoing active growth, suggesting that the genes are involved

10.3389/fgene.2022.933560

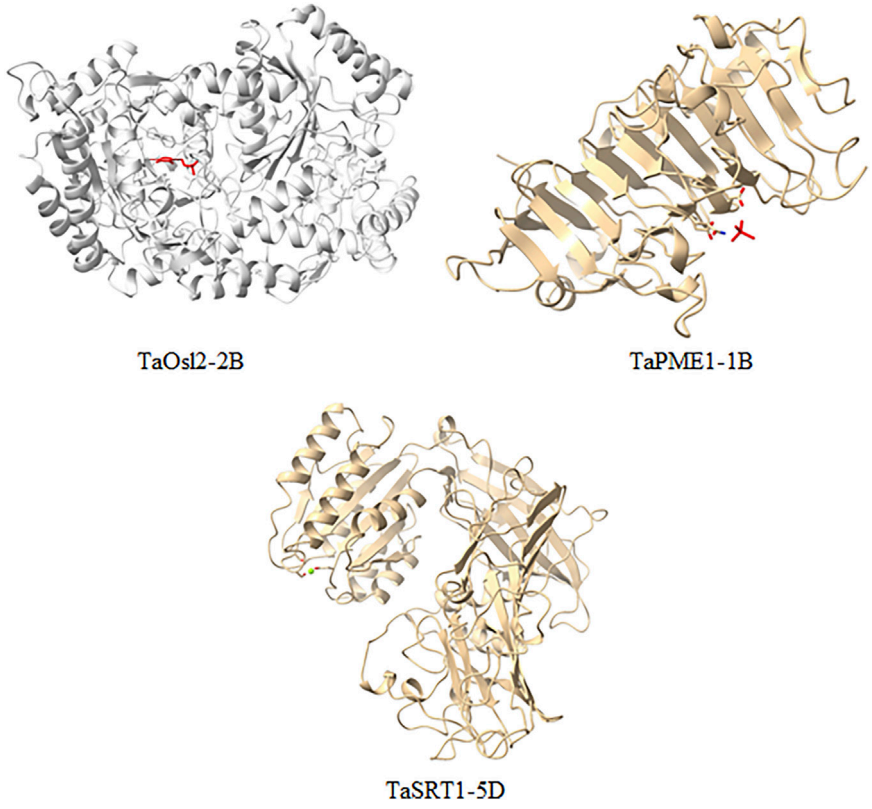



FIGURE 5
Localization of ligands interacting with the modeled protein structure through ChimeraX.

TABLE 4 Catalog of dihedral properties of candidate proteins elucidated through Ramachandran plot analysis.

| S No. | Protein | Most favored region | Generously allowed region | Additionally allowed region | Disallowed region | G score |
|---|---|---|---|---|---|---|
| 1 | TaAct1-4B | 94.2 | 0.0 | 5.80 | 0.0 | −0.07 |
| 2 | TaBri1-3D | 77.2 | 0.7 | 22.1 | 0.0 | −0.22 |
| 3 | TaGATA12-3D | 75.0 | 3.6 | 21.4 | 0.0 | −0.10 |
| 4 | TaNAP1-7B | 89.1 | 1.1 | 9.10 | 0.7 | −0.15 |
| 5 | TaNfl1-2B | 95.2 | 1.4 | 3.40 | 0.0 | 0.07 |
| 6 | TaNOL-4D | 89.2 | 1.6 | 9.00 | 0.6 | −0.20 |
| 7 | TaNyc1-3D | 88.1 | 0.7 | 10.6 | 0.6 | −0.13 |
| 8 | TaNyc3-7A | 86.1 | 0.5 | 12.5 | 1.0 | −0.32 |
| 9 | TaOsh1-4A | 96.3 | 0.0 | 3.70 | 0.0 | −0.07 |
| 10 | TaOsl2-2B | 87.8 | 0.8 | 10.8 | 0.6 | −0.12 |
| 11 | TaPME1-1B | 87.1 | 0.4 | 12.5 | 0.0 | −0.16 |
| 12 | TaPNH1-7B | 88.9 | 1.4 | 9.40 | 0.3 | −0.09 |
| 13 | TaRCCR1-7D | 90.2 | 0.0 | 9.60 | 0.2 | −0.14 |
| 14 | TaSCR-5B | 95.6 | 0.0 | 4.40 | 0.0 | 0.02 |
| 15 | TaSGR-5D | 82.4 | 0.0 | 14.7 | 2.9 | −0.15 |
| 16 | TaSRT1-5D | 86.5 | 1.9 | 11.0 | 0.6 | −0.23 |
| 17 | TaTSD2-6B | 88.8 | 1.7 | 8.60 | 0.9 | −0.28 |

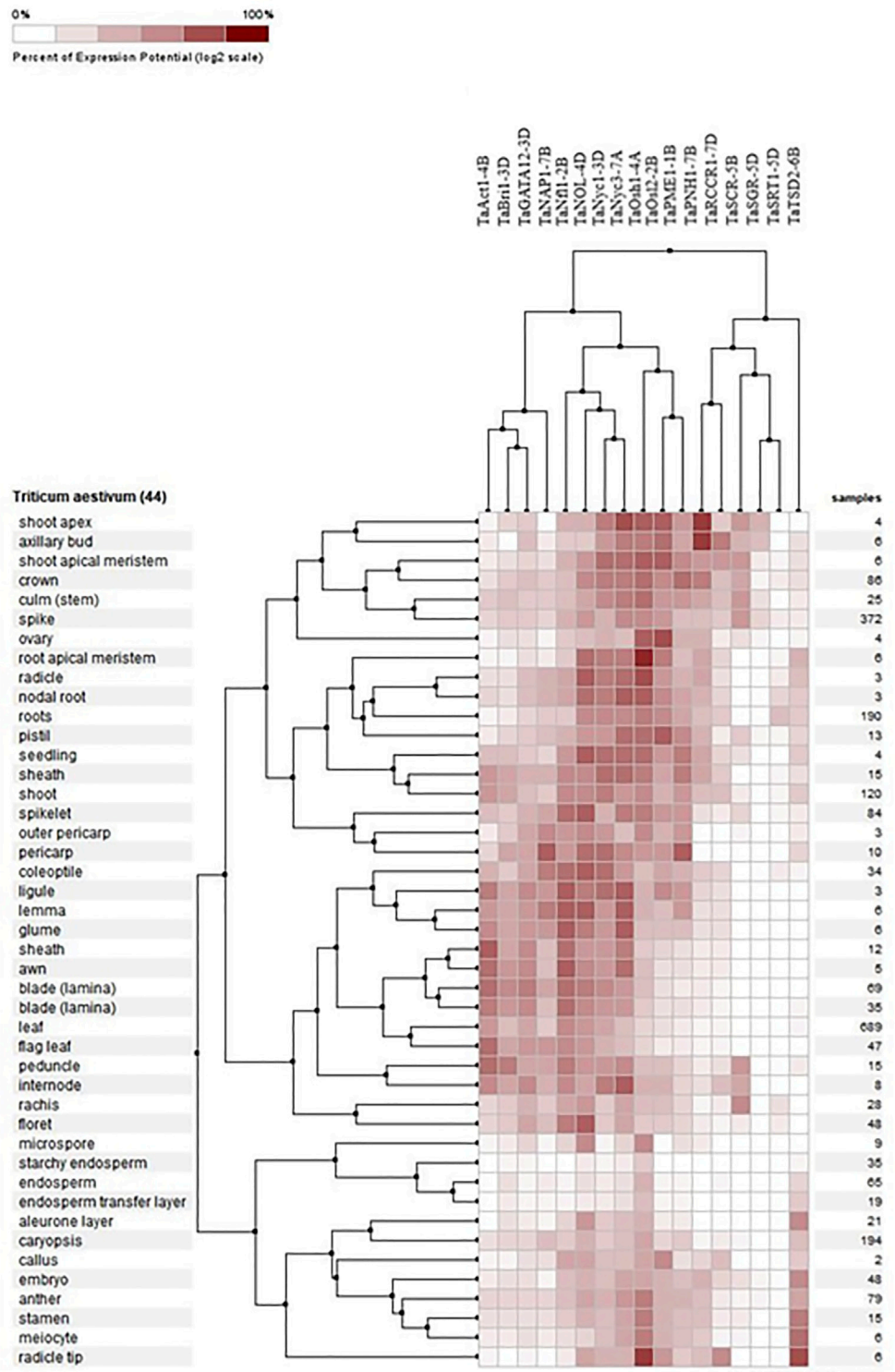



FIGURE 6
Hierarchical clustering of wheat genes based on their expression in 44 different tissues.

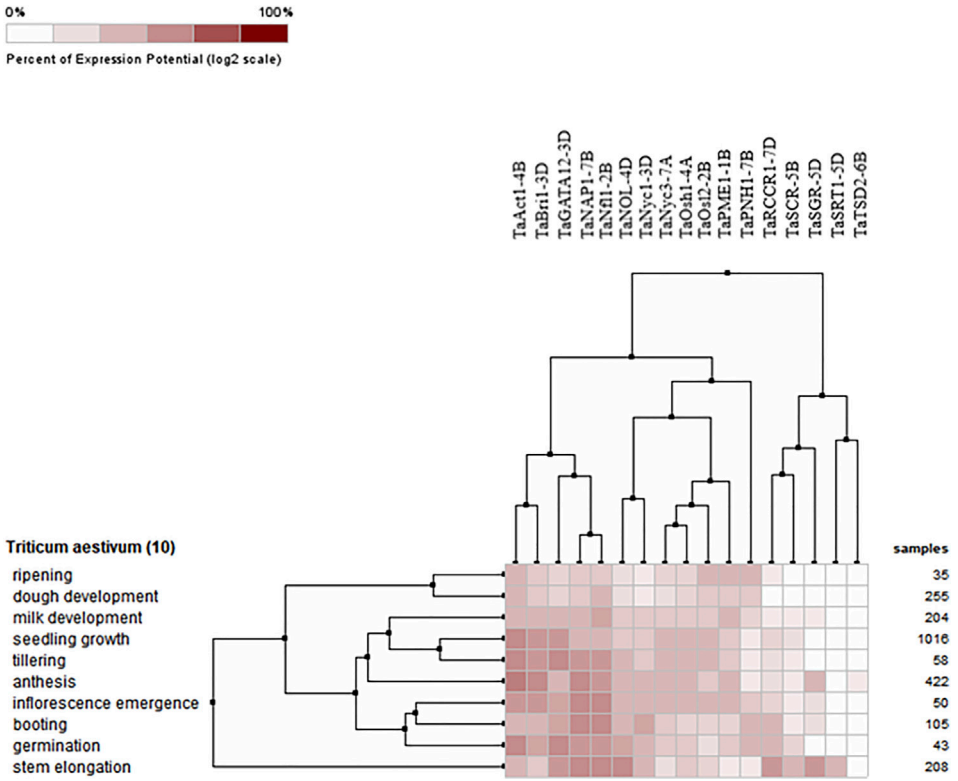



FIGURE 7
Hierarchical clustering of wheat genes based on their expression in ten different development stages.

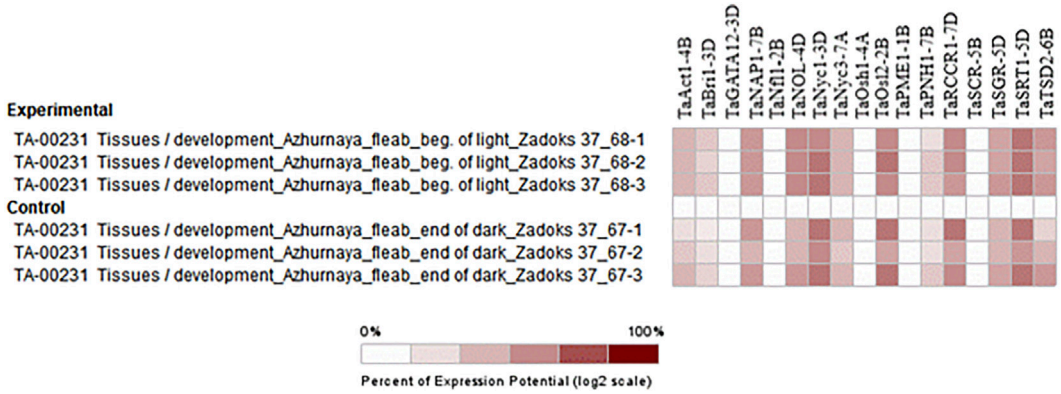



FIGURE 8
Hierarchical clustering of wheat genes based on their expression under the emergence stage in flag leaf.

in the development of tissue at the seedling stage and at later stages. TaAct1-4B was found mostly in the growing parts of the plants, especially in the flag leaf sheath and internodal areas, and it was found to be the most active gene in the flag leaf compared to the rest of the candidate genes. TaBri1-3D and TaRCCR1-7D were prominently present in the shoot apex and axillary buds (Figure 6).

### 3.8.2 Expression of candidate genes at different developmental stages

The expression profiles of wheat genes associated with the development of flag leaf were analyzed at ten different developmental stages, including milk development, seedling growth, tillering, anthesis, inflorescence emergence, booting,

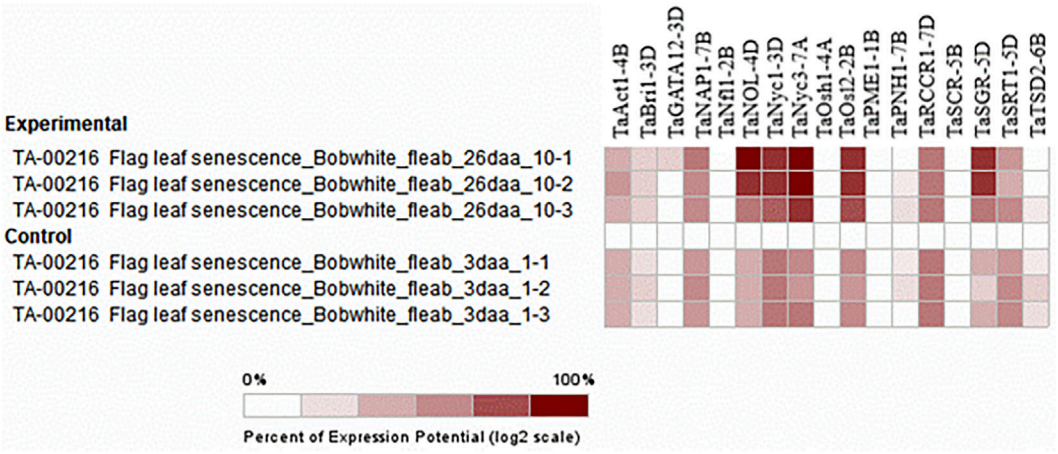



FIGURE 9
Hierarchical clustering of wheat genes based on their expression under senescence perturbations.

germination ripening, dough development, and stem elongation. All the candidate genes either upregulated or downregulated were found to be expressed at all developmental stages. However, the expression of these genes was found to be highly prominent during germination and seed growth, deciphering their role in development; tillering, booting, and anthesis, depicting their role in the development of flag leaf, and functioning as a source–sink pathway at the time of anthesis, when the stored energy in flag leaf starts accumulating in grains (Figure 7).

## 3.8.3 Expression during flag leaf development and senescence

### 3.8.3.1 Flag leaf development

Absolute expression analysis was performed by selecting perturbations where the flag leaf blade was harvested at the beginning of the light period (15 min after lights went on) from Azhurunaya plants grown to Zadoks 37 (flag leaf just visible) in a growth chamber under 16 h light/8 h dark cycles compared to 15 min before lights went on. The higher expression of genes such as *TaNAP1-7B*, *TaNOL-4D*, *TaNyc1-3D*, *TaOsl2-2B*, *TaSRT1-5D*, and *TaTSD2-6B* confirms their involvement in flag leaf development in wheat through processes such as cellular organization, leaf development, and photosynthesis (Figure 8).

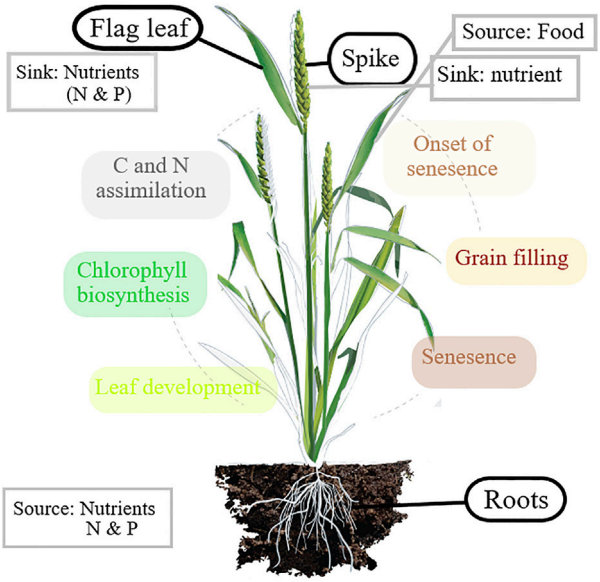



FIGURE 10
Developmental stages of flag leaf in wheat.

### 3.8.3.2 Flag leaf senescence

Selective expression analysis was performed by selecting particular perturbations involving the sampling of a 3-cm long section from the middle of the flag leaf blade 26 days after anthesis (dough development) from the main tiller of Bobwhite plants grown under 16 h light at 20°C/8 h dark at 15°C cycles in 1-L pots filled with Peters Field Cereal Mix, and 3 days after anthesis, sampling was taken as control (milk development). The chlorophyll content was ten units higher in the control sample than in the treated sample, which was 26 days after anthesis (Warburton et al., 2002). Similarly, the chlorophyll content in the flag leaf was found to increase from the time of heading to the seed setting stage and declined thereafter in cereal crops (Liu et al., 2008; Derkx et al., 2012). The biosynthesis and degradation of chlorophyll are catalyzed by a unique set of enzymes (Reinbothe et al., 2010; Hörtensteiner and

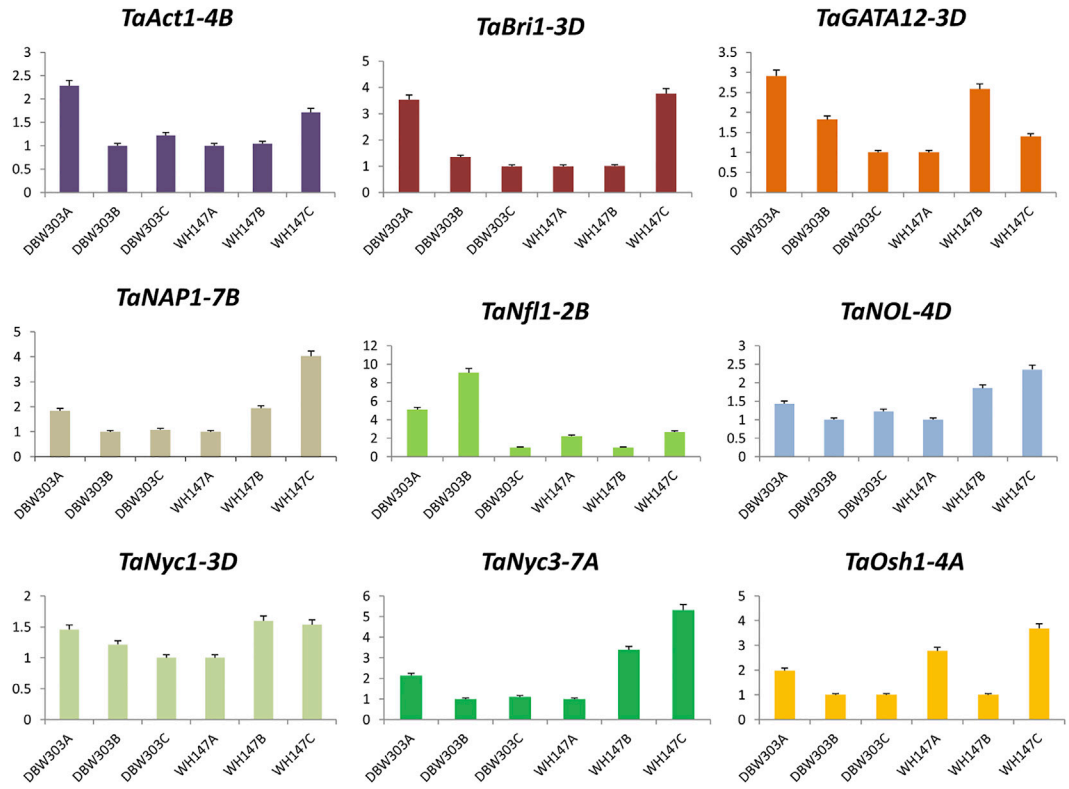



FIGURE 11
Graphical representation of relative expression levels of nine candidate genes in two wheat varieties namely DBW303 and WH147 at three different developmental stages namely flag leaf emergence (DBW303A, WH147A); fully developed flag leaf (DBW303B, WH147B) and flag leaf at the time of senescence (DBW303C, WH147C).

Kräutler, 2011). The expression of genes associated with the degradation of chlorophyll is the first molecular indication of the onset of senescence (Hörtensteiner and Kräutler, 2011). As previously stated, the function of these genes was in metabolism, specifically the catabolism of the photosynthetic apparatus and chlorophyll, so the function of these genes can be inferred perfectly from the experiment (Figure 9).

At the molecular level, during the development of the flag leaf, there are different phases controlled by different sets of genes. Phase 1 starts with the increase in the expression of genes involved in the development of the leaf. Phase 2 encounters an increase in the expression of genes associated with the biosynthesis of chlorophyll and other leaf functions. Phase 3 is the most active phase, involving the assimilation of carbon and nitrogen in the leaves, as mature leaves serve as a sink to store nitrogen before the anthesis phase; phase 4 begins with the onset of senescence and is characterized by the decline in leaf chlorophyll content; and upon anthesis, these leaves serve as a source of nitrogen to support the process of grain filling and utilization. Phase 5 involves the remobilization of nutrients from senescing flag leaf to the grain and other developing parts, which will ultimately lead to the complete senescence of the flag leaf. Leaf senescence is known to be an active process until death, with the main functions of recycling and reusing nutrients for the newly developing organs and enhancing the chances of survival of plants under abiotic stress (Figure 10).

## 3.9 The expression profile of candidate genes using RT-PCR

To assess the reliability and validity of *in silico* expression data and to obtain comprehensive insight into the expression profile of candidate genes in wheat in relation to high-yielding and low-yielding varieties, quantitative real-time PCR was performed using gene-specific primers for all 17 candidate genes. The fold change of the genes was analyzed relative to the other two stages of the same variety only.

The expression of *TaAct-4B* was highest in the flag leaf emergence stage of the high-yielding wheat variety and was similar in *TaBri-3D*, *TaNfl1-2B*, *TaSGR-5D*, *TaSRT1-5D*, and *TaTSD2-6B genes*, indicating their constructive roles in the growth of flag leaf. The fold expression of genes such as *TaNyc1-3D*, *TaNyc3-7A*, *TaSCR-5B*, and *TaOsl2-2B* was

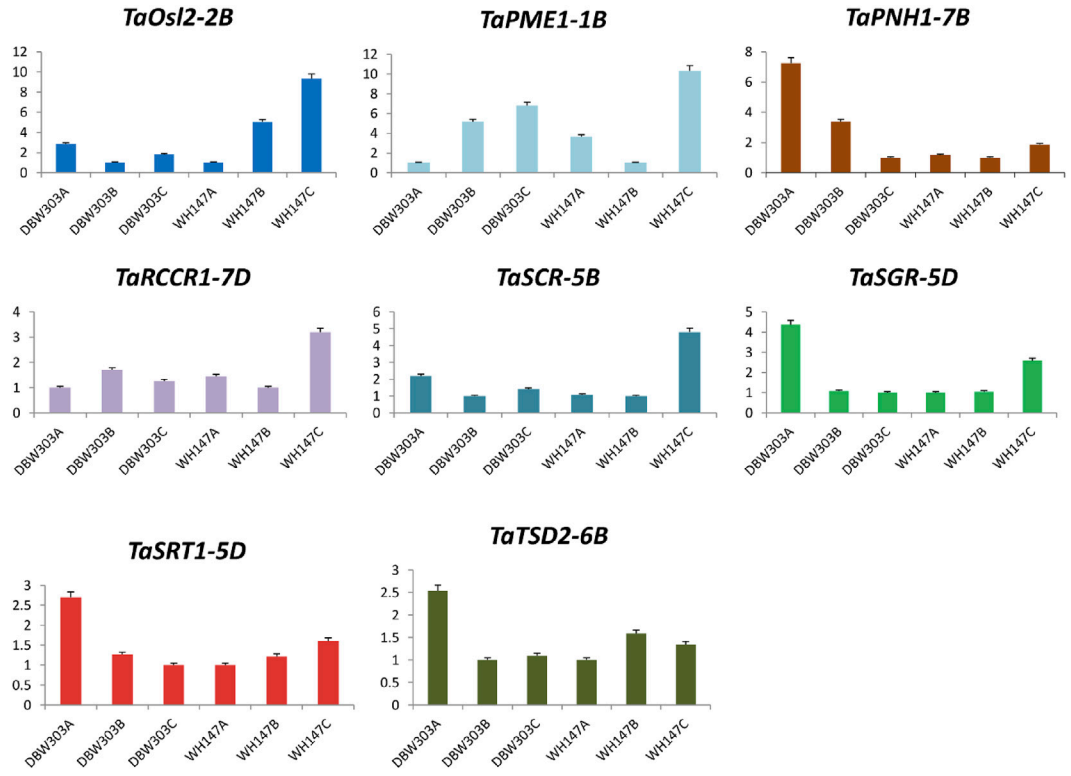



FIGURE 12
Graphical representation of relative expression levels of 8 candidate genes in two wheat varieties namely DBW303 and WH147 at three different developmental stages namely flag leaf emergence (DBW303A, WH147A); fully developed flag leaf (DBW303B, WH147B) and flag leaf at the time of senescence (DBW303C, WH147C) in the manuscript.

more prominent during the senescence of flag leaf, revealing their important roles in senescence-related processes. Not just this expression level of these senescence-related genes is much higher in low-yielding wheat variety, that is, WH147 than the high-yield wheat variety DBW303 (Figures 11, 12).

## 3.10 Complex regulatory networks of flag leaf development and associated proteins

### 3.10.1 Network of protein–protein interactions

Due to the unavailability of data for the wheat interactome in GeneMANIA, an interactome study was performed with reference to Arabidopsis. These putative interactor proteins were predicted at the *in silico* level to have regulatory-associated proteins based on physical interactions, shared functional domains, and co-expression. Of 17 proteins, interactome data for nine proteins were available in Arabidopsis. GeneMANIA (Franz et al., 2018) interactome analysis revealed interactions with the 20 most closely related proteins for each gene, allowing the elucidation of different functions based on these interactions.

TaBri1-3D has been discovered to interact with BAK1, CPI1, BSK1, BIK1, CRT3, SERK4, and 14 other proteins. The green lines show their interaction at the genetic level. The interactome analysis of TaBri1-3D revealed that this protein is involved in steroid metabolism, specifically the brassinosteroid-mediated signaling pathway, which is responsible for the negative regulation of cell death in plants. TaGATA12-3D was found to interact with 20 proteins of the same GATA family, including GATA18, GATA4, GATA5, and GATA2, which are basically involved in the circadian rhythm of plants. TaNyc1-3D was found to interact with NOL, HCAR, SGR1, SGRL, RCCR, and 15 other genes that were found to be involved in chlorophyll metabolic processes and nitrogen-containing compound catabolic processes. TaSRT1-5D was found to interact with SRT2, ETFA, HDA15, NUP62, GLDP1, PHB, DHS, ALS, HACL, PDC, CCA, and nine more genes. These were found to be involved in carboxy-lyase activity, the oxidoreductase complex, mitochondrial respiratory chain complex I, the NADH dehydrogenase complex, and the negative regulation of nitrogen compound metabolic processes. TaOsh1-4A was found to interact with a group of 20 genes, such as GILT, SHM, RHM, ETFB, and ETFQO involved in functions like

hydrolase activity, nucleotide biosynthetic processes, and oxidoreductase activity acting on the CH-OH group of donors, NNAD, or NADP as acceptors. TaSGR-5D was found to interact with proteins such as SGRL, NOL, PPH, RCCR, HO1, NYC1, and HCAR, which were found to be involved in chlorophyll metabolic processes, mainly catabolic ones. TaNAP1-7B was found to interact with 20 proteins, including BRK1, FLP, ABIL1, MYB88, PNM1, and HDT2. These proteins were found to be involved in the positive regulation of protein polymerization, cellular component morphogenesis, and cellular component organization. TaSCR-5B was found to be involved in asymmetric cell division by interacting with proteins such as SCL, GASA, PER32, SHR, and different SCL elements. TaRCCR1-7D was found to interact with 20 proteins, including ALB, NYC, RVE, SGRL, NOL, PPH, and PAO. These were found to be involved in chlorophyll metabolic processes and cellular nitrogen compound catabolic processes (Supplementary Figure S2).

These proteins were selected together for interactome analysis, and two different clusters were found, one with Nyc1, NOL, SGR, and RCCR, which are involved in chlorophyll metabolic processes, and the second cluster includes NAP, Bri1, and SCR, which are involved in cellular organization and brassinosteroid-mediated cell signaling.

### 3.10.2 Networks of chemical–protein interactions

The chemical–protein interaction analysis was performed with the help of the STITCH v5.0 server against *Hordeum vulgare* (Kuhn et al., 2007). The protein sequence of TaBri1-3D was found to interact with manganese and MgATP; TaGATA12-3D with MgATP, nitrate, ammonia, nitrite, and cycGMP; TaNOL-4D with nicotinamine, diphosphate, and reduced nitric acid; TaNyc1-3D with nicotinamine, reduced nitric acid, and TPNH; TaNyc3-7A with red chlorophyll, benzoic acid, sphinganine, and 7-keto-8-amine; TaPME1-1B with methanol, pectate, and distilled water; TaRCCR1-7D with red chlorophyll, hydrogen, sodium, TPNH, nicotinamine e and p, and SeMet; TaSCR-5B with R-rolipram and 1,2 dibromo 1; TaSGR-5D with CDCs, DMFE, and ketoglutarate; TaSRT1-5D with magnesium, vitamin B, carbamoyl phosphate, citrulline, ketoglutarate, hydrogen, ammonia, and phosphate; and TaTSD2-6B with mifepristone, trichlorobiphenyl, and Pentachlorobiphenyl (Supplementary Figure S3).

After the analysis of interactions, it was quite evident that most of the proteins were found to be associated with various N-containing compounds and enzymes related to nitrogen metabolism. Therefore, the importance of nitrogen and chlorophyll metabolism in flag leaf development was quite evident from the predicted results.

Nitrogen-use efficiency depends on nitrogen uptake, assimilation, and remobilization. In cereal crops such as wheat, mature leaves work as a sink to store nitrogen before the stage of anthesis, and upon anthesis, these leaves serve as the source of nitrogen to support the process of grain filling. Plants absorb N in their inorganic forms such as nitrate and ammonia, most of which are assimilated into organic forms. Mature tissue stores all organic and unassimilated inorganic nitrogen, either directly or indirectly utilized by the expanding tissue.

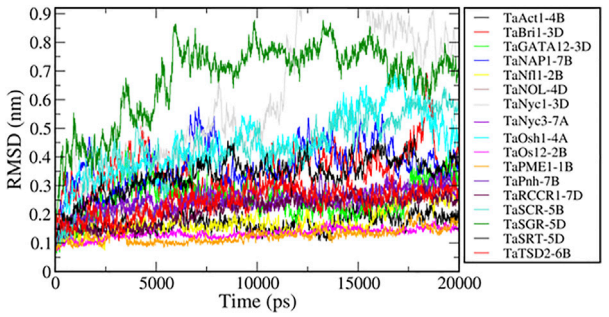



FIGURE 13
MD simulations of genes associated with flag leaf development.

## 3.11 Molecular dynamics simulation of predicted proteins

Molecular dynamics simulations have been extensively used to explore the conformational behavior of proteins. Here, we performed 20-nanosecond MD simulation studies of seventeen modeled flag leaf genes to understand their structural behavior. To investigate the equilibration and protein stability during the simulations of these proteins, the Cα RMSDs were calculated and monitored over the course of 20-nanosecond simulations. The assessment of the structural change was carried out by the analysis of the RMSDs from the starting structures as a function of simulation time. The resulting RMSD plots, presenting the conformational changes during simulation, showed that all the proteins (except two) achieved equilibrium at ~5 ns and remained stable for a period of 20 ns (Figure 13). These analyses also suggest that there are no large structural changes observed during simulation.

The RMSD values of two proteins, TaNyc1-D and TaSGR-5D, show higher values (>0.6 nm), indicating that these proteins have large conformational changes during the simulation. We used principal component analysis to better understand conformational changes during the simulation. PCA is a widely used method to reveal concerted motions (fluctuations) with large amplitudes (structural variations) from a set of configurations (Hess et al., 2008). All the simulated trajectories were projected along with the first principal component (PC1) which represents the largest structural variations. The motions along PC1 for all the trajectories were rendered (Supplementary Figure S4). Motions along PC1 showed

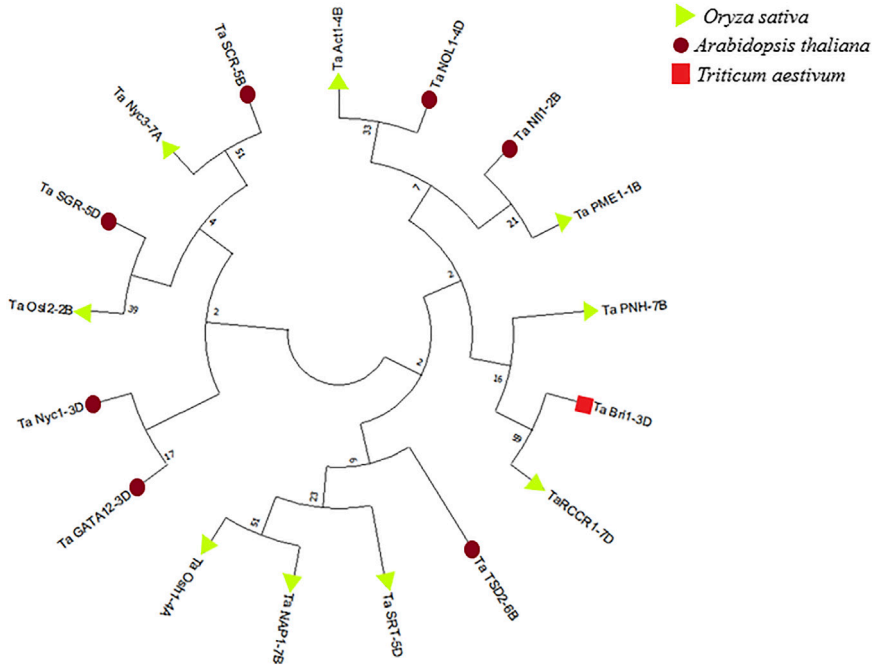



FIGURE 14
Phylogenetic tree of proteins associated with the development of flag leaf constructed through MegaX software.

that the large conformational changes in TaNyc1-D were due to the relative movement of different subunits, and the intra subunit conformational changes were not significant in TaSGR-5D due to the movement of flexible loops present at the terminal regions.

## 3.12 Phylogenetic analysis

Phylogenetic analysis was performed by using sequences of predicted proteins associated with flag leaf development through Mega-X software (Kumar S. et al., 2018). A phylogenetic tree was constructed using the neighbor-joining method with 1,000 bootstraps for inferring evolutionary relationships. Two distinct clusters were obtained, one with the genes involved in developmental processes and the other with the genes involved in degradation processes such as chlorophyll degradation and cell wall degradation. As inferred from Figure 14, the genes involved in the developmental processes made up a larger cluster than those in the other cluster.

# 4 Discussion

All 17 candidate genes associated with flag leaf development in wheat were characterized with the help of various bioinformatics tools. Phylogenetic analysis was the most prominent basis for the classification of these genes, under two distinct clusters, one with most of the genes associated with development and the other with genes associated with the regulation of senescence in flag leaf.

The first cluster included development-associated genes such as *TaAct1-4B*, *TaNfl1-2B*, *TaPME1-1B*, *TaPNH1-7B*, *TaBri1-3D*, *TaRCCR1-7D*, *TaTSD2-6B*, *TaSRT-5D*, and *TaOsh1-4A*, except *TaNAP1-7B* and *TaNOL1-4D*, whereas the second cluster included senescence-associated genes such as *TaSCR-5B*, *TaNyc3-7A*, *TaSGR-5D*, *TaOsl2-2B*, *TaNyc1-3D*, and *TaGATA12-3D*. Overall, starting from the functional enrichment analysis followed by the expression analysis, interaction studies, and structure-based functional analysis, this classification of genes fitted the best in understanding the whole picture of flag leaf development to a larger extent. The role of these genes as individuals or in collaboration with other genes was too complex to distinctively group them; hence, the genes are discussed in accordance with phylogenetic clusters.

TaAct1-4B, a member of the NBD-sugar kinase-Hsp70_act superfamily, plays an important role in cytoskeletal organization, localization, and transport of photoassimilates. Gui et al. (2015) reported that plasmodesmata conductance for the transport of photoassimilates in rice is regulated by the interaction of grain setting defect 1 with *OsACT1* (Gui et al., 2015). Moreover, it was the most prominent gene that was found to be expressed in the flag leaf of wheat, as was deciphered through *in silico* gene expression analysis and was the most prominent during the flag leaf emergence stage in the high-yielding wheat variety.

Nfl in tobacco was found to specify determinacy for both flowers and leaves in progenitor cells as Nfl was found to be expressed in vegetative tissues (Hofer et al., 1997). TaNfl1-2B was found to be a member of the C-LFY-FLO superfamily, expressed predominantly during stages such as inflorescence emergence and germination as studied through gene ontology and expression analysis. The highest amount of TaNfl1-2B was found during the developmental stages of flag leaf in the high-yielding wheat variety. Similar results were revealed by Ahearn et al. (2001) signifying the critical role of Nfl1 in the allocation of meristematic cells that are responsible for the differentiation of lateral structures such as leaves and branches (Ahearn et al., 2001).

TaPME1-1B was found to be a pectinesterase, and its interaction with pectate was confirmed through the STITCH-based chemical protein analysis. Pectin methylesterase (PME) is a carbohydrate esterase family member that cleaves the ester bond between a methyl group and galacturonic acid. The optimal pectin methyl esterification in each cell type is determined by the balance between PME activity and PME inhibitors regulated by posttranslational PME inhibition. The overexpression of PMEI28 was found to result in reduced culm diameter and dwarf phenotypes in transgenic rice plants. It was found to function as a critical structure modulator by regulating the degree of pectin methyl esterification. Impairment in pectin methyl esterification affects physiological properties of the cell wall components and causes abnormal cell extensibility in culm tissue (Nguyen et al., 2017).

TaPNH1-7B, an Argonaute (PAZ Piwi domain) protein, found in the cytoplasm and intracellular spaces, has nuclease activity and acts as a positive regulator of leaf development, as evident through its predominant expression in the apex and axillary bud in *in silico* analysis. The fold expression of *TaPNH1-7B* was found to be highest during flag leaf emergence followed by fully grown flag leaf stages in the high-yielding wheat variety. The expression of *TaPNH1-7B* was optimally lower during all stages of flag leaf development in the low-yielding wheat variety, so it can be a potent gene to work with to increase the wheat yield. It was found by Nishimura et al. (2002) that OsPnh not only functions in S-adenosyl methionine (SAM) maintenance but also in leaf formation directly through vascular development. Malformed leaves with abnormal internal vascular structure were observed in antisense *OsPnh1* plants (Nishimura et al., 2002). TaBri1-3D, an LRR from the ribonuclease inhibitor-like superfamily, was found to positively regulate the development of the flag leaf in wheat. Decreased plant height with compact stature, narrow and short leaves, short internodes, decreased brassinosteroid response, and expression of BR-related genes were reported in *Brachypodium distachyon* with the BRI1-RNAi mutation (Feng et al., 2015). This supports the increased expression of *TaBri1-3D* during flag leaf development. Although most researchers primarily stress the influence of brassinosteroid-signaling genes on factors such as plant height, the impact on leaf architecture is also apparent. Altered leaf architecture, along with other brassinosteroid-insensitive phenotypic traits, was shown in Uzu 1, a mutant of barley with altered signaling of brassinosteroid due to the exchange of amino acids in the *HvBRI1* gene (Dockter et al., 2014). Similarly, shorter leaf blades and sheaths after the knockdown of *BRI1* homologs were found in maize plants (Kir et al., 2015). In ryegrass genotypes, the deletion in the LpBRI1 locus resulted in significantly narrower leaves than that of the genotypes in which the deletion was harbored (Statkevičiūtė et al., 2018). Chemical network analysis of TaBri1-D revealed its involvement in carboxylase and nitrate domains, both for synthase and transferase. A dwarf and low tillering (dlt) mutant of rice was characterized, and cloning of the dlt gene was performed through map-based cloning. DLT was found to encode a new member of the plant-specific GRAS family. The dwarf phenotype of *dlt* is similar to that of the BR-deficient mutant of rice (Tong et al., 2009).

TaRCCR1-7D was found to be a red chlorophyll catabolite reductase. Through microarray analysis, it was found to be predominantly present during developmental stages such as germination, stem elongation, and booting. The *in silico* expression of *TaRCCR1-7D* during the development of the flag leaf was found to be highly increased compared to that of other genes. In contrast, Sakuraba et al. (2013) reported less-pronounced changes in RCCR in developing or senescing leaves (Sakuraba et al., 2013). Through PPI and PCI network analyses, TaRCCR1-7D was found to interact with red chlorophyll and genes responsible for its degradation, such as *Nyc, SGRL*, and *NOL*, indicating its potential role in the reduction of chlorophyll catabolic genes and indirectly helping in the developmental processes.

*TaTSD2-6B* expression, which is a member of the methyltransferase 29 family, was found to be increased during the emergence of the flag leaf and unaffected during the senescence stage when analyzed through GENEVESTIGATOR. Similar results were found through expression analysis in high-yielding wheat varieties. Through chemical network analysis, TaTSD2-6B was found to interact with reduced cell adhesion and noncordial shoot development, similar to what was observed in the mutant of the TSD2 gene (Krupková et al., 2007).

TaSRT-5D, an inter-alpha trypsin inhibitor, was involved in stem elongation, and an increase in its expression during flag leaf emergence was also elucidated through a microarray analysis. TaSRT1-5D was found to be the most prominent in the high-yielding wheat variety during flag leaf emergence. Through network analysis, it was found to interact with intermediates of different metabolic cycles, such as the citric acid cycle. *OsSRT1*, one of the homologs of silent information regulator 2 (*SIR2*) in rice, was found to negatively regulate the process of leaf senescence by repressing the expression of the biosynthetic genes of the leaf senescence metabolic cascade and partially

through H3K9 deacetylation of *OsPME1* (Fang et al., 2016). TaOsh1-4A, a homeobox family member, is highly important for the development of the plant. An Osh15 mutant revealed the importance of the *KNOX* gene in the self-regulation of other genes and SAM (Tsuda et al., 2011).

Of all the genes in cluster 1 discussed before, *TaNAP1-7B* and *TaNOL1-4D* were the only exceptions that were found to be not associated with developmental processes. The expression of *TaNAP1-7B* increased during the emergence stage of flag leaf development. It was found to be involved in cellular component morphogenesis and organization. OsNAP was found to act as an important link between ABA and leaf senescence in rice. The mutant named ps1, that is, a gain of function mutant, significantly exhibited premature senescence, in which the expression of ABA biosynthesis genes was found to be affected. The knockdown of OsNAP was found to produce an obvious delay in the process of senescence, which most importantly slowed the decrease in the functional photosynthetic capacity and highly influenced the seed setting ratio to a higher extent (Liang et al., 2014). The expression analysis of TaNAP1-7B deciphered its negative role in stay-green character in wheat, as it was found to be highest during the flag leaf senescence stage of the low-yielding variety and lowest during that in the high-yielding variety. TaNOL1-4D was found to be a member of the SDR family, whose expression most likely increased during flag leaf senescence (Kusaba et al., 2007; Park et al., 2007). Through network analysis, it was found to play an important role in carboxylic acid synthesis. When analyzed using a microarray, its increased level of expression during senescence was quite evident through its increased level of expression during senescence. *TaNOL-4D* was also found to be highly expressed in the low-yield wheat variety during flag leaf senescence; hence, its role in senescence is quite clear. The *Nyc3* mutant was found to retain more chlorophyll a and b content than the wild-type, and a significant decrease in senescence parameters during dark incubation suggested that it is a nonfunctional stay-green mutant. In addition, a small amount of chlorophyll a, pheophytin a, and Mg derivative without $^{2+}$ in its tetrapyrrole ring accumulated in the senescent leaves of the nyc3 mutant (Sato et al., 2009). Similarly, its expression was highest during the flag leaf senescence stage in the low-yield wheat variety.

Regardless of decrease in leaf functionality, the retention of chlorophyll during leaf senescence was increased in the mutant of the stay-green *SGR* gene in rice and its orthologs in *A. thaliana* (nonyellowing), *Festuca pratensis* (senescence-induced deficiency), tomato (green-fresh), pepper (chlorophyll retainer, and pea (I; known as Mendel's green cotyledon gene) (Armstead et al., 2006; Jiang et al., 2007; Kusaba et al., 2007; Park et al., 2007; Aubry et al., 2008; Barry et al., 2008). *SGR* overexpression in rice seedlings resulted in the generation of singlet oxygen and activation of other reactive oxygen species and resulted in a chlorophyll-dependent regional cell death phenotype in the leaves (Jiang et al., 2011). It was evident that *OsSGR*, a member of the stay-green superfamily, plays an important role in the regulation of chlorophyll degradation, and the changes in the transcript expression levels of *OsSGR* are reflected in the regulation of jasmonic acid and cytokinin on chlorophyll degradation and leaf senescence. It was found to catalyze a key conversion from a red chlorophyll catabolite to a primary fluorescent catabolite during the degradation of chlorophyll (Liu et al., 2016). Chlorophyll was found to be degraded through the induction of light-harvesting chlorophyll a/b protein complex II (LHCPII) disassembly leading to the degradation of chlorophyll and chlorophyll-free LHCPII by proteases and catabolic enzymes, respectively (Park et al., 2007). However, the expression was highest in the flag leaf emergence stage in the high-yield wheat variety and flag leaf senescence stages in the low-yield wheat variety, depicting its contrary nature. TaOsl2-2B, an aspartate aminotransferase, was deciphered from network analysis and gene ontology analysis as being connected to various nitrogen-containing compounds. The overexpressed *OsL2* fusion protein in recombinant *E. coli* showed pyruvate-dependent-aminobutyric acid (GABA) transaminase activity. Induced *Osl2*-specific transcripts were found to be induced in the leaves that were senescing, and the chronological profile of accumulation of Osl2 protein was found to be linked with that of pyruvate-dependent GABA transaminase activity in the rice leaf (Ansari et al., 2005). The expression of TaOsl2-2B was also found to be elevated during the senescence of flag leaf, specifically in the low-yield wheat variety, as well as through microarray analysis indicating its significance as a senescence-associated gene.

TaNyc1-3D was found to interact with senescence-associated genes such as HCAR, SGR, and various nitrogen metabolites. Its expression was also prominently increased during flag leaf senescence. *Nyc1* encodes a membrane-localized short-chain dehydrogenase/reductase (SDR) that represents a chlorophyll b reductase that is required to catalyze the initial step of chlorophyll degradation. The *Nyc1* mutant, when analyzed, showed stay-green characters phenotypically. Similar stay-green phenotypes were reported in rice nol mutant such as *nyc1* mutants, that is, confirming that chlorophyll b degradation was selectively retained during senescence, resulting in the retention of thylakoid grana even at a later stage of senescence. It was found to severely inhibit the light-harvesting complex (Sato et al., 2009).

When over*expressed* in a semi-dwarf rice variety, *OsGATA12,* a zinc-finger transcription factor, caused an increase in the greenness of the leaf and a reduction in the tiller number and other yield-related attributes. The transgenic plants were found to be comparatively distinct from the wild type due to a significant increase in yield per area and harvest index as a result of reduced tillering. The increased greenness observed in the transformed plants was mostly due to decreased chlorophyll degradation along with chlorophyll synthesis, as the expression

of genes involved in the degradation pathway of chlorophyll was also reduced (Lu et al., 2017). Therefore, TaGATA12-3D was found to regulate the processes of catabolism in plants, as revealed through the gene ontology analysis.

Functional enrichment analysis revealed the enzymatic functions of TaPME1-1B, TaNyc3-7A, TaNyc-3D1, TaNOL-4D, TaTSD2-6B, TaOsl2-2B, and TaRCCR1-7D and the regulatory functions of TaPnNH1-7B, TaBri1-3D, TaOsh1-4A, TaSRT1-5D, and TaAct1-4B in aiding cytoskeleton organization. Overexpression of genes associated with photosynthetic activity, nutrient transport indirectly maintaining greeness of plant tissue and downregulation or knocking out of the genes responsible for senescence associated activities like chlorophyll degradation and plant cell death. Both approaches can be given as individual trials or can be combined to obtain the best results regarding higher yields. It is quite evident from the results of MD simulations that all the candidate genes except *TaNyc1-3D* and *TaSGR-5D* were found to be stable under 20s RMAD analysis, and hence, they can be validated and used for further applications.

## 5 Conclusion and outlook for the future

With the concomitant development of genetic and molecular studies for the development of plants, a huge amount of understanding about the expression patterns of genes, their cellular locations, and molecular functions is necessary. It is obvious to realize that *in silico* approaches, involving the use of bioinformatics tools, remain cost-effective methods to make rapid analysis regarding the expected effect of genes; the more factors that are taken into account, the more accurate the prediction will be. Therefore, in this article, we performed the complete structural and functional annotation of genes related to different developmental stages of flag leaf so that direct or indirect higher grain yield can be achieved in the future by exploiting the information of these genes. Several genes associated with the development of flag leaf were identified and mapped to the genome. From the variations in characteristics from the structure of genes to microarray gene expression studies and phylogeny, it is quite evident that there is a great deal of complexity within the processes controlled by these genes. The 3D structures of proteins encoded by candidate genes were formed through a homology modeling-based approach. Modeled 3D structures of all the proteins were evaluated by using dihedral analysis of the Ramachandran plot. To assess the stability of modeled structures, molecular dynamics simulations after energy minimization were performed. By analyzing the results of both MD and homology modeling, the results were found to be consistent with the known set of investigational data. Based on our results, it has been concluded that these genes can be considered important candidates for the regulation of the development of the flag leaf in wheat. The thorough investigation of these genes associated with different processes such as leaf emergence, development, nutrient remobilization, and senescence indicated that the improvement in yield production can be attained by regulating the genes associated with the source–sink pathway and senescence of flag leaf. Leaf senescence is a major determinant of yield in many cereal crops. Stay-green, delayed senescence is associated with the retention of high-yield increments due to higher photosynthetic capacity during the active stage (Gentinetta et al., 1986; Thomas and Howarth, 2000). The stay-green sorghum variety B35 has been reported by Rosenow et al. (1983). It was found to show postflowering and drought resistance with a high contribution to stable and high-yield production (Rosenow, 1983). Through gene expression analysis, it was clear that by enhancing the expression of genes such as *TaSRT1-5D*, *TaPNH1-7B*, and *TaNfl1-2B* and by downregulating genes such as *TaNAP1-7B*, *TaNOL-4D*, and *TaOsl-2B*, high-yielding wheat varieties could be generated. Therefore, the data regarding the candidate genes can be utilized to manipulate these complex processes such as the source–sink pathway and senescence for the enhancement of yield in wheat. Despite comprehensive analysis in the current study, there is still much room for improvement before a complete array of regulation of these complex processes involved in flag leaf development through validating these genes through reverse genetics and biotechnological approaches.

## 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

SM performed bioinformatics data analysis. NA performed the molecular dynamics simulation of modeled structures. SM, PK, YS, PS, NL, and VS drafted the manuscript. UK served as the principal investigator and also revised the manuscript. OD, AK, RM, KS, PB, VP, and ID reviewed the manuscript. All authors agreed to the final manuscript.

## Funding

This work was supported by SPARC, the Ministry of Education, GOI, in collaboration with the University of Massachusetts, Amherst (SPARC/2018-2019/P854/SL).

## Acknowledgments

The corresponding author is highly grateful to the Head, Department of Molecular Biology, Biotechnology & Bioinformatics, and Directorate of Research, CCS Haryana Agricultural University, for providing all the necessary facilities during the course of this work.

## Conflict of interest

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

## Publisher's note

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

## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.933560/full#supplementary-material

## References

agrochart (2021). The estimate for global cereal production in 2020 raised sharply while early prospects for cereal production in 2021 are positive. Available at: https://www.agrochart.com/en/news/7340/fao-the-estimate-for-global-cereal-production-in-2020-raised-sharply-while-early-prospects-for-cereal-production-in-2021-are-positive.html.fao-the-estimate-for-global-cereal-production-in-2020-raised-sharply-while-early-prospects-for-cereal-production-in-2021-are-positive @ www.agrochart.com

Ahearn, K. P., Johnson, H. A., Weigel, D., and Wagner, D. R. (2001). NFL1, a Nicotiana tabacum LEAFY-like gene, controls meristem initiation and floral structure. *Plant Cell Physiol.* 42, 1130–1139. doi:10.1093/pcp/pce143

Ansari, M. I., Lee, R.-H., and Chen, S.-C. G. (2005). A novel senescence-associated gene encoding gamma-aminobutyric acid (GABA):pyruvate transaminase is upregulated during rice leaf senescence. *Physiol. Plant.* 123, 1–8. doi:10.1111/j.1399-3054.2004.00430.x

Armstead, I., Donnison, I., Aubry, S., Harper, J., Hörtensteiner, S., James, C., et al. (2006). From crop to model to crop: Identifying the genetic basis of the staygreen mutation in the lolium/festuca forage and amenity grasses. *New Phytol.* 172, 592–597. doi:10.1111/j.1469-8137.2006.01922.x

Aubry, S., Mani, J., and Hörtensteiner, S. (2008). Stay-green protein, defective in Mendel's green cotyledon mutant, acts independent and upstream of pheophorbide a oxygenase in the chlorophyll catabolic pathway. *Plant Mol. Biol.* 67, 243–256. doi:10.1007/s11103-008-9314-8

Bahar, I., Atilgan, A. R., Demirel, M. C., and Erman, B. (1998). Vibrational dynamics of folded proteins: Significance of slow and fast motions in relation to function and stability. *Phys. Rev. Lett.* 80, 2733–2736. doi:10.1103/physrevlett.80.2733

Banitaba, A., Naderi, M., Javanmard, H., and Emami, B. (2007). *Effect of flag leaf and awn removal on vegetative traits, grain yield and yield components of bread wheat*. Colombia: Triticum aestivum L.

Barry, C. S., McQuinn, R. P., Chung, M.-Y., Besuden, A., and Giovannoni, J. J. (2008). Amino acid substitutions in homologs of the STAY-GREEN protein are responsible for the green-flesh and chlorophyll retainer mutations of tomato and pepper. *Plant Physiol.* 147, 179–187. doi:10.1104/pp.108.118430

Biswal, A. K., and Kohli, A. (2013). Cereal flag leaf adaptations for grain yield under drought: Knowledge status and gaps. *Mol. Breed.* 31, 749–766. doi:10.1007/s11032-013-9847-7

Bírsîn, M. A. (2005). Effects of removal of some photosynthetic structures on some yield components in wheat. *J. Agric. Sci.* 11, 364–367. doi:10.1501/Tarimbil_0000000559

Blade, S. F., and Baker, R. J. (1991). Kernel weight response to source-sink changes in spring wheat. *Crop Sci.* 31, 1117–1120. doi:10.2135/cropsci1991.0011183x003100050005x

Bolser, D. M., Kerhornou, A., Walts, B., and Kersey, P. (2015). Triticeae resources in ensembl plants. *Plant Cell Physiol.* 56, e3. doi:10.1093/pcp/pcu183

Derkx, A. P., Orford, S., Griffiths, S., Foulkes, M. J., and Hawkesford, M. J. (2012). Identification of differentially senescing mutants of wheat and impacts on yield, biomass and nitrogen partitioning. *J. Integr. Plant Biol.* 54, 555–566. doi:10.1111/j.1744-7909.2012.01144.x

Dhaliwal, A. K., Mohan, A., and Gill, K. S. (2014). Comparative analysis of ABCB1 reveals novel structural and functional conservation between monocots and dicots. *Front. Plant Sci.* 5, 657. doi:10.3389/fpls.2014.00657

Dockter, C., Gruszka, D., Braumann, I., Druka, A., Druka, I., Franckowiak, J., et al. (2014). Induced variations in brassinosteroid genes define barley height and sturdiness, and expand the green revolution genetic toolkit. *Plant Physiol.* 166, 1912–1927. doi:10.1104/pp.114.250738

Duwayri, M. (1984). Effect of flag leaf and awn removal on grain yield and yield components of wheat grown under dryland conditions. *Field Crops Res.* 8, 307–313. doi:10.1016/0378-4290(84)90077-7

Fan, X., Cui, F., Zhao, C., Zhang, W., Yang, L., Zhao, X., et al. (2015). QTLs for flag leaf size and their influence on yield-related traits in wheat (*Triticum aestivum* L.). *Mol. Breed.* 35, 24. doi:10.1007/s11032-015-0205-9

Fang, C., Zhang, H., Wan, J., Wu, Y., Li, K., Jin, C., et al. (2016). Control of leaf senescence by an MeOH-jasmonates cascade that is epigenetically regulated by OsSRT1 in rice. *Mol. Plant* 9, 1366–1378. doi:10.1016/j.molp.2016.07.007

FAO (2020). En @ Www.Fao.Org. Available at:http://www.fao.org/fishery/statistics/software/fishstatj/en.

Feng, Y., Yin, Y., and Fei, S. (2015). Down-regulation of BdBRI1, a putative brassinosteroid receptor gene produces a dwarf phenotype with enhanced drought tolerance in Brachypodium distachyon. *Plant Sci.* 234, 163–173. doi:10.1016/j.plantsci.2015.02.015

Franz, M., Rodriguez, H., Lopes, C., Zuberi, K., Montojo, J., Bader, G. D., et al. (2018). GeneMANIA update 2018. *Nucleic Acids Res.* 46, W60–W64. doi:10.1093/nar/gky311

Gajula, M. N. V. P., Kumar, A., and Ijaq, J. (2016). Protocol for molecular dynamics simulations of proteins. *Bio-protocol* 6, e2051. doi:10.21769/bioprotoc.2051

Garg, V. K., Avashthi, H., Tiwari, A., Jain, P. A., Ramkete, P. W., Kayastha, A. M., et al. (2016). MFPPI-multi FASTA ProtParam interface. *Bioinformation* 12, 74–77. doi:10.6026/97320630012074

Gautam, T., Saripalli, G., Gahlaut, V., Kumar, A., Sharma, P. K., Balyan, H. S., et al. (2019). Further studies on sugar transporter (SWEET) genes in wheat (*Triticum aestivum* L.). *Mol. Biol. Rep.* 46, 2327–2353. doi:10.1007/s11033-019-04691-0

Gentinetta, E., Ceppl, D., Lepori, C., Perico, G., Motto, M., Salamini, F., et al. (1986). A major gene for delayed senescence in maize. Pattern of photosynthates accumulation and inheritance. *Plant Breed.* 97, 193–203. doi:10.1111/j.1439-0523.1986.tb01053.x

Giraldo, P., Benavente, E., Manzano-Agugliaro, F., and Gimenez, E. (2019). Worldwide research trends on wheat and barley: A bibliometric comparative analysis. *Agronomy* 9, 352. doi:10.3390/agronomy9070352

Grennan, A. K. (2006). Genevestigator. Facilitating web-based gene-expression analysis. *Plant Physiol.* 141, 1164–1166. doi:10.1104/pp.104.900198

Gui, J., Zheng, S., Shen, J., and Li, L. (2015). Grain setting defect1 (GSD1) function in rice depends on S-acylation and interacts with actin 1 (OsACT1) at its C-terminal. *Front. Plant Sci.* 6, 804. doi:10.3389/fpls.2015.00804

Hofer, J., Turner, L., Hellens, R., Ambrose, M., Matthews, P., Michael, A., et al. (1997). UNIFOLIATA regulates leaf and flower morphogenesis in pea. *Curr. Biol.* 7, 581–587. doi:10.1016/s0960-9822(06)00257-0

Hörtensteiner, S., and Kräutler, B. (2011). Chlorophyll breakdown in higher plants. *Biochim. Biophys. Acta* 1807, 977–988. doi:10.1016/j.bbabio.2010.12.007

Hu, B., Jin, J., Guo, A.-Y., Zhang, H., Luo, J., Gao, G., et al. (2015). Gsds 2.0: An upgraded gene feature visualization server. *Bioinformatics* 31, 1296–1297. doi:10.1093/bioinformatics/btu817

Humann, J. L., Lee, T., Ficklin, S., and Main, D. (2019). Structural and functional annotation of eukaryotic genomes with GenSAS. *Methods Mol. Biol.* 1962, 29–51. doi:10.1007/978-1-4939-9173-0_3

importance (2019). The importance of protecting the flag leaf in winter wheat. Available at: https://www.fas.scot/news/the-importance-of-protecting-the-flag-leaf-in-winter-wheat/.bf642e28db4188a97b09ffb580409bd5e7a2a4 @ www.fas.scot

Jenuth, J. P. (2000). *Bioinformatics methods and protocols*. New york: Springer, 301–312.

Jiang, H., Chen, Y., Li, M., Xu, X., and Wu, G. (2011). Overexpression of SGR results in oxidative stress and lesion-mimic cell death in rice seedlings. *J. Integr. Plant Biol.* 53, 375–387. doi:10.1111/j.1744-7909.2011.01037.x

Jiang, H., Li, M., Liang, N., Yan, H., Wei, Y., Xu, X., et al. (2007). Molecular cloning and function analysis of the stay green gene in rice. *Plant J.* 52, 197–209. doi:10.1111/j.1365-313X.2007.03221.x

Kir, G., Ye, H., Nelissen, H., Neelakandan, A. K., Kusnandar, A. S., Luo, A., et al. (2015). RNA interference knockdown of BRASSINOSTEROID INSENSITIVE1 in maize reveals novel functions for brassinosteroid signaling in controlling plant architecture. *Plant Physiol.* 169, 826–839. doi:10.1104/pp.15.00367

Koshkin, E. I., and Tararina, V. V. (1989). Yield and source/sink relations of spring wheat cultivars. *Field Crops Res.* 22, 297–306. doi:10.1016/0378-4290(89)90029-4

Krupková, E., Immerzeel, P., Pauly, M., and Schmülling, T. (2007). The tumorous shoot Development2 gene of Arabidopsis encoding A putative methyltransferase is required for cell adhesion and Co-ordinated plant development. *Plant J.* 50, 735–750. doi:10.1111/j.1365-313X.2007.03123.x

Kuhn, M., von Mering, C., Campillos, M., Jensen, L. J., and Bork, P. (2007). Stitch: Interaction networks of chemicals and proteins. *Nucleic Acids Res.* 36, D684–D688. doi:10.1093/nar/gkm795

Kumar, A., Kumar, S., Kumar, A., Sharma, N., Sharma, M., Singh, K. P., et al. (2018a). Homology modeling, molecular docking and molecular dynamics based functional insights into rice urease bound to urea. *Proc. Natl. Acad. Sci. India Sect. B. Biol. Sci.* 88, 1539–1548. doi:10.1007/s40011-017-0898-0

Kumar, A., Kumar, S., Kumar, U., Suravajhala, P., and Gajula, M. N. V. P. (2016). Functional and structural insights into novel DREB1A transcription factors in common wheat (*Triticum aestivum* L.): A molecular modeling approach. *Comput. Biol. Chem.* 64, 217–226. doi:10.1016/j.compbiolchem.2016.07.008

Kumar, A., Sharma, M., Kumar, S., Tyagi, P., Wani, S. H., Gajula, M. N. V. P., et al. (2018b). Functional and structural insights into candidate genes associated with nitrogen and phosphorus nutrition in wheat (*Triticum aestivum* L.). *Int. J. Biol. Macromol.* 118, 76–91. doi:10.1016/j.ijbiomac.2018.06.009

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

Kusaba, M., Ito, H., Morita, R., Iida, S., Sato, Y., Fujimoto, M., et al. (2007). Rice NON-YELLOW COLORING1 is involved in light-harvesting complex II and grana degradation during leaf senescence. *Plant Cell* 19, 1362–1375. doi:10.1105/tpc.106.042911

Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993). Procheck: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26, 283–291. doi:10.1107/s0021889892009944

Laskowski, R. A., Watson, J. D., and Thornton, J. M. (2005). ProFunc: A server for predicting protein function from 3D structure. *Nucleic Acids Res.* 33, W89–W93. doi:10.1093/nar/gki414

Liang, C., Wang, Y., Zhu, Y., Tang, J., Hu, B., Liu, L., et al. (2014). OsNAP connects abscisic acid and leaf senescence by fine-tuning abscisic acid biosynthesis and directly targeting senescence-associated genes in rice. *Proc. Natl. Acad. Sci. U. S. A.* 111, 10013–10018. doi:10.1073/pnas.1321568111

Liu, K., Xu, H., Liu, G., Guan, P., Zhou, X., Peng, H., et al. (2018a). QTL mapping of flag leaf-related traits in wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 131, 839–849. doi:10.1007/s00122-017-3040-z

Liu, L., Xu, W., Hu, X., Liu, H., and Lin, Y. (2016). W-box and G-box elements play important roles in early senescence of rice flag leaf. *Sci. Rep.* 6, 20881. doi:10.1038/srep20881

Liu, L., Zhou, Y., Zhou, G., Ye, R., Zhao, L., Li, X., et al. (2008). Identification of early senescence-associated genes in rice flag leaves. *Plant Mol. Biol.* 67, 37–55. doi:10.1007/s11103-008-9300-1

Liu, Y., Tao, Y., Wang, Z., Guo, Q., Wu, F., Yang, X., et al. (2018b). Identification of QTL for flag leaf length in common wheat and their pleiotropic effects. *Mol. Breed.* 38, 11. doi:10.1007/s11032-017-0766-x

Lu, G., Casaretto, J. A., Ying, S., Mahmood, K., Liu, F., Bi, Y.-M., et al. (2017). Overexpression of OsGATA12 regulates chlorophyll content, delays plant senescence and improves rice yield under high density planting. *Plant Mol. Biol.* 94, 215–227. doi:10.1007/s11103-017-0604-x

Mathpal, P., Kumar, U., Kumar, A., Kumar, S., Malik, S., Kumar, N., et al. (2018). Identification, expression analysis, and molecular modeling of Iron-deficiency-specific clone 3 (Ids3)-like gene in hexaploid wheat. *Biotech* 8, 219. doi:10.1007/s13205-018-1230-2

ncbi The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information. Available at: https://www.ncbi.nlm.nih.gov.a47505ef5f64d6f8117bd892e386d0af972e @ www.ncbi.nlm.nih.gov

Nguyen, H. P., Jeong, H. Y., Jeon, S. H., Kim, D., and Lee, C. (2017). Rice pectin methylesterase inhibitor28 (OsPMEI28) encodes a functional PMEI and its overexpression results in a dwarf phenotype through increased pectin methylesterification levels. *J. Plant Physiol.* 208, 17–25. doi:10.1016/j.jplph.2016.11.006

Nishimura, A., Ito, M., Kamiya, N., Sato, Y., and Matsuoka, M. (2002). OsPNH1 regulates leaf development and maintenance of the shoot apical meristem in rice. *Plant J.* 30, 189–201. doi:10.1046/j.1365-313x.2002.01279.x

Palmer, N. A., Donze-Reiner, T., Horvath, D., Heng-Moss, T., Waters, B., Tobias, C., et al. (2015). Switchgrass (Panicum virgatum L) flag leaf transcriptomes reveal molecular signatures of leaf development, senescence, and mineral dynamics. *Funct. Integr. Genomics* 15, 1–16. doi:10.1007/s10142-014-0393-0

Park, S.-Y., Yu, J.-W., Park, J.-S., Li, J., Yoo, S.-C., Lee, N.-Y., et al. (2007). The senescence-induced staygreen protein regulates chlorophyll degradation. *Plant Cell* 19, 1649–1664. doi:10.1105/tpc.106.044891

Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., et al. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* 30, 70–82. doi:10.1002/pro.3943

Reinbothe, C., El Bakkouri, M., Buhr, F., Muraki, N., Nomata, J., Kurisu, G., et al. (2010). Chlorophyll biosynthesis: Spotlight on protochlorophyllide reduction. *Trends Plant Sci.* 15, 614–624. doi:10.1016/j.tplants.2010.07.002

Rosenow, D., Mitchell, R., and Patancheru, A. P. (1983). *Breeding for resistance to root and stalk rots in Texas*. Sorghum root Stalk rots. India: ICRISTAT. 209–217.

Sakuraba, Y., Kim, Y.-S., Yoo, S.-C., Hörtensteiner, S., and Paek, N.-C. (2013). 7-Hydroxymethyl chlorophyll a reductase functions in metabolic channeling of chlorophyll breakdown intermediates during leaf senescence. *Biochem. Biophys. Res. Commun.* 430, 32–37. doi:10.1016/j.bbrc.2012.11.050

Sato, Y., Morita, R., Katsuma, S., Nishimura, M., Tanaka, A., Kusaba, M., et al. (2009). Two short-chain dehydrogenase/reductases, NON-YELLOW COLORING 1 and NYC1-LIKE, are required for chlorophyll b and light-harvesting complex II degradation during senescence in rice. *Plant J.* 57, 120–131. doi:10.1111/j.1365-313X.2008.03670.x

Schwede, T., Kopp, J., Guex, N., and Peitsch, M. C. (2003). SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res.* 31, 3381–3385. doi:10.1093/nar/gkg520

Statkevičiūtė, G., Kemešytė, V., Aleliūnas, A., Jonavičienė, K., and Brazauskas, G. (2018). LpBRI1 polymorphism association with flag leaf architecture in perennial ryegrass. *Zemdirbyste* 105, 33–38. doi:10.13080/z-a.2018.105.005

Thomas, H., and Howarth, C. J. (2000). Five ways to stay green. *J. Exp. Bot.* 51, 329–337. doi:10.1093/jexbot/51.suppl_1.329

Tong, H., Jin, Y., Liu, W., Li, F., Fang, J., Yin, Y., et al. (2009). Dwarf and low-tillering, A new member of the gras family, plays positive roles in brassinosteroid signaling in rice. *Plant J.* 58, 803–816. doi:10.1111/j.1365-313X.2009.03825.x

Tsuda, K., Ito, Y., Sato, Y., and Kurata, N. (2011). Positive autoregulation of a KNOX gene is essential for shoot apical meristem maintenance in rice. *Plant Cell* 23, 4368–4381. doi:10.1105/tpc.111.090050

Warburton, M., Skovmand, B., and Mujeeb-Kazi, A. (2002). The molecular genetic characterization of the'Bobwhite'bread wheat family using AFLPs and the effect of the T1BL. 1RS translocation. *Theor. Appl. Genet.* 104, 868–873. doi:10.1007/s00122-001-0816-x

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., et al. (2018). SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46, W296–W303. doi:10.1093/nar/gky427

Yan, X., Zhao, L., Ren, Y., Zhang, N., Dong, Z., Chen, F., et al. (2020). Identification of genetic loci and a candidate gene related to flag leaf traits in common wheat by genome-wide association study and linkage mapping. *Mol. Breed.* 40, 58. doi:10.1007/s11032-020-01135-7

Youssef, S., and Salem, A. (1976). Contribution of individual plant parts to grain yield and protein content and quality of two wheat varieties [Giza 155, Chenab] differing in protein content. *Alex J. Agric. Res.* 24, 55-61.

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 30 September 2022
DOI 10.3389/fgene.2022.941954



OPEN ACCESS

EDITED BY
Dwijesh Chandra Mishra,
Indian Council of Agricultural Research, India

REVIEWED BY
Jun You,
Oil Crops Research Institute (CAAS), China
Luís R. Silva,
University of Beira Interior, Portugal

*CORRESPONDENCE
Xuexiao Zou,
zouxuexiao428@163.com
Cheng Qin,
qincheng1001@163.com

†These authors have contributed equally to this work

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

RECEIVED 12 May 2022
ACCEPTED 13 September 2022
PUBLISHED 30 September 2022

CITATION
Luo Y, Yang S, Luo X, Li J, Li T, Tang X, Liu F, Zou X and Qin C (2022), Genome-wide analysis of OFP gene family in pepper (Capsicum annuum L.). *Front. Genet.* 13:941954. doi: 10.3389/fgene.2022.941954



# Genome-wide analysis of *OFP* gene family in pepper (*Capsicum annuum* L.)

Yin Luo[1†], Shimei Yang[2†], Xirong Luo[2†], Jing Li[2], Tangyan Li[2], Xiangqun Tang[3], Feng Liu[1,4], Xuexiao Zou[1,4*] and Cheng Qin[2*]

[1]Longping Branch, College of Biology, Hunan University, Changsha, China, [2]Engineering Research Center of Zunyi Pepper Germplasm Resources Conservation and Breeding Cultivation of Guizhou Province, Department of Modern Agriculture, Zunyi Vocational and Technical College, Zunyi, China, [3]Key Lab of Zunyi Crop Gene Resource and Germplasm Innovation, Zunyi Academy of Agricultural Sciences, Zunyi, China, [4]College of Horticulture, Hunan Agricultural University, Changsha, China

Ovate family proteins (OFPs) are transcriptional inhibitors that regulate plant growth and development and play important roles in the synthesis of secondary cell walls during pollen development. This study identified the pepper *OFP* gene family based on the genome-wide analysis and used bioinformatics methods to provide a fundamental profile of the gene family. 74 *OFP* genes with typical Ovate domain were identified in cultivated pepper Zunla-1, wild pepper Chiltepin and CM334. Chromosome mapping revealed that *CazOFP* genes were unevenly distributed on 11 chromosomes and Chr00 in Zunla-1, *CacOFP* genes on 12 chromosomes in Chiltepin, and *CamOFP* genes on 12 chromosomes and two Scaffflods in CM334. Gene structure analysis revealed that *CaOFP* genes possessed 1-3 exons, and the analysis of physicochemical properties suggested that *CaOFP*s were hydrophilic. Many *cis*-acting elements were identified in the promoter region of *CaOFP* genes, including ABRE, ARE, Box 4, G-box, TC-rich, and TCT-motif. The expression patterns of pepper at different growth stages showed that *CaOFP* genes were actively involved in the growth and fruit development of pepper, and *CazOFP16* and *CazOFP17* were actively involved in response to multiple hormones and stress events. qRT-PCR was also used to verify the expression of *CazOFP* gene in two developmental stages of seven pepper varieties with different fruit shapes, and it was found that *CaOFP* genes may be involved in the formation of fruit type in pepper. This study provides theoretical and practical evidence for future research on the *OFP* gene family.

KEYWORDS

**pepper, *OFP* genes, phylogenetic tree, expression characteristics analysis, qRT-PCR**

# Introduction

Transcription factors (TFs) play fundamental roles in the growth and development of higher plants, as well as in their responses to the external environment. Typical TFs in higher plants possess DNA-binding domains, transcription regulation domains, oligomerization sites, and nuclear localization signals; the interaction between these domains and *cis*-acting elements enables TFs to regulate gene expression (Jiang et al., 2011). Ovate family proteins (OFPs) are plant-specific, multigene family members usually with a conserved Ovate domain. Homologs of *OFPs* have been reported in higher plants, mosses, and lycophytes (Wang et al., 2016).

Pepper is an important vegetable with high standards for its quality, spiciness, and shape, but less research has been done to study its fruit shape. Fruit shape is one of the most important quality traits of pepper, and is also one of the main attributes for the classification of pepper fruit types (Paran and Van Der Knaap, 2007). In crops, yield is regulated by environmental factors and fruit morphology. Genes regulating fruit morphology were first identified in tomatoes, and *OFP* genes are highly involved in this process (Liu et al., 2002; Snouffer et al., 2020). According to previous studies, *OFP* genes were found to be involved in the regulation of shape in different varieties of crops, such as rice, tomato, and melon. Moreover, overexpression of *OFP* destroyed plant organ shape in the process of phytohormone regulation, which suggested that *OFP* genes are involved in regulating hormones and signaling pathways within developing organs (Snouffer et al., 2020). *OFP* genes related to pepper fruit shape were cloned from different pepper varieties, and there were significant differences among the different fruit shapes (Tsaballa et al., 2011). A recent study revealed that virus-induced gene silencing of the pepper ortholog *CaOFP20* resulted in increased fruit elongation on two independent backgrounds (Borovsky et al., 2021).

*OFP* genes play important roles in the growth and development of plants, including flowering (Yuan et al., 2016), growth of roots and stems (Xu et al., 2018), and formation of secondary cell walls (Li et al., 2011). Yuan et al. (2016) detected 12 *VvOFP* genes in different tissues, and the expression levels of *VvOFP1*, *VvOFP2*, *VvOFP4*, *VvOFP5*, *VvOFP6*, *VvOFP7*, *VvOFP12*, and *VvOFP17* were relatively high in the week before flowering, suggesting that *OFP* TFs regulate fruit development during anthesis and the week before anthesis, whereas *VvOFP5* and *VvOFP2* are involved in the regulation of flowering and reproductive growth, respectively (Yuan et al., 2016). Li et al. (2011) also found that *AtOFP4* regulates the formation of secondary cell walls *via* its interaction with the KNAX7 protein.

In this study, we systematically characterized the *CaOFP* transcription factors at the genome-wide level, analyzed its phylogenetic tree, physicochemical properties, conserved structure, chromosomal localization, response to biotic and abiotic stresses, and provided a theoretical basis for studying the function of *OFP* genes in pepper.

# Materials and methods

## Material and RNA extraction

Seven pepper varieties with different fruit shapes were used in this study. Plants were grown in the greenhouse of the department of modern agriculture, Zunyi vocational and technical college in the spring of 2022 (Zunyi, Guizhou, 107°045 ′E, 27°710′ N). The pepper fruits were harvested at two different developmental stages including fruit with mature green and fruit at breaker plus 5 days. All experiments contained three biological repeats, and each replicate was a pooled sample of 10 fruits of uniform size from five individual plants. In total, 42 samples were immediately frozen in liquid nitrogen. RNA was extracted from the collected samples using the TianGene RNA Extraction Kit (DP432, Beijing, China). We then added fruit material with a weight of 50–100 ng for aseptic freezing grinding; 450 μL for oscillation mixing. This was transferred to the CS filter column and centrifuged for 3 min (12,000 rpm). The supernatant was transferred from the collection tube with a pipette gun to the Rnase-free centrifuge tube. Then, the supernatant was added and 0.5 times of anhydrous ethanol was mixed into the centrifuge tube and then transferred to the adsorption column CR3 for centrifugation for 30 s (12,000 rpm). A drop of 80 μL DNase I was added to the center of the collecting tube and left at room temperature for 15 min 250 μL of protein-removing solution RW1 was added to the adsorption column CR3, and left to stand at room temperature for 2 min, before being centrifuged for 30 s (12,000 rpm) (this procedure was repeated once). We then took an enzyme-free centrifuge tube and placed the adsorption column in a new centrifuge tube for several minutes (until the rinsing solution RW was dried). 50 μL Rnase-free dd$H_2O$ was then vertically added to the adsorption column, and the obtained RNA was stored at −80°C for further Quantitative Real-Time Polymerase Chain Reaction (qRT-PCR).

## Identification and characterization of *OFP* TFs in pepper

The whole-genome data (v2.0) of *Capsicum annuum* L. (Zunla-1 and Chiltepin) were downloaded from the database (http://peppersequence.genomics.cn) (Qin et al., 2014). CM334 genome sequence was downloaded from National Center for Biotechnology Information (https://www.ncbi.nlm.nih.gov/, GCA_000512255.2) (S Kim et al., 2014). To identify the *OFP* TFs, the Pfam model (PF04844) was downloaded from the Pfam database (Mistry et al., 2021), and the HMMER package was used for whole genomic and proteomic alignments. The verification was performed using an online tool, the HMMER web server (https://www.ebi.ac.uk/Tools/hmmer/) (Potter et al., 2018). *AtOFPS* were obtained from the *Arabidopsis* reference genome (https://www.arabidopsis.org/).

## Gene structure, conserved motifs, and phylogenetic analysis of *OFP* genes in pepper

The phylogenetic tree was constructed based on the *CaOFPs* protein sequences in pepper and *AtOFPs* protein sequences in *Arabidopsis thaliana*. MEGA X was used to construct a phylogenetic tree of *CaOFP* genes according to the neighbor-joining method with 1000 bootstrap reiterations (Kumar et al., 2016). The structures of *CaOFP* genes were analyzed using the online tool GSDS 2.0 (http://gsds.gao-lab.org/) (Hu et al., 2015), and the analysis of conserved motifs was conducted using Meme (https://meme-suite.org/meme/) (Bailey et al., 2015) and visualized using TBtools (Chen et al., 2020). Physicochemical properties of the *OFP* gene family were analyzed using the online tool Expasy (Artimo et al., 2012).

## Chromosome mapping and collinearity analysis of *OFP* genes in pepper

The structure file of *OFP* genes in pepper was extracted by TBtools to determine their starting and ending position in the chromosomes (Chen et al., 2020). The protein database was constructed using the Zunla-1 genomic protein file, Chiltepin genomic protein file and CM334 genomic protein file for comparison. The collinearity analysis was performed using MCScanX software (Wang et al., 2013). The chromosome mapping and collinearity results of *OFP* genes were visualized using TBtools (Chen et al., 2020).

## Selection pressure and *cis*-acting element analysis of *OFP* genes in pepper

The online tool Pal2nal (http://www.bork.embl.de/pal2nal/) (Suyama et al., 2006) was used for the analysis of selection pressure of nine pairs of paralogous genes. The 2000-bp upstream sequences of the *OFP* genes were extracted using Bedtools (Quinlan, 2014), analyzed using the online tool Plantcare (Lescot et al., 2002), and visualized using TBtools.

## Characterization of *OFP*-gene expression in pepper

Taking zunla-1 as the reference genome, the gene expression data of pepper were obtained from the NCBI (Accession No. GSE45037) and Pepper Information Data Center (Qin et al., 2014; Liu et al., 2017) (http://pepperhub.hzau.edu.cn). Experimental treatment and data analysis were conducted as described by (Liu et al., 2017), and the gene expression data were visualized using TBtools (Chen et al., 2020).

## cDNA synthesis and qRT-PCR analysis

Expression pattern analysis was investigated using qRT-PCR. Its amplification was carried out using SYBR Green Pro Taq HS (Takara, Dalian, China) according to the manufacturer's instructions. qRT-PCR was performed on an Applied Biosystems 7500 Real-Time PCR System (Applied Biosystems, Foster City, CA, United States) using the following program: 95°C for 30 s, followed by 40 cycles of 95°C for 5 s and 60°C for 30 s. The *CaActin* (GenBank No. DQ832719) and *CaUbiquitin* (GenBank No. AY496112) pepper genes was amplified as two control genes. Three biological replicates and three measurements for each replicate were performed under identical conditions. Analysis of the relative mRNA expression data was performed using $2^{-\Delta\Delta Ct}$ (Livak and Schmittgen, 2001). All primers, designed by Primer3plus (http://www.primer3plus.com/cgi-bin/dev/primer3plus.cgi) and used in this study, were listed in Supplementary Table S1.

# Results

## Identification, phylogenetic analysis, and physicochemical properties of pepper *OFP* genes

After aligning to the pepper proteomic database, 74 *CaOFP* genes were identified in Zunla-1, Chiltepin and CM334 (Table 1; Supplementary Table S2), and the conserved structures were found in all *CaOFP* genes. Phylogenetic data showed that the pepper *OFP* genes were divided into two subfamilies, including Class I and II. The number of *CaOFP* genes of Zunla-1, Chiltepin and CM334 in two subfamilies was 12, 12, and 13, respectively (Figure 1). All *OFP* genes were named according to their positions on the chromosomes. *CaOFPs* consisted of 103–436 amino acids; their molecular mass ranged from 11.82 to 49.45 (kDa), and the isoelectric point (pI) ranged from 4.33 to 9.94. The results indicated that *CaOFPs* are amphoteric proteins, with the instability index ranging from 31.34 to 76.51. The total average hydrophilicity was between -1.030–0.072, and almost all *CaOFPs* were hydrophilic proteins except for *CamOFP25*, which was 0.072 hydrophobic protein.

## Structure characterization and motif analysis of *CaOFP* genes

The motif analysis of *OFP* genes based on 10 motifs showed that the highest number of motifs was recorded for *CazOFP17*, *CazOFP21*, *CacOFP17*, *CacOFP18*, *CacOFP20*, *CacOFP21*, *CamOFP10* and *CamOFP20*, containing 9 motifs (Figure 2C). According to their motifs, Class I can be divided into motif 1, motif 2, motif 4, motif 6 and motif 9 as Class I-A subclade and

10.3389/fgene.2022.941954

TABLE 1 OFP gene information of pepper.

| Gene_id | Rename | Number of amino acids | Molecular weight | pI | Instability index | GRAVY |
|---|---|---|---|---|---|---|
| Capana01g003669 | CazOFP1 | 190 | 21,309.00 | 9.54 | 75.18 | −0.502 |
| Capana01g003670 | CazOFP2 | 241 | 27,214.99 | 4.33 | 46.62 | −0.509 |
| Capana02g001663 | CazOFP3 | 381 | 44,895.92 | 8.40 | 66.03 | −1.020 |
| Capana02g002672 | CazOFP4 | 308 | 35,371.74 | 9.64 | 59.98 | −1.009 |
| Capana03g000334 | CazOFP5 | 281 | 31,625.45 | 8.66 | 47.31 | −0.691 |
| Capana03g003702 | CazOFP6 | 231 | 26,239.81 | 4.67 | 52.50 | −0.630 |
| Capana04g000381 | CazOFP7 | 274 | 31,578.51 | 7.09 | 65.11 | −0.717 |
| Capana05g002442 | CazOFP8 | 228 | 25,629.91 | 4.68 | 53.89 | −0.358 |
| Capana06g000283 | CazOFP9 | 339 | 38,208.89 | 9.66 | 63.37 | −0.797 |
| Capana06g000645 | CazOFP10 | 155 | 17,809.74 | 5.38 | 44.69 | −0.048 |
| Capana07g001811 | CazOFP11 | 164 | 18,936.63 | 8.99 | 66.82 | −0.678 |
| Capana09g000310 | CazOFP12 | 282 | 32,488.54 | 9.14 | 43.56 | −0.873 |
| Capana09g001195 | CazOFP13 | 239 | 26,380.44 | 5.37 | 38.63 | −0.308 |
| Capana10g001230 | CazOFP14 | 176 | 20,071.91 | 9.11 | 54.07 | −0.520 |
| Capana10g002095 | CazOFP15 | 145 | 16,616.02 | 8.67 | 51.31 | −0.385 |
| Capana10g002096 | CazOFP16 | 145 | 16,616.02 | 8.67 | 51.31 | −0.385 |
| Capana10g002097 | CazOFP17 | 320 | 36,192.28 | 4.55 | 46.98 | −0.531 |
| Capana10g002098 | CazOFP18 | 168 | 19,101.53 | 4.42 | 38.06 | −0.330 |
| Capana10g002099 | CazOFP19 | 165 | 18,469.72 | 5.20 | 52.40 | −0.420 |
| Capana10g002100 | CazOFP20 | 111 | 12,768.27 | 4.66 | 35.64 | −0.372 |
| Capana10g002104 | CazOFP21 | 320 | 36,161.30 | 4.83 | 46.56 | −0.529 |
| Capana11g000348 | CazOFP22 | 135 | 15,068.29 | 7.64 | 48.28 | −0.591 |
| Capana12g002165 | CazOFP23 | 250 | 28,182.00 | 4.35 | 55.62 | −0.657 |
| Capana00g002905 | CazOFP24 | 406 | 46,225.56 | 9.72 | 51.58 | −0.823 |
| Capang01g001844 | CacOFP1 | 231 | 26,215.81 | 4.62 | 51.89 | −0.594 |
| Capang01g003395 | CacOFP2 | 406 | 46,197.50 | 9.68 | 51.17 | −0.821 |
| Capang02g001492 | CacOFP3 | 381 | 44,981.97 | 8.21 | 67.09 | −1.030 |
| Capang02g002400 | CacOFP4 | 335 | 38,506.07 | 8.92 | 67.33 | −0.986 |
| Capang03g000309 | CacOFP5 | 281 | 31,548.33 | 8.67 | 47.72 | −0.707 |
| Capang04g000354 | CacOFP6 | 264 | 30,636.70 | 8.75 | 68.19 | −0.672 |
| Capang05g001990 | CacOFP7 | 227 | 25,532.79 | 4.68 | 53.23 | −0.353 |
| Capang06g000272 | CacOFP8 | 339 | 38,208.89 | 9.66 | 63.37 | −0.797 |
| Capang06g000600 | CacOFP9 | 155 | 17,809.74 | 5.38 | 44.69 | −0.048 |
| Capang07g001809 | CacOFP10 | 164 | 18,936.63 | 8.99 | 66.82 | −0.678 |
| Capang08g000595 | CacOFP11 | 192 | 21,483.16 | 9.54 | 76.51 | −0.505 |
| Capang08g000596 | CacOFP12 | 274 | 30,519.62 | 4.33 | 43.22 | −0.407 |
| Capang09g000292 | CacOFP13 | 282 | 32,520.60 | 9.14 | 43.56 | −0.881 |
| Capang09g000909 | CacOFP14 | 239 | 26,394.46 | 5.37 | 37.50 | −0.308 |
| Capang10g001043 | CacOFP15 | 340 | 38,301.09 | 9.75 | 68.07 | −0.807 |
| Capang10g001855 | CacOFP16 | 147 | 16,850.28 | 8.68 | 50.75 | −0.366 |
| Capang10g001856 | CacOFP17 | 320 | 36,274.43 | 4.55 | 44.01 | −0.515 |
| Capang10g001857 | CacOFP18 | 315 | 35,270.45 | 4.94 | 33.27 | −0.520 |
| Capang10g001859 | CacOFP19 | 238 | 27,195.28 | 4.77 | 44.44 | −0.541 |
| Capang10g001861 | CacOFP20 | 313 | 35,222.48 | 5.25 | 43.47 | −0.406 |
| Capang10g001863 | CacOFP21 | 320 | 36,140.41 | 4.97 | 46.63 | −0.509 |
| Capang11g000323 | CacOFP22 | 128 | 14,703.09 | 8.77 | 45.06 | −0.669 |
| Capang11g000324 | CacOFP23 | 120 | 13,854.05 | 8.25 | 45.97 | −0.594 |

(Continued on following page)

TABLE 1 (*Continued*) OFP gene information of pepper.

| Gene_id | Rename | Number of amino acids | Molecular weight | pI | Instability index | GRAVY |
|---|---|---|---|---|---|---|
| Capang12g001856 | CacOFP24 | 250 | 28,210.05 | 4.39 | 54.54 | −0.661 |
| PHT93985.1 | CamOFP1 | 368 | 40,626.74 | 4.91 | 46.10 | −0.635 |
| PHT95520.1 | CamOFP2 | 126 | 14,074.08 | 6.40 | 63.68 | −0.112 |
| PHT95521.1 | CamOFP3 | 304 | 33,830.27 | 4.40 | 41.39 | −0.427 |
| PHT90291.1 | CamOFP4 | 421 | 49,484.22 | 8.93 | 63.72 | −1.017 |
| PHT91194.1 | CamOFP5 | 335 | 38,478.01 | 8.92 | 67.91 | −0.985 |
| PHT85943.1 | CamOFP6 | 231 | 26,242.83 | 4.62 | 53.79 | −0.605 |
| PHT88768.1 | CamOFP7 | 281 | 31,602.46 | 8.67 | 45.94 | −0.682 |
| PHT84793.1 | CamOFP8 | 436 | 49,428.19 | 5.99 | 55.79 | −0.356 |
| PHT82131.1 | CamOFP9 | 145 | 16,616.02 | 8.67 | 51.31 | −0.385 |
| PHT82132.1 | CamOFP10 | 320 | 36,192.28 | 4.55 | 46.98 | −0.531 |
| PHT82493.1 | CamOFP11 | 226 | 25,444.77 | 4.68 | 54.56 | −0.318 |
| PHT79195.1 | CamOFP12 | 103 | 11,817.49 | 4.59 | 35.57 | −0.359 |
| PHT79899.1 | CamOFP13 | 130 | 15,152.68 | 8.22 | 31.34 | −0.165 |
| PHT80370.1 | CamOFP14 | 339 | 38,294.99 | 9.70 | 62.54 | −0.810 |
| PHT76741.1 | CamOFP15 | 164 | 18,936.63 | 8.99 | 66.82 | −0.678 |
| PHT74641.1 | CamOFP16 | 337 | 38,375.76 | 9.62 | 51.84 | −0.859 |
| PHT72525.1 | CamOFP17 | 239 | 26,380.44 | 5.37 | 38.63 | −0.308 |
| PHT70189.1 | CamOFP18 | 342 | 38,563.50 | 9.94 | 66.25 | −0.833 |
| PHT71118.1 | CamOFP19 | 190 | 20,883.52 | 4.63 | 39.97 | −0.253 |
| PHT71146.1 | CamOFP20 | 320 | 36,161.30 | 4.83 | 46.56 | −0.529 |
| PHT71148.1 | CamOFP21 | 111 | 12,768.27 | 4.66 | 35.64 | −0.372 |
| PHT71152.1 | CamOFP22 | 165 | 18,376.64 | 5.35 | 50.32 | −0.408 |
| PHT69076.1 | CamOFP23 | 135 | 15,353.64 | 8.61 | 47.65 | −0.748 |
| PHT65371.1 | CamOFP24 | 254 | 28,626.59 | 4.43 | 55.85 | −0.671 |
| PHT62009.1 | CamOFP25 | 120 | 13,567.65 | 4.36 | 50.43 | 0.072 |
| PHT60812.1 | CamOFP26 | 129 | 14,752.58 | 4.54 | 35.48 | −0.457 |

motif 1, motif 2, motif 4, motif 6, motif 7, motif 9 as Class I-B subclade, and *OFP* genes in Group I only contained 4–5 motifs. The 16 *OFP* genes in Class I-A subclade did not contain introns, while the Class I-B subclade had 1-2 exons. The largest number of exons was presented by *CacOFP18*, which contained five exons, and the number of exons in Class II was higher than that in Class I. Sequence comparison of 74 pepper *OFP* genes showed that all *OFP* genes contained motif1 and 69 *OFP* genes contained motif2 with the typical conserved structure of OVATE domain except for *CazOFP24*, *CacOFP2*, *CamOFP4*, *CamOFP8*, and *CamOFP19* (Figure 3). Gene structure analysis showed that the *OFP* genes possessed 1-5 exons (Figure 2D).

## Chromosome mapping and collinearity analysis of *CaOFP* genes

Chromosome mapping analysis showed that 24 *CazOFP* genes were unevenly distributed on 11 chromosomes and Chr00 in Zunla-1, and 24 *CacOFP* genes unevenly on 12 chromosomes in Chiltepin (Figure 4A). Among them, ZChr10 and CChr10 had the highest number of *OFP* genes (8 and 7 genes, respectively), while Chr04, Chr05, Chr07, and Chr012 in Zunla-1 and Chiltepin had only one *OFP* gene each. Interestingly, only ZChr00 contained one *OFP* gene Six genes (*CaOFP15–CaOFP21*) on ZChr10 and CChr10 were tandem repeats with colinear regions as well as other chromosomes in Zunla-1 and Chiltepin, but ZChr08 vs. CChr08 and CChr00 vs. ZChr00 did not have colinear regions (Figure 4A). In another comparison (Figure 4B), most chromosomes had colinear regions except for ZChr00, ZChr08, MChr08, and scaffold2890 in Zunla-1 or CM334.

## Analysis of evolutionary selection pressure of *CaOFPs*

Gene synonymous substitution rate ($K_s$), non-synonymous substitution rate ($K_a$), and their ratio ($K_a/K_s$)

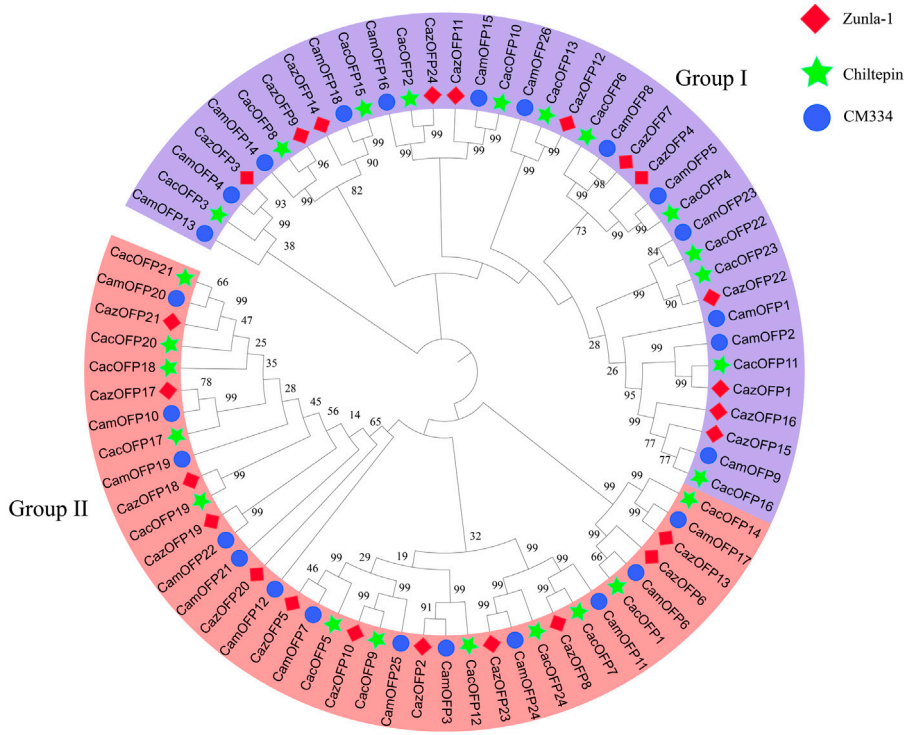



FIGURE 1
Phylogenetic analysis of *OFP* genes in pepper. Caz: Zunla-1; Cac: Chiltepin; Cam: CM334. The solid diamond with red color represented *CazOFP* gene of Zunla-1; the solid star with green color represented *CacOFP* gene of Chiltepin; and the solid circle with blue color represents *CamOFP* gene of CM334.

TABLE 2 Analysis of the evolutionary pressure of selection on the OFP genes in pepper.

| Collateral homologous | Synonymous substitution rate (Ks) | Non-synonymous substitution rate (Ka) | Selective pressure ratio (Ka/Ks) |
|---|---|---|---|
| *CacOFP4-CacOFP8* | 2.3995 | 4.2491 | 1.7709 |
| *CacOFP6-CacOFP4* | 59.1311 | 3.2503 | 0.055 |
| *CacOFP16-CacOFP11* | 2.1397 | 2.8974 | 1.3541 |
| *CacOFP23-CacOFP22* | 0.409 | 0.977 | 2.389 |
| *CazOFP8-CazOFP23* | 2.4928 | 16.504 | 6.6207 |
| *CazOFP17-CazOFP21* | 1.2441 | 1.3557 | 1.0897 |
| *CamOFP11_CamOFP24* | 1.1054 | 6.2469 | 5.6512 |
| *CamOFP10_CamOFP20* | 1.2441 | 1.3557 | 1.0897 |
| *CamOFP14_CamOFP18* | 2.3596 | 2.1842 | 0.9256 |

can reflect the selection pressure in the evolutionary process. The evolutionary selection pressure ($K_a/K_s$) of the seven pairs of paralogous genes identified among 74 *CaOFP* genes were greater than one, suggesting that *OFP* genes were subjected to positive selection. Out of these, nine pairs were under evolutionary selection pressure. The Ka/Ks of two pairs (*CacOFP6-CacOFP4* and *CamOFP14-CamOFP18*) were less than one, indicating that they were under purification selection, while their Ka/Ks of the other seven pairs were more than one, while the other seven pairs were all above 1, indicating that these genes were under positive selection during the evolution of the *CaOFP* genes in pepper (Table 2; Supplementary Table S3).

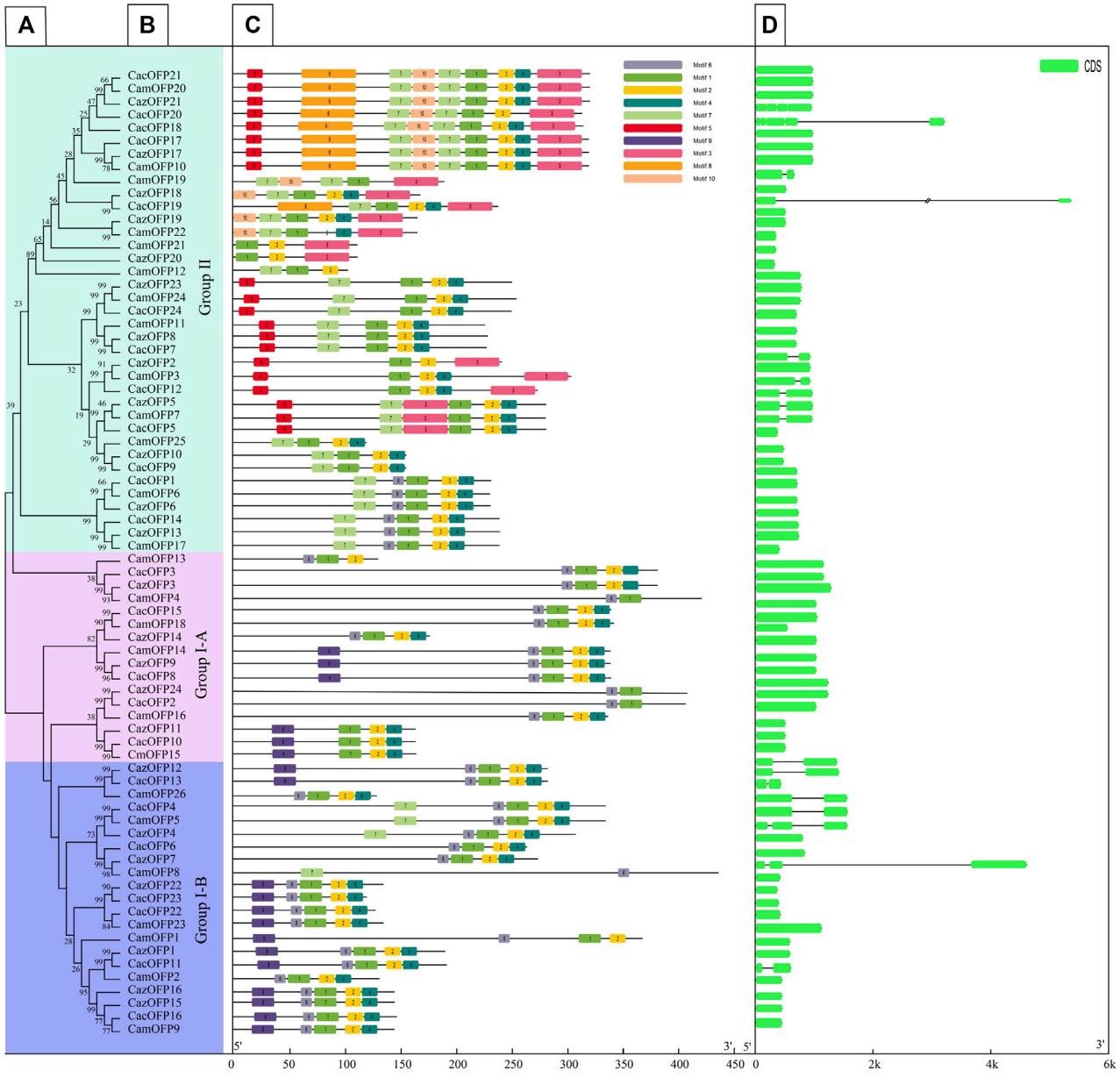



FIGURE 2
Characteristics analysis of all OFP gene family in pepper. [**(A,B)** Phylogenetic tree analysis, **(C)** Gene motifs, **(D)** Gene structure].

## *Cis*-acting element analysis of *CaOFP* genes

The analysis of the 2000-bp upstream sequence of *CaOFP* genes showed that the *OFP* gene family contained numerous basic elements, such as the CAAT box and TATA box (Figure 5), as well as the ABA-related *cis*-acting element ABRE, anaerobic induction-required *cis*-acting element ARE, photoresponse-related conserved element Box 4, photoresponse-related *cis*-acting element G-box, defense and stress response-related *cis*-acting element TC-rich repeats, photoresponse element TCT-motif, cis-acting element involved in salicylic acid responsiveness TCA-element, MYB binding site involved in drought-inducibility MBS and cis-acting element involved in low-temperature responsiveness LTR (Supplementary Table S4). It was

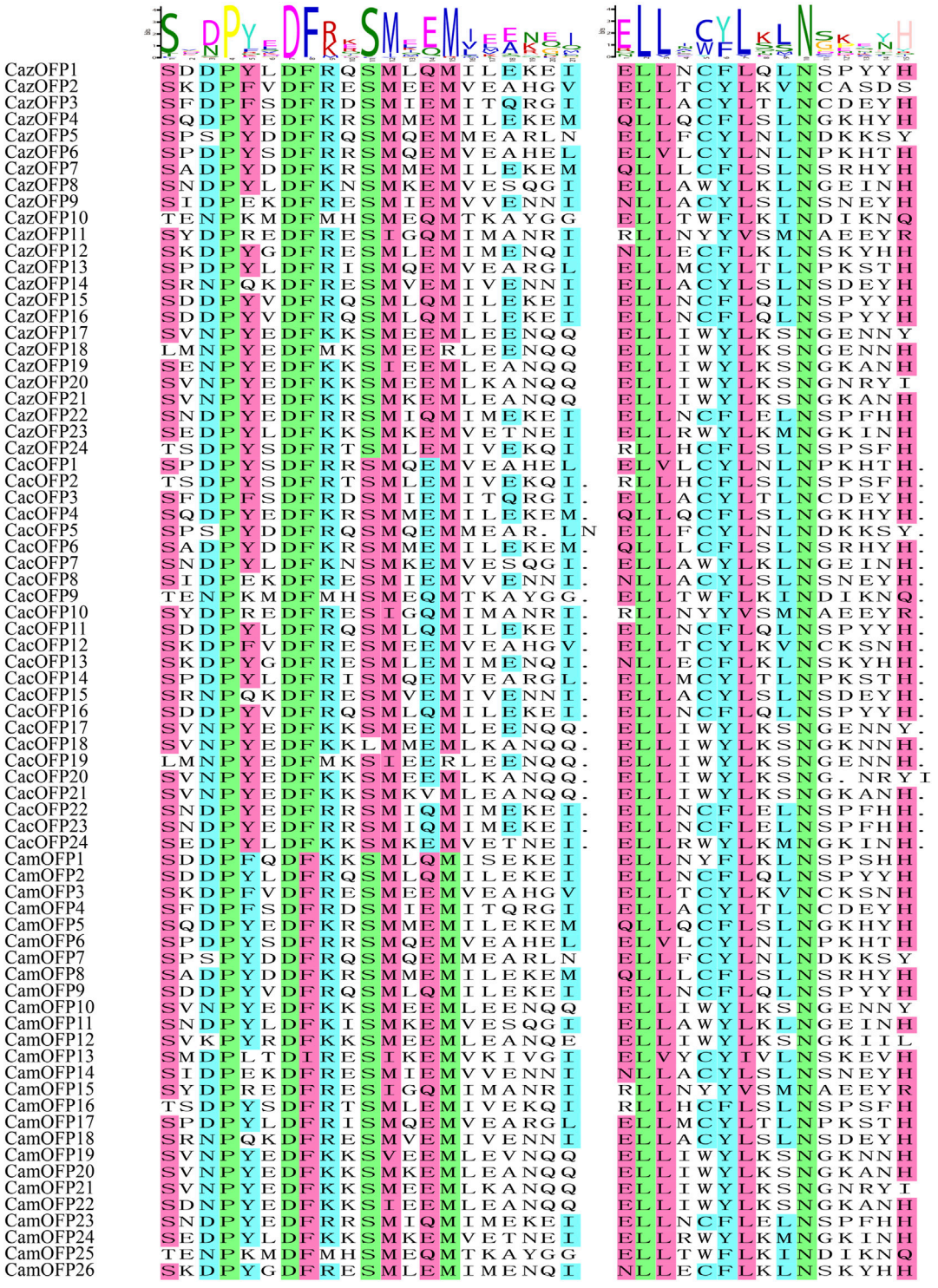



FIGURE 3
Sequence alignment of OFP transcription factor motifs in pepper.

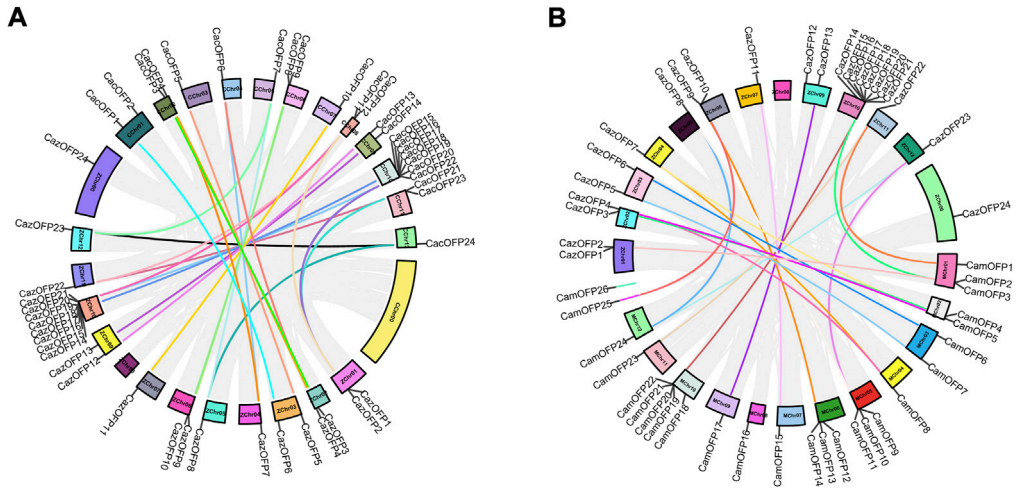



FIGURE 4
Chromosomal localization and collinearity analysis of *OFP* genes in pepper. [**(A)** ZChr: Zunla-1; CChr: Chiltepin. **(B)** ZChr: Zunla-1; MChr: CM334].

speculated that these genes may respond to biological stress and be involve in plant growth.

## Expression patterns of *OFP* genes in different tissues and development stages of pepper

The expression levels of 24 *CazOFP* genes in five tissues including developing roots, developing stems, mature leaves, closed flower buds and open flowers from Zunla-1 pepper data were analyzed. It showed that *CazOFP17~CazOFP21* were not expressed, and other 19 *CazOFP* were expressed in five different tissues on different levels (Figure 6A). During the nine stages of fruit development in Zunla-1, 22 *CazOFP* genes except for *CazOFP18* and *CazOFP20* were expressed in the first eight stages, showing that these *CazOFP* genes were mainly involved in early fruit development. One gene (*CazOFP10*), by contrast, was mainly involved in late fruit development (Figure 6B).

## Characterization of *CazOFP* genes expression patterns under different stresses in pepper

As showed in Figure 7, the expression of *CazOFP16, CazOFP17, and CazOFP18* in pepper leaves was up-regulated upon salicylic acid (SA, 2 mM) treatment. Jasmonic acid (JA, 10 μM) treatment similarly up-regulated the expression of *CazOFP16 and CazOFP17*. Pepper *OFP* genes expression was mainly up-regulated in roots after indole acetic acid (IAA, 2 μM) treatment; the expression of *CazOFP24* in roots was higher than in leaves at a late expression stage. When treated with gibberellic acid ($GA_3$, 2 μM), the expression of *CazOFP16, CazOFP17, and CazOFP21* was higher than that of other *OFP* genes, and the level of up-regulation was greater in roots than in leaves. Abscisic acid (ABA, 30 μM) treatment up-regulated the expression of *CazOFP11 and CazOFP12* in roots at stage AR4, while *CazOFP17* had the highest expression level in leaves at stage AL5. The expression of *CazOFP17* was significantly up-regulated upon salicylic acid (SA, 2 mM), JA, IAA, $GA_3$, and ABA treatments, which suggested that *CazOFP17* participates in plant response to the hormone, whereas *CazOFP18, CazOFP19, and CazOFP20* were not expressed.

The expression of *CazOFP17* was up-regulated in roots under cold stress at stage FR1 (Root/Freezing 1), and the expression of *CazOFP16* was up-regulated in leaves at stage RL3 when treated with $H_2O_2$ (Figure 7). Heat stress up-regulated the expression of *CazOFP16* in roots at stage HR3 and that of *CazOFP17* in leaves at stage HL5 (Figure 7). When treated with D-mannitol, the expression of *CazOFP15* was up-regulated in leaves; the expression of *CazOFPs* was higher in roots than in leaves. Salt stress up-regulated the expression of *CazOFP16* in leaves at stage NL4 (Figure 7). In summary, *CazOFP18-21* were not expressed, and the number of genes expressed in roots was higher than that in leaves.

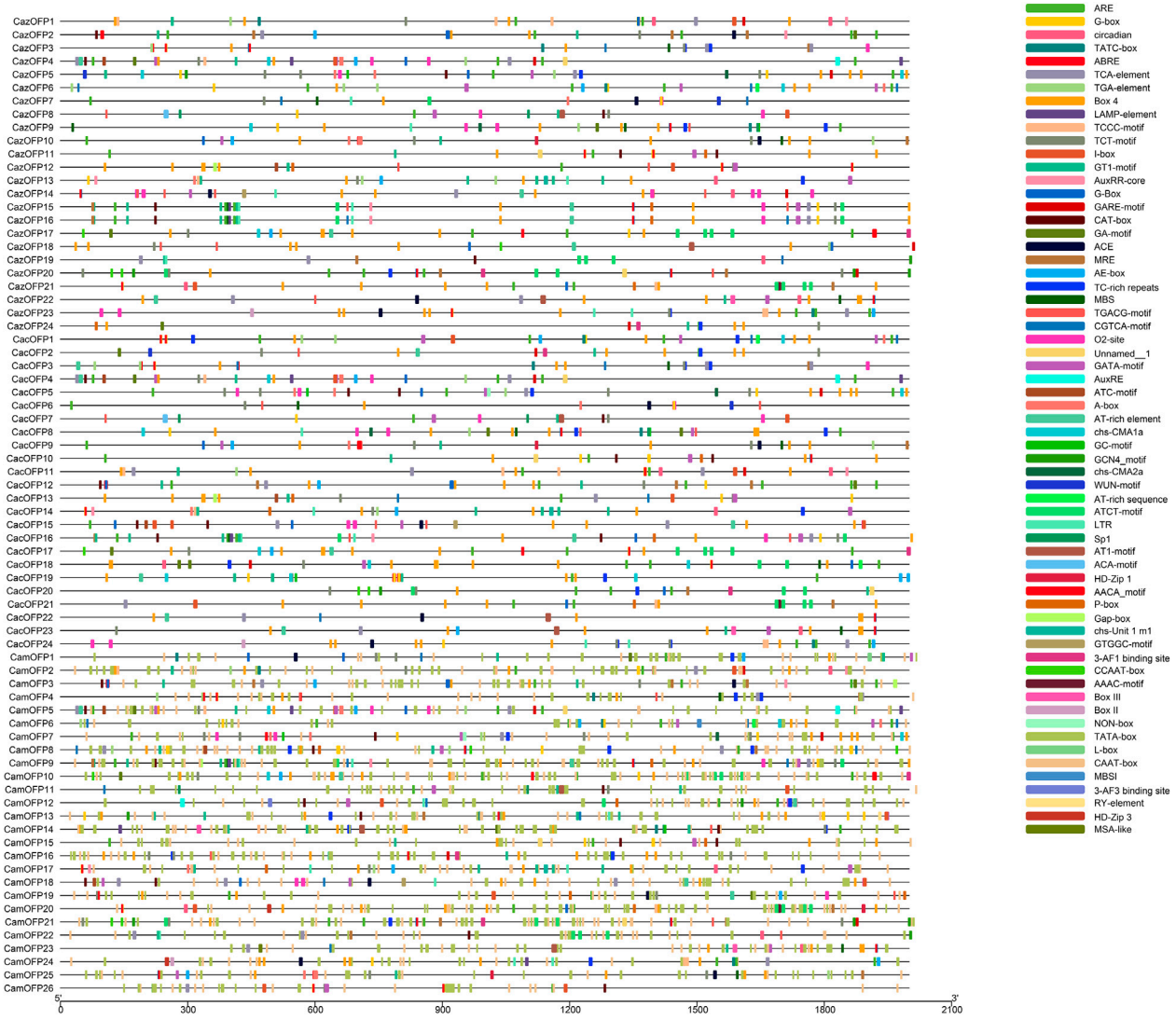



FIGURE 5
Analysis of *Cis*-acting element in the *OFP* gene family of pepper.

## Expression patterns of *OFP* genes in seven pepper varieties with different fruit shapes by qRT-PCR experiments

In order to verify wheather *CaOFP* genes involve in pepper development and fruit shape formation (Figure 8A), eight *CazOFP* genes were randomly selected and qRT-PCR performed. These genes except for *CazOFP12* were mainly expressed in developmental stage 1 (fruit with mature green) of seven pepper varieties with different fruit shapes (Figure 8B), indicating that these genes were mainly involved in early fruit development, which were consistent with transcriptome results mentioned above. In seven pepper varieties with different fruit shapes, the expression of *CazOFP3* and *CazOFP12* at developmental stage 2 (fruit at breaker plus 5 days) in T803 was higher than that at developmental stage 1 (fruit with mature green), as well as *CazOFP14* in Zunla-1. Moreover, *CazOFP3*, *CazOFP9* and *CazOFP14* were expressed in two developmental stages of seven pepper varieties, showing that the three genes may be involved in the formation of pepper fruit shape.

# Discussion

## Characteristics of *OFP* gene family in pepper

We identified 24 *CaOFP* genes in cultivated pepper Zunla-1, 24 in wild pepper Chiltepin, and 26 in another cultivated pepper

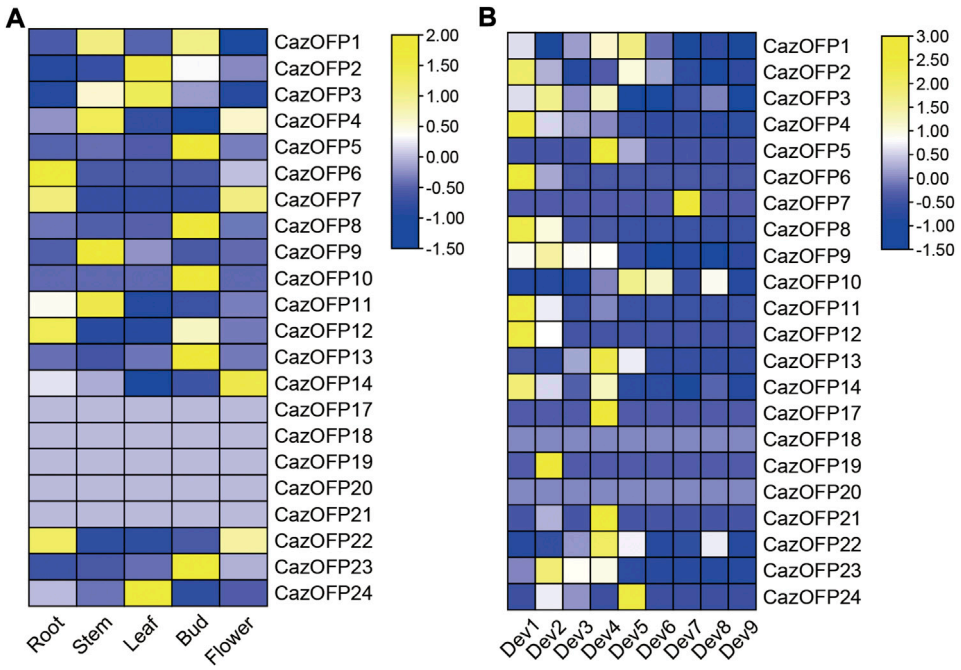



FIGURE 6
Expression pattern of *OFP* genes in pepper.[**(A)** Roots: developing roots; Stems: developing stems; Leaves: mature leaves; Bud: closed flower buds; Flowers: open flowers. **(B)** Dev1: Fruits with length between 0–1 cm; Dev2: Fruits with length between 1–3 cm; Dev3: Fruits with length between 3–4 cm; Dev4: Fruits with length between 4–5 cm; Dev5: mature green fruit; Dev6: Fruit at breaker (fruit turning red); Dev7: Fruit at breaker plus 3 days; Dev8: Fruit at breaker plus 5 days; Dev9: Fruit at Breaker plus 7 days].

CM334. So far, *OFP* genes have been identified in many species with the development of genome sequencing, such as 18 *AtOFP* genes in *A. thaliana* (Wang et al., 2011), 35 *RsOFP* genes in *Raphanus sativus*, 31 *SlOFP* genes in *Solanum lycopersicum* (Huang et al., 2013), 29 *BrOFP* genes in *Brassica rapa* (Wang et al., 2021), and 8 *VvOFP* genes in *Vitis vinifera* (Wang et al., 2018).

In reported studies, the tomato *SlOFP* genes can be divided into three subfamilies (Huang et al., 2013), and the apple *OFP* genes can be divided into 15 subfamilies by carrying out evolutionary analysis of 727 protein sequences from 32 species (Li et al., 2019), while Liu et al. found that *OFP* genes in angiosperms were divided into 11 subfamilies by phylogenetic tree analysis (Liu et al., 2014). In the study, the *CaOFP* genes were divided into two subfamilies, namely Class I and Class II by identifying and classifying three varieties of pepper (Zunla-1, Chiltepin and CM334). However, these *CaOFP* genes can be subdivided according to their conserved domains and motifs. For example, Class I contains only six motifs, while Class II contains motif1, motif5, motif8, and motif10. These results indicated that the subfamily classification of *OFP* genes was quite different according to their plant properties, conserved domains and exon size. In a previous study, *OFP* genes of *Selaginella moellendorffii*, an early terrestrial plant, were found to have more than dozens of exons (Dangwal and Das, 2018), while there were 1-5 exons in the *CaOFP* gene family of pepper, and Class I-A subclade did not contain introns according to its subgroup classification.

Some studies have found that *OFP* was not only involved in the regulation of plant growth and development, but also in plant stress response and fruit morphology formation. Liu et al. cloned OVATE in tomato for the first time and found that it could control the shape of tomato fruit. A point mutation of OVATE gene led to premature termination of translation, which increased the longitudinal diameter of wild tomato fruit and limited the growth of its neck, thus developing from round fruit to pear fruit (Liu et al., 2002). In the yeast two-hybrid assay, nine OVATE proteins were found to interact with TALE homeobox proteins, and AtOFP1 and AtOFP5 co-regulated the subcellular localization of a TALE homeobox protein BLH1. When these two genes were co-expressed in tobacco leaves, The BLH1 protein was transported from the nucleus to the cytoplasm (Hackbusch et al., 2005). Numerous studies have shown that OFP played a role by interacting with KNOX and BELL transcription factors. For example, Schmitz et al. found that OsOFP2 could regulate KNOX-BELL function to participate in the development process of rice (Schmitz et al., 2015). Another study found that the *CmOFP13* gene *fsqs8.1* could alter the recessive structure of the

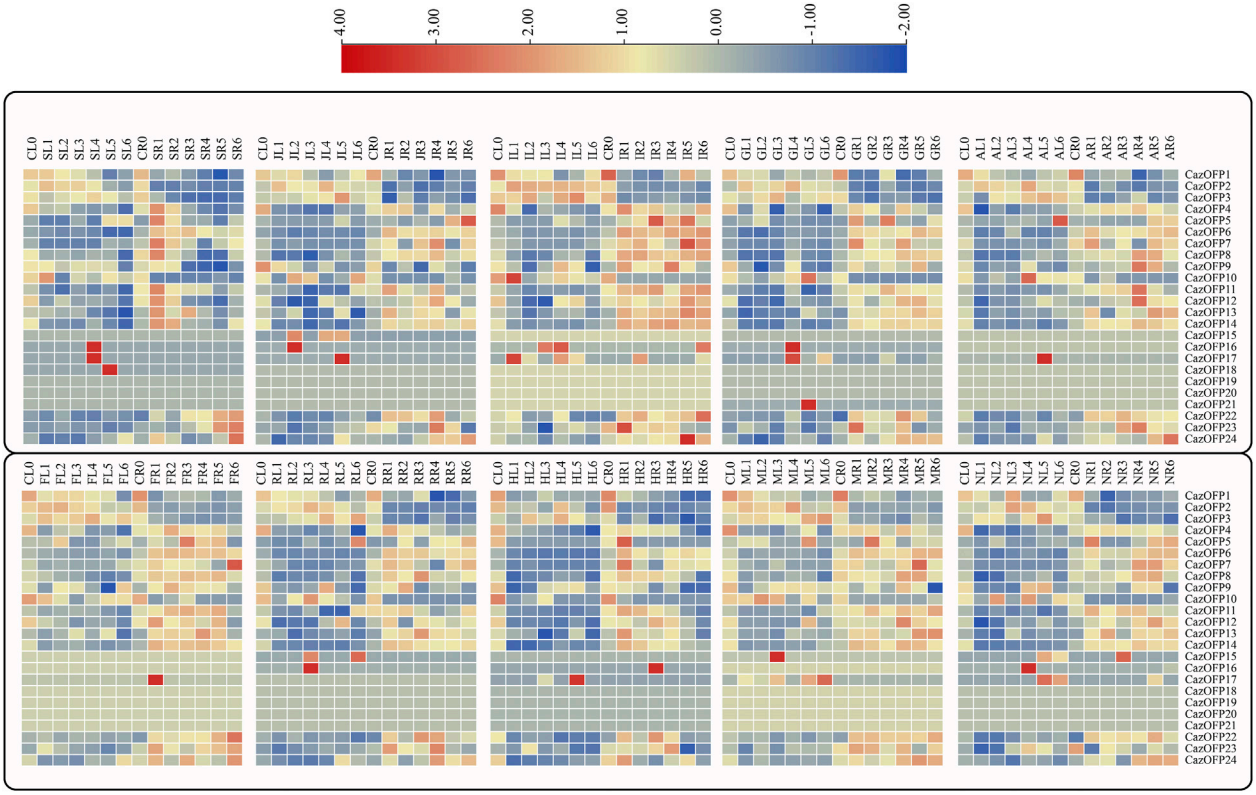



FIGURE 7
Analysis of *CazOFP* gene expression profiles in pepper. Red: up-regulated expression; Gray: not expressive; Blue: down-regulated expression. CL: Control Leaf; SL: SA Leaf; JL:JA Leaf; IL: IAA Leaf; GL: GA3 Leaf; CR: Control Root; SR:SA Root; JR: JA Root; IR: IAA Root; GR: GA3 Root; AL: ABA Leaf; FL: Cold Leaf; RL: H2O2 Leaf; HL: Heat Leaf; ML: Mannitol Leaf; NL: NaCl Leaf; AR: ABA Root; FR: Cold Root; RR: H2O2 Root; HR: Heat Root; MR: Mannitol Root; NR: NaCl Root.

gene in melon, which could lead to structural variation and affect the melon fruit morphology (Martínez et al., 2022). Current biochemical evidence suggests that OFPs may function in this feedback through interactions with three amino acid loop extension (TALE) proteins, as well as through interactions with additional proteins within signaling pathways (Snouffer et al., 2020). Expression patterns of *OFP* genes in different tissues and development stages of pepper and seven pepper varieties with different fruit shapes indicated that some *CaOFP* genes were actively involved in the growth and development of pepper, and participated in plant stress response process and fruit morphological development through mediating mediates.

## Expression profiles of *CaOFP* genes in pepper

Recent studies have found that *OFP* gene was responsive to salt stress in plants. Tang et al. found that *PpOFP1* had salt tolerance through yeast experiments *in vitro* (Tan et al., 2021). Ma et al. found that the *OsOFP6* overexpression lines had slower water loss and less $H_2O_2$ accumulation under drought condition, while the RNAi lines had faster water loss and higher $H_2O_2$ content, which indicated that *OsOFP6* may have drought resistance in rice (Ma et al., 2017). Characterization of *CaOFP* genes expression patterns under different stresses in pepper showed that *CazOFP15*, *CazOFP16* and *CazOFP17* can respond to regulation and participate in the salt stress process. Aabscisic acid responsiveness, salicylic acid responsiveness, defense and stress responsiveness, drought-inducibility, low-temperature responsiveness were also found in the upstream 2000 bp of pepper *CaOFP* genes. Defense and stress responsiveness, drought-inducibility, low-temperature responsiveness, etc. Now studies on OFP mainly focus on growth and development, including defense and stress response, low temperature response and drought induction, while no other studies under SA, JA, IAA, GA3, ABA, $H_2O_2$, Mannitol stress have been reported. According to the analysis of cis-acting elements in pepper, It is speculated that *CazOFP* plays a functional role in responding to relevant elements, which provides new insights for subsequent studies on pepper's response to biotic and abiotic stresses. By studying the expression profile of *CazOFP* genes under different tissues

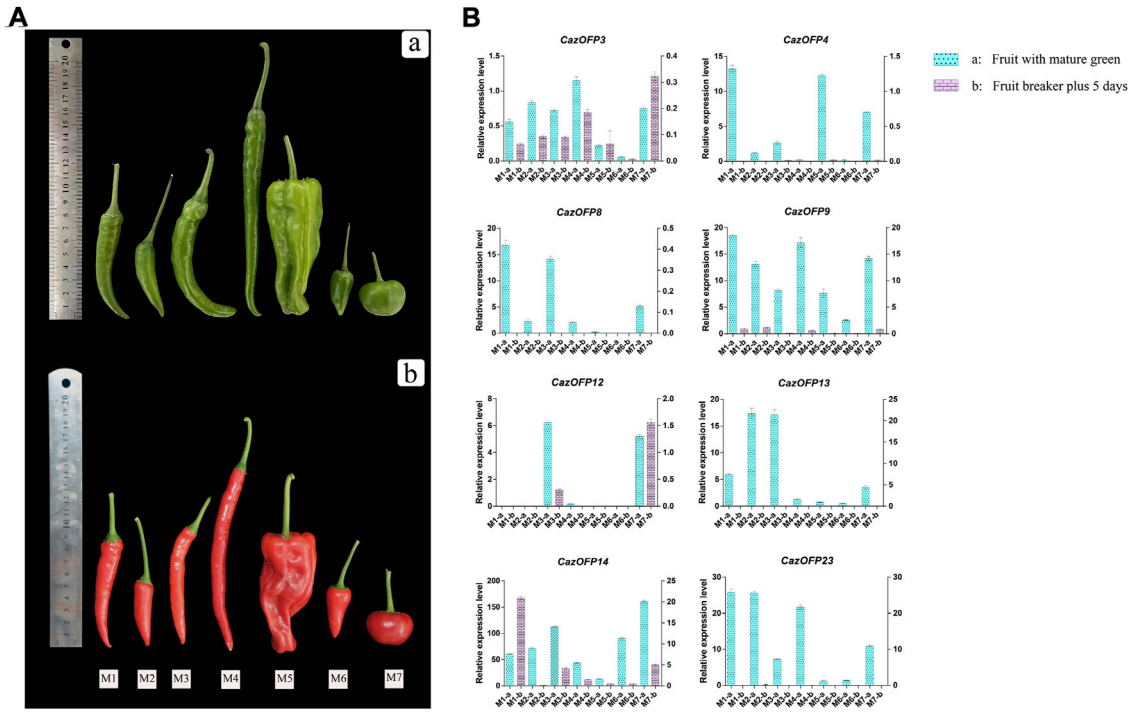



FIGURE 8
Expression of eight *CazOFP* genes in two developmental stages of seven pepper varieties with different fruit shapes. **(A)** Fruit diagrams of seven varieties with different fruit shapes. These varieties (M1-M7) are zunla-1, ZJ-5, 11C255-1, ZJ11, HYL, ZJ-1 and T803. Above is the fruit development stage 1 (fruit with mature green), Below is the fruit development stage 2 (fruit at breaker plus 5 days). **(B)** The levels of these eight *CazOFP* genes detected by qRT-PCR.

and stress conditions, this study provides theoretical support for studying the stress response process of *OFP* and its participation in the regulation of growth and development in pepper.

*OFP*genes have been found to reduce the length of hypocotyls, rosette leaves, stem leaves, inflorescence stems and floral organs in Arabidopsis mutants, while *AtOFP1* regulates the target gene *AtGA20ox1*, resulting in insufficient gibberellin synthesis. These results indicated that OFP transcription factors negatively regulated plant growth and development (Wang et al., 2007). OFP can also promote chlorophyll accumulation and delay leaf senescence. Zhou et al. overexpressed *SlOFP20* in tomato and found that the expressions of genes encoding transcription factors SlGLK1, SlGLK2 and HY5 related to chloroplast development and chlorophyll level were significantly up-regulated, which indicated that OFP had a positive regulatory effect on chlorophyll accumulation and retarding leaf senescence (Zhou et al., 2019). At present, it had been reported that *OFP* was involved in the fruit shape development process of pepper, and CaOVATE was found to be involved in the fruit shape process of tomato through the study of pepper varieties with different fruit shapes (Tsaballa et al., 2011). Expression pattern analysis of *Citrus OFPs* (*CitOFPs*) showed that *CitOFP19* had significantly higher expression level in the ovaries of round pummelo than in those of pear-shaped pummelo (Wu et al., 2022). We also found that the *CazOFP* genes were up-regulated in the early stages of pepper fruit development (Dev1~Dev5) in the study. However, *CazOFPs* expression tended to be down-regulated in the late fruit stages. The expression levels of *CazOFP9* and *CazOFP13* in thin pepper varieties were higher than those in large pepper varieties, suggesting that *CazOFP9* and *CazOFP13* were involved in the fruit shape development of thin pepper varieties. In brief, *CazOFP* genes with different expression levels in pepper may be involved in plant growth and development, thus playing both positive and negative regulatory roles.

Fruit shape is an important quality and yield trait of pepper. High quality pepper and excellent fruit shape help to improve its market competitiveness and increase economic output value. The mechanism and functional verification of OFP transcription factors regulating the growth and development of pepper will become the focus of future research. With the development of plant genome field, it is helpful to explore more *OFP* gene

functions and provide theoretical and practical support for the study of plant shape and morphological development.

## Data availability statement

Publicly available datasets were analyzed in this study. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

## Author contributions

YL, SY, XL, XZ, and CQ conceived and designed the experiments and drafted the manuscript. JL, TL, XT, and FL performed the experiments. YL, SY, and CQ analyzed the data. All authors read and approved the manuscript.

## Funding

This work was supported by the Guizhou Province Science and technology plan program of China (Qian Kehe support Nos. [2021] 1Y088, [2021] yiban267, [2019]2413 and [2022] yiban117), the Zunyi Excellent Young Scientific and Technological Innovative Talents Training Project (Zunyouqingke No. 201806), and the Zunyi Innovative Talent Team Training Project of China (Zunshike rencai No. 201904).

## 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/fgene.2022.941954/full#supplementary-material

## References

Artimo, P., Jonnalagedda, M., Arnold, K., Baratin, D., Csardi, G., De Castro, E., et al. (2012). Expasy: Sib bioinformatics resource portal. *Nucleic Acids Res.* 40 (W1), W597–W603. doi:10.1093/nar/gks400

Bailey, T. L., Johnson, J., Grant, C. E., and Noble, W. S. (2015). The meme suite. *Nucleic Acids Res.* 43 (W1), W39–W49. doi:10.1093/nar/gkv416

Borovsky, Y., Raz, A., Doron-Faigenboim, A., Zemach, H., Karavani, E., and Paran, I. (2021). Pepper fruit elongation is controlled by Capsicum annuum ovate family protein 20. *Front. Plant Sci.* 12, 815589. doi:10.3389/fpls.2021.815589

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

Dangwal, M., and Das, S. (2018). Identification and analysis of OVATE family members from genome of the early land plants provide insights into evolutionary history of OFP family and function. *J. Mol. Evol.* 86, 511–530. doi:10.1007/s00239-018-9863-7

Hackbusch, J., Richter, K., Müller, J., Salamini, F., and Uhrig, J. F. (2005). A central role of *Arabidopsis thaliana* ovate family proteins in networking and subcellular localization of 3-aa loop extension homeodomain proteins. *Proc. Natl. Acad. Sci. U. S. A.* 102 (13), 4908–4912. doi:10.1073/pnas.0501181102

Hu, B., Jin, J., Guo, A., Zhang, H., Luo, J., and Gao, G. (2015). GSDS 2.0: An upgraded gene feature visualization server. *Bioinformatics* 31 (8), 1296–1297. doi:10.1093/bioinformatics/btu817

Huang, Z., Houten, J. V., Gonzalez, G., Xiao, H., and Knaap, E. (2013). Genome-wide identification, phylogeny and expression analysis of SUN, OFP and YABBY gene family in tomato. *Mol. Genet. Genomics* 288 (3), 111–129. doi:10.1007/s00438-013-0733-0

Jiang, T., Lin, Y. X., Liu, X., Jiang, H. Y., and Zhu, S. W. (2011). Genome-wide analysis of the WRKY transcription factor family in Medicago truncatula. *Acta Prat. Sin.* 20 (3), 211–218. doi:10.11686/cyxb20110326

Kim, S., Park, M., Yeom, S. I., Kim, Y. M., Lee, J. M., Lee, H-A., et al. (2014). Genome sequence of the hot pepper provides insights into the evolution of pungency in Capsicum species. *Nat. Genet.* 46, 270–278. doi:10.1038/ng.2877

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33 (7), 1870–1874. doi:10.1093/molbev/msw054

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

Li, E., Wang, S., Liu, Y., Chen, J. G., and Douglas, C. J. (2011). Ovate family protein4 (OFP4) interaction with KNAT7 regulates secondary cell wall formation in *Arabidopsis thaliana*. *Plant J.* 67 (2), 328–341. doi:10.1111/j.1365-313X.2011.04595.x

Li, H., Dong, Q., Zhao, Q., and Ran, K. (2019). Genome-wide identification, expression profiling, and protein-protein interaction properties of ovate family proteins in apple. *Tree Genet. Genomes* 15, 45. doi:10.1007/s11295-019-1354-5

Liu, D., Sun, W., Yuan, Y., Zhang, N., Hayward, A., Liu, Y., et al. (2014). Phylogenetic analyses provide the first insights into the evolution of OVATE family proteins in land plants. *Ann. Bot.* 113 (7), 1219–1233. doi:10.1093/aob/mcu061

Liu, F., Yu, H., Deng, Y., Zheng, J., Liu, M., Ou, L., et al. (2017). Pepperhub, an informatics hub for the chili pepper research community. *Mol. Plant* 10 (8), 1129–1132. doi:10.1016/j.molp.2017.03.005

Liu, J., Van Eck, J., Cong, B., and Tanksley, S. D. (2002). A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. *Proc. Natl. Acad. Sci. U. S. A.* 99 (20), 13302–13306. doi:10.1073/pnas.162485999

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

Ma, Y., Yang, C., He, Y., Tian, Z., and Li, J. (2017). Rice OVATE family protein 6 regulates plant development and confers resistance to drought and cold stresses. *J. Exp. Bot.* 68, 4885–4898. doi:10.1093/jxb/erx309

Martínez, C., Gonzalo, M., Sipowicz, P., Campos, M., Martínez, F. I., Leida, C., et al. (2022). A cryptic variation in a member of the Ovate Family Proteins is underlying the melon fruit shape QTL fsqs8.1. *Theor. Appl. Genet.* 135, 785–801. doi:10.1007/s00122-021-03998-6

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L., et al. (2021). Pfam: The protein families database in 2021. *Nucleic Acids Res.* 49 (D1), D412–D419. doi:10.1093/nar/gkaa913

Paran, I., and Van Der Knaap, E. (2007). Genetic and molecular regulation of fruit and plant domestication traits in tomato and pepper. *J. Exp. Bot.* 58 (14), 3841–3852. doi:10.1093/jxb/erm257

Potter, S. C., Luciani, A., Eddy, S. R., Park, Y., Lopez, R., and Finn, R. D. (2018). Hmmer web server: 2018 update. *Nucleic Acids Res.* 46 (W1), W200-W204–W204. doi:10.1093/nar/gky448

Qin, C., Yu, C., Shen, Y., Fang, X., Chen, L., Min, J., et al. (2014). Whole-genome sequencing of cultivated and wild peppers provides insights into capsicum domestication and specialization. *Proc. Natl. Acad. Sci. U. S. A.* 111 (14), 5135–5140. doi:10.1073/pnas.1400975111

Quinlan, A. R. (2014). BEDTools: The swiss-army tool for genome feature analysis. *Curr. Protoc. Bioinforma.* 47 (1), 11–34. doi:10.1002/0471250953.bi1112s47

Schmitz, A. J., Begcy, K., Sarath, G., and Walia, H. (2015). Rice Ovate Family Protein 2 (OFP2) alters hormonal homeostasis and vasculature development. *Plant Sci.* 241, 177–188. doi:10.1016/j.plantsci.2015.10.011

Snouffer, A., Kraus, C., and Knaap, E. (2020). The shape of things to come: Ovate family proteins regulate plant organ shape. *Curr. Opin. Plant Biol.* 53, 98–105. doi:10.1016/j.pbi.2019.10.005

Suyama, M., Torrents, D., and Bork, P. (2006). PAL2NAL: Robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* 34 (2), W609–W612. doi:10.1093/nar/gkl315

Tan, Q., Jiang, S., Wang, N., Liu, X., Zhang, X., Wen, B., et al. (2021). OVATE family protein PpOFP1 physically interacts with PpZFHD1 and confers salt tolerance to tomato and yeast. *Front. Plant Sci.* 12, 759955. doi:10.3389/fpls.2021.759955

Tsaballa, A., Pasentsis, K., Darzentas, N., and Tsaftaris, A. S. (2011). Multiple evidence for the role of an ovate-like gene in determining fruit shape in pepper. *BMC Plant Biol.* 11 (1), 46–16. doi:10.1186/1471-2229-11-46

Wang, L., Zhang, S., Zhang, X., Hu, X., Guo, C., Wang, X., et al. (2018). Evolutionary and expression analysis of Vitis vinifera OFP gene family. *Plant Syst. Evol.* 304, 995–1008. doi:10.1007/s00606-018-1528-x

Wang, R., Han, T., Sun, J., Xu, L., Liu, C., Cao, H., et al. (2021). Genome-wide identification and characterization of the OFP gene family in Chinese cabbage ( Brassica rapa L. ssp. pekinensis ). *PeerJ* 9, e10934. doi:10.7717/peerj.10934

Wang, S., Chang, Y., and Ellis, B. (2016). Overview of ovate family proteins, a novel class of plant-specific growth regulators. *Front. Plant Sci.* 7, 417. doi:10.3389/fpls.2016.00417

Wang, S., Chang, Y., Guo, J., and Chen, J. G. (2007). Arabidopsis Ovate Family Protein 1 is a transcriptional repressor that suppresses cell elongation. *Plant J.* 50, 858–872. doi:10.1111/j.1365-313X.2007.03096.x

Wang, S., Chang, Y., Guo, J., Zeng, Q., Ellis, B. E., and Chen, J. G. (2011). Arabidopsis ovate family proteins, a novel transcriptional repressor family, control multiple aspects of plant growth and development. *Plos One* 6, e23896. doi:10.1371/journal.pone.0023896

Wang, Y., Li, J., and Paterson, A. H. (2013). MCScanX-transposed: Detecting transposed gene duplications based on multiple colinearity scans. *Bioinformatics* 29 (11), 1458–1460. doi:10.1093/bioinformatics/btt150

Wu, Q., Sun, u., Fu, J., Yu, H., Wang, X., Wang, S., et al. (2022). Genome-wide identification of ovate family in Citrus and functional characterization of CitOFP19. *Plant Sci.* 321, 111328. doi:10.1016/j.plantsci.2022.111328

Xu, R., Li, R., Wang, X., and Hao, Y. (2018). Identification and expression analysis under abiotic stresses of ofp gene family in apple. *Sci. Agr. Sin.* 51 (10), 1948–1959. doi:10.3864/j.issn.0578-1752.2018.10.014

Yuan, Y., Zhang, Y. G., Gao, S. M., and Tao, J. M. (2016). Bioinformatics and expression of the ovate gene family in grape. *Sci. Agr. Sin.* 49 (19), 3786–3797. doi:10.3864/j.issn.0578-1752.2016.19.010

Zhou, S., Cheng, X., Li, F., Feng, P., Hu, G., Chen, G., et al. (2019). Overexpression of SlOFP20 in tomato affects plant growth, chlorophyll accumulation, and leaf senescence. *Front. Plant Sci.* 10, 1510. doi:10.3389/fpls.2019.01510

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 30 September 2022
DOI 10.3389/fgene.2022.988264



OPEN ACCESS

EDITED BY
Sundeep Kumar,
National Bureau of Plant Genetic Resources (ICAR), India

REVIEWED BY
Suraj Sapkota,
United States Department of Agriculture, United States
Spurthi N. Nayak,
University of Agricultural Sciences, India

*CORRESPONDENCE
Xu Zhang,
xuzhang@jaas.ac.cn
Pawan K. Singh,
pk.singh@cgiar.org

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

RECEIVED 07 July 2022
ACCEPTED 13 September 2022
PUBLISHED 30 September 2022

CITATION
Wu L, He X, He Y, Jiang P, Xu K, Zhang X and Singh PK (2022), Genetic sources and loci for Fusarium head blight resistance in bread wheat. *Front. Genet.* 13:988264. doi: 10.3389/fgene.2022.988264



# Genetic sources and loci for Fusarium head blight resistance in bread wheat

Lei Wu[1], Xinyao He[2], Yi He[1], Peng Jiang[1], Kaijie Xu[3], Xu Zhang[1]* and Pawan K. Singh[2]*

[1]CIMMYT-JAAS Joint Center for Wheat Diseases, Jiangsu Academy of Agricultural Sciences, Nanjing, China, [2]International Maize and Wheat Improvement Center (CIMMYT), Texcoco, Mexico, [3]Institute of Cotton Research, Chinese Academy of Agricultural Sciences, Anyang, China

Fusarium head blight (FHB) of wheat is an important disease worldwide, affecting the yield, end-use quality and threatening food safety. Genetic resources or stable loci for FHB resistance are still limited in breeding programs. A panel of 265 bread wheat accessions from China, CIMMYT-Mexico and other countries was screened for FHB resistance under 5 field experiments in Mexico and China, and a genome-wide association analysis was performed to identify QTLs associated with FHB resistance. The major locus *Fhb1* was significantly associated with FHB severity and Deoxynivalenol content in grains. FHB screening experiments in multiple environments showed that *Fhb1*-harbouring accessions Sumai3, Sumai5, Ningmai9, Yangmai18 and Tokai66 had low FHB index, disease severity and DON content in grains in response to different *Fusarium* species and ecological conditions in Mexico and China. Accessions Klein Don Enrique, Chuko and Yumai34 did not have *Fhb1* but still showed good FHB resistance and low mycotoxin accumulation. Sixteen loci associated with FHB resistance or DON content in grains were identified on chromosomes 1A, 1B, 2B, 3A, 3D, 4B, 4D, 5A, 5B, 7A, and 7B in multiple environments, explaining phenotypic variation of 4.43–10.49%. The sources with good FHB resistance reported here could be used in breeding programs for resistance improvement in Mexico and China, and the significant loci could be further studied and introgressed for resistance improvement against FHB and mycotoxin accumulation in grains.

KEYWORDS

**Fusarium head blight, resistance, genetic sources, genome-wide association study, deoxynivalenol**

## Introduction

Fusarium head blight (FHB) is one of the most devastating diseases of wheat and other small grain cereals, threatening global wheat production and food security and safety (Bai and Shaner, 2004; Ma et al., 2020). Many *Fusarium* spp. can infect wheat head, of which *F. graminearum* species complex (FGSC) is the predominant FHB pathogen in most wheat production areas (Van Der Lee et al., 2015). The distribution of various FGSC sub-species depends on geographies, climatic conditions and cropping systems, e.g., *F. asiaticum* is

mainly in East Asia, while *F. meridionale* and *F. boothii* are mainly in South America and Africa (Zhang et al., 2012; Backhouse, 2014; Van Der Lee et al., 2015). FHB leads to severe yield loss, poor grain quality, and more importantly contamination of the infected grains with mycotoxins like deoxynivalenol (DON) or nivalenol, for which public concerns have prompted governments to set upper limits for DON in wheat grain and its products (Mesterházy et al., 1999; He et al., 2016; Ma et al., 2020).

Breeding for FHB resistance by using QTL/genes in genetic sources is one of the effective approaches to control this disease and prevent toxins contamination in grains after harvest. FHB resistance in wheat is a complex quantitative trait with strong genotype-by-environment interactions, resulting in different response to *Fusarium* pathogens across different environments (Mesterhazy, 2020). Host resistance to FHB involves five resistant types, i.e. Type I for initial infection, Type II for disease spread, Type III for toxin accumulation, Type IV for kernel infection, and Type V for yield reduction (Schroeder and Christensen, 1963; Miller and Amison, 1986; Mesterhazy, 1995). Type II and III resistance have been widely studied (Buerstmayr et al., 2020; Ma et al., 2020; Mesterhazy, 2020), and over 600 loci for Type II or III resistance have been mapped on all 21 wheat chromosomes (Venske et al., 2019; Zheng et al., 2021). Of the nominated *Fhb* genes, *Fhb1* (Cuthbert et al., 2006; Liu et al., 2006), *Fhb2* (Cuthbert et al., 2007), *Fhb4* (Xue et al., 2010) and *Fhb5* (syn. *Qfhs.ifa-5A*) (Xue et al., 2011; Buerstmayr et al., 2018; Steiner et al., 2019) are derived from common wheat, of which the former two mainly confer Type II resistance and the latter two mainly confer Type I resistance. The wild relatives of wheat have also contributed several resistance genes/loci, like *Fhb3* (Qi et al., 2008) from *Leymus racemosus*, *Fhb6* (Cainong et al., 2015) from *Elymus tsukushiensis* and *Fhb7* (Guo et al., 2015) from *Thinopyrum elongatum*, all conferring Type II resistance. Type II and Type III resistance are usually highly correlated in field experiments as observed by most researchers, although QTL exclusively for Type III resistance have been reported (He et al., 2019).

Thousands of wheat accessions have been screened for FHB resistance throughout the world over the last decades, but resistant sources for FHB improvement in breeding programs remain limited. *Fhb1* (Syn. *Qfhs.ndsu-3BS*) derived from Sumai3 or its derivative Ning7840 is recognized as a locus with major effect and stable resistance and has been widely used in wheat breeding programs worldwide (Zhu et al., 2019; Buerstmayr et al., 2020; Hao et al., 2020; Ma et al., 2020; Gaire et al., 2022b). Several research groups have cloned the candidate genes of *Fhb1* (Rawat et al., 2016; Li et al., 2019; Su et al., 2019) and designed diagnostic markers for marker-assisted selection (MAS) (He et al., 2018; Su et al., 2018; Singh et al., 2019). Sumai3 that has *Fhb1*, *Fhb2*, and *Fhb5* is a widely used donor for FHB resistance improvement, and more than 20 released spring wheat varieties have Sumai3 in their pedigrees in North America and Canada (Zhu et al., 2019; Ghimire et al., 2020). The donor parent of *Fhb1* in the Chinese wheat breeding programs is, however, not Sumai3 but Ningmai9 (developed by Jiangsu Academy of Agriculture Sciences, JAAS) because of its better agronomic traits and high yield potential (Zhu et al., 2018). More recently, *Fhb7* derived from *Th. elongatum* was reported to confer broad resistance to *Fusarium* species by detoxifying DON without yield penalty (Wang et al., 2020), and several locally adaptive lines with *Fhb7* in their pedigrees have been tested in Regional Yield Trials in China (Prof. Hongwei Wang, Shandong Agricultural University, personal communication).

Despite the achievement, highly resistant sources and major loci for FHB resistance are scarce for wheat breeding, it is therefore worthwhile to identify additional FHB resistance sources and loci for gene pyramiding. The objectives of this study were to screen a collection of common wheat varieties and elite breeding lines of worldwide origin for FHB resistance in Mexico and China, and to conduct a genome-wide association study (GWAS) to identify QTL for FHB resistance in field experiments.

## Materials and methods

### Plant materials

In this study, 265 wheat accessions (CIMMYT-China panel) were screened for FHB resistance under field conditions in Mexico and China. The panel included commercial varieties and breeding lines of worldwide origin, i.e., 131 Chinese accessions mainly from the Yellow and Huai River Valley Region and Middle-lower Yangtze Valley Region, 71 from CIMMYT-Mexico, 41 from South America, 10 from North America, five each from Asia and Europe, and one each from Oceania and Africa (Supplementary Table S1). The accessions were mostly of spring type along with a few of winter type. In Mexico, Sumai3 and Heilo were used as resistant checks, while Gamenya and Ocoroni were included as susceptible checks. In China, Sumai3 and Yangmai158 served as resistant checks with Annong8455 and Jimai22 as susceptible checks.

### Field trials and Fusarium head blight screening

The CIMMYT-China panel was evaluated for FHB resistance in CIMMYT's El Batan research station (altitude of 2,240 masl, 19.5ºN/98.8ºW, with an average annual precipitation of 625 mm) in Mexico, and JAAS's FHB nursery (altitude of 22 masl, 32.0ºN/118.8ºE, with an annual precipitation of 800–1034 mm) in Nanjing, China. In each location, the accessions were planted in 1-m double rows with two replications in randomized complete block design. The trials were conducted in the

2018 and 2019 cropping cycles (from May to September) in Mexico, and in the 2018–19, 2019–20, and 2020–21 cropping cycles (i.e. November to May) in China.

In Mexico, inoculum comprised a mixture of five *F. graminearum* isolates, CIMFU No. 85, 89, 108, 162, and 222, following the protocols described by (He et al., 2013). These isolates were obtained from naturally infected wheat heads in the El Batan station in 2017 and were selected based on their high DON productivity and high aggressiveness in greenhouse experiments. At anthesis, the field plots were sprayed with an inoculum of 50,000 spores/ml, and the procedure was repeated 2 days later to reinforce the infection. A misting system was set up in the nursery, operational from 9 am to 8 pm with 10 min of spraying each hour, to maintain a humid environment conducive for FHB infection. Field evaluation of FHB infection was carried out at 25 days post-inoculation (dpi) on the 10 spikes tagged at anthesis (Feekes 10.5.1). Numbers of total and infected spikelets of each spike were scored for the calculation of FHB index with the formula: FHB index = severity × incidence, where severity means the averaged percentage of diseased spikelets, and incidence the percentage of symptomatic spikes. About 20g grain sample was ground with a coffee mill and a 2g sub-sample was measured for DON quantification with the Ridascreen Fast DON ELISA kit (RBiopharm GmbH, Darmstadt, Germany) following the manufacturer's instructions.

In China, *F. asiaticum* isolates Fa0609, Fa1312 and Fa0980 were used in this study. These isolates were isolated from different wheat plots in JAAS (Nanjing, China) in 2006, 2009 and 2013, respectively, and had shown strong virulence in greenhouse and field experiments. The fungal isolates were applied to wheat kernels to produce spawn inoculum, which was distributed on the soil surface with a density of 30 g/m$^2$ on the 20th day before flowering and the inoculation was repeated 2 weeks later. All the plots were misted for 15 min per hour during the day to create a humid environment for fungal development and spread. The numbers of diseased spikelets of 10 random spikes were recorded to calculate the disease severity for each accession at the milky ripe stage (Feekes 11.2). The spawn inoculation experiments were performed only during 2018–19 and 2019–20 cropping cycles. Besides, point inoculation with a single isolate Fa0609 was also conducted to evaluate the Type II resistance of the panel during 2018–19, 2019–20, and 2020–21 cropping cycles. Ten spikes per accession were inoculated via injecting 10 µl suspension of *F. asiaticum* (100,000 spores/ml) into a central spikelet at anthesis (Feekes 10.5.1), and then the spikes were covered with a plastic bag for 3 days to meet the moisture requirement for fungal infection. The moisturizing measure was the same as that mentioned above. The numbers of diseased spikelets were recorded as FHB severity on the 21st day after inoculation. The inoculated spikes were harvested to detect DON content with DON ELISA test kit (Huaan Magnech Bio-tech, China) using an indirect competitive enzyme-labeled immunoassay.

The phenotypic dataset including the mean values of FHB index, FHB severity and DON content in grains in each environment for each accession were used in subsequent analysis. The datasets of Mexico and China were calculated separately. Pearson's correlation coefficients among environments were computed by R to test the correlation of FHB responses across different environments (R Team, 2011). Turkey's mean comparison tests were performed among subgroups using R (R Team, 2011).

## Genotyping

The GWAS panel was genotyped with the DArTSeq technology at the Genetic Analysis Service for Agriculture (SAGA) at CIMMYT, Mexico. Markers with more than 10% missing data or minor allele frequency less than 1% were eliminated, resulting in 18,436 high quality markers, of which 14,195 SNPs with known physical positions in Chinese Spring reference genome v.1.0 were used for population structure and LD analysis. PCR-based marker JAASM395 was used to test whether an accession carried *Fhb1* (Chinese patent, ZL201811515195.8), and the primer pair was as follows: JAASM395F: GTCTCCGTTCAA TTCGGTGAGT; JAASM395R: GACAATGTGAAGGCGTTGTCT A. Analysis on population structure and linkage disequilibrium followed the method in (Wu et al., 2021). Briefly, a Bayesian model-based method was used to infer the number of subpopulation among all accessions with the software STRUCTURE v 2.3.4 (Pritchard et al., 2000). Five independent analyses for the assumed number of subpopulations (K value) from 1 to 8 were performed based on an admixture model with MCMC replications and burn-in time number set at $1\times10^5$ and $1\times10^4$, respectively (Pritchard et al., 2000). Optimal value of population size was inferred by CLUMPP based on the data of STRUCTURE (Jakobsson and Rosenberg, 2007). Pairwise LD ($r^2$) for all pairwise comparisons in a distance of 10,000 kb was estimated by the software PLINK (Chang et al., 2015) and were plotted against the physical distances, and then a nonlinear regression was fitted in R (R Team, 2011). The critical $r^2$ value was determined as the 95th quantile for all $r^2$ values between unlinked SNPs. The intersection between the critical $r^2$ value and the regression line was used to estimate the average size for LD blocks in this panel.

## Genome-wide association analysis

Genome-wide association for FHB resistance across environments was performed using software TASSEL v5.0 pipeline command line interface (Bradbury et al., 2007). Mixed linear models (MLM) with K/Q and K/P matrices as covariates were chosen for all GWAS analyses. The Q matrix was calculated and estimated by STRUCTURE v2.3.4 (Pritchard et al.,

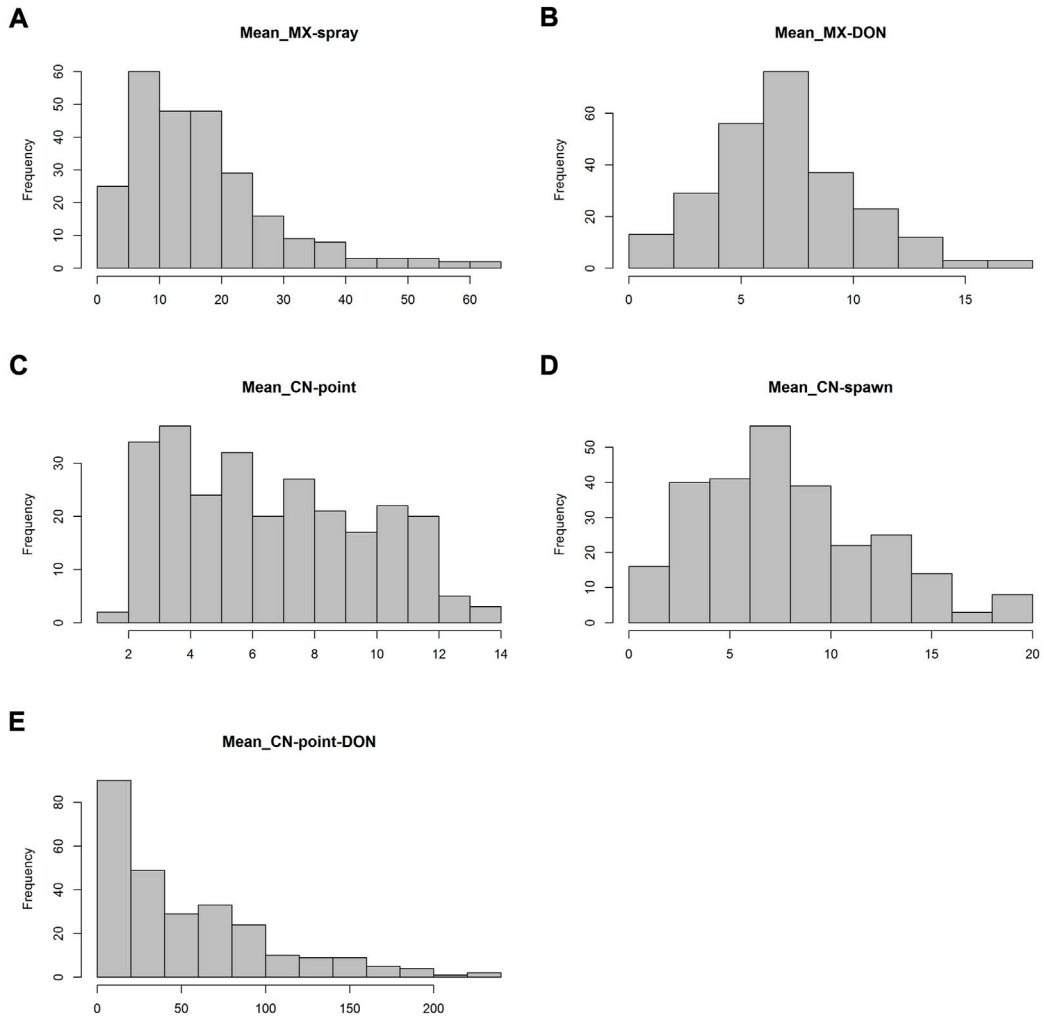



FIGURE 1
Frequency distribution of FHB index **(A)**, number of diseased spikelets **(C and D)** and DON content in grains **(B and E)** based on the mean data in Mexico (MX, spray inoculation) and China (CN, point and spawn inoculation).

2000). Physical positions of significant SNPs were determined by sequence alignments with "Chinese Spring" reference genome v1.0 on the website of EnsemblPlants (http://plants.ensembl.org/index.html) using BLAST program with default parameters. Physical positions of significant MTAs were compared to meta-QTL reported by (Zheng et al., 2021) to see their novelty.

# Results

## Fusarium head blight screening in Mexico and China

In spray inoculation experiments in Mexico, FHB index showed a skewed distribution toward the low disease direction and the panel exhibited a grand mean FHB index of 16.83%. The accession Ning894013 had the lowest FHB index of 0.97%, while Norseman had the maximum of FHB index of 62.87% (Figure 1A, Supplementary Table S1). DON content in grains showed a continuous distribution with a mean of 6.91 μg/g (Figure 1B, Supplementary Table S1). In point inoculation experiments in China, numbers of diseased spikelets evenly distributed ranging from 2 to 13 with a mean of 6.6 (Figure 1C, Supplementary Table S1). In spawn inoculation experiments, the mean numbers of diseased spikelets were higher than that in point inoculation experiments (Figure 1D). The DON content in grains in point inoculation experiments was skewed to the low content direction, though, the mean value 54.11 μg/g was much higher than that in spray inoculation experiments (Figure 1E, Supplementary Table S1).

TABLE 1 Pearson's correlation coefficients among experiments in Mexico (MX) and China (CN).

| | MX.spray | MX.DON | CN.point | CN.spawn |
|---|---|---|---|---|
| MX.DON | 0.32* | | | |
| CN.point | 0.09 | 0.44* | | |
| CN.spawn | 0.44* | 0.47* | 0.34* | |
| CN.pointDON | 0.07 | 0.45* | 0.79* | 0.33* |

Mean values across years were used here for FHB, index (spray inoculation in Mexico), number of diseased spikelets (point and spawn inoculation in China), DON, content in Mexico and China. * indicates significant correlations at $p < 0.001$.

Phenotypic correlation varied greatly among different experiments, from low non-significant correlation (r = 0.07) to high significant correlation (r = 0.79). Generally, low to moderate correlations were observed among different inoculation methods, and the highest correlation was found between number of diseased spikelets and DON content in point inoculation experiments in China (Table 1).

Genetic resources with moderate resistance to FHB and low DON content in grains were listed in Table 2. As expected, *Fhb1*-carriers Sumai3, Sumai5, Ningmai9, Yangmai18 and Tokai66 exhibited good resistance in all experiments in Mexico and China. Accessions Klein Don Enrique, Chuko and Yumai34 that do not have *Fhb1* still showed good FHB resistance and low toxin accumulation.

Population structure analysis indicated that the panel could be divided into two subpopulations, which could be further divided into five groups 1A, 1B, 2A, 2B and 2C. The clustering of subpopulations mainly associated with geographical origins. Subpopulation 1 (130 accessions) involved accessions mainly from CIMMYT, South America, North America, and Europe, while most accessions in subpopulation 2 (135 accessions) were from Asian countries including China, Japan and India (Supplementary Table S1).

There were significant differences in FHB resistance or DON content in grains among the groups 1A, 1B, 2A, 2B and 2C ($p < 0.01$) (Figures 2A–E). In spray inoculation experiments in Mexico, group 2A had lower mean FHB index than 1B and 2C, and 2B had the lowest DON content in grains (Figures 2A,B). In the experiments in China, however, subpopulation 2 outperformed subpopulation 1 significantly in most cases, and significant differences were also found within subpopulations (Figures 2C–E).

Significant differences in FHB index or number of diseased spikelets were observed between *Fhb1* and non-*Fhb1* groups, where the former outperformed the latter; the same trend applied to DON content in both Mexico and China (Figure 3).

## Loci associated with Fusarium head blight resistance

Marker-trait associations (MTAs) were tested separately for the dataset of mean value in Mexico and China, and a total of 16 MTAs (24 SNPs) were significantly associated with FHB resistance or DON content in grains (Figures 4, 5), explaining phenotypic variation between 4.43% and 10.49% (Table 3). Four of the SNPs located in annotated gene regions without known functions. Ten MTAs (18 SNPs) on chromosomes 1A, 1B, 2B, 7A, 4D, 4B, 5A and 7B were significantly associated with FHB resistance ($p < 0.001$), of which markers 3026949, 979146, 1157139 and 10334520 were significantly associated with both FHB resistance and DON content using point inoculation data in China. Marker

TABLE 2 Top performers in FHB and DON traits across experiments.

| Name | Origin | Group | *Fhb1* | FHB index (%) | Number of disease spikelets | | DON content (μg/g) | |
|---|---|---|---|---|---|---|---|---|
| | | | | Spray inoculation | Point inoculation | Spawn inoculation | Point inoculation | Spray inoculation |
| Klein Don Enrique | Argentina | 2A | No | 3.34 | 3.33 | 1.60 | 9.79 | 2.75 |
| Chuko | Japan | 2B | No | 5.72 | 3.88 | 2.13 | 6.94 | 2.45 |
| Tokai66 | Japan | 2B | Yes | 5.79 | 3.46 | 2.22 | 6.12 | 1.05 |
| Sumai3 | China | 2C | Yes | 7.21 | 3.48 | 1.13 | 11.33 | 1.24 |
| 7P3 | China | 2C | Yes | 1.14 | 2.85 | 2.29 | 4.99 | 0.74 |
| Ningmai18 | China | 2B | Yes | 7.91 | 3.80 | 3.50 | 14.42 | 3.68 |
| Ningmai9 | China | 2C | Yes | 5.25 | 2.30 | 2.39 | 6.07 | 3.80 |
| Ningyan1 | China | 2C | Yes | 1.89 | 2.10 | 1.58 | 3.55 | 7.48a |
| Sumai5 | China | 2C | Yes | 7.57 | 2.90 | 3.50 | 24.18a | 5.35a |
| Yangmai18 | China | 2B | Yes | 5.73 | 2.25 | 7.24a | 8.82 | 3.68 |
| Yumai34 | China | 2C | No | 6.88 | 2.65 | 2.50 | 26.42a | 4.60 |

Numbers marked with "a" rank between the first and second quantiles in their own datasets.

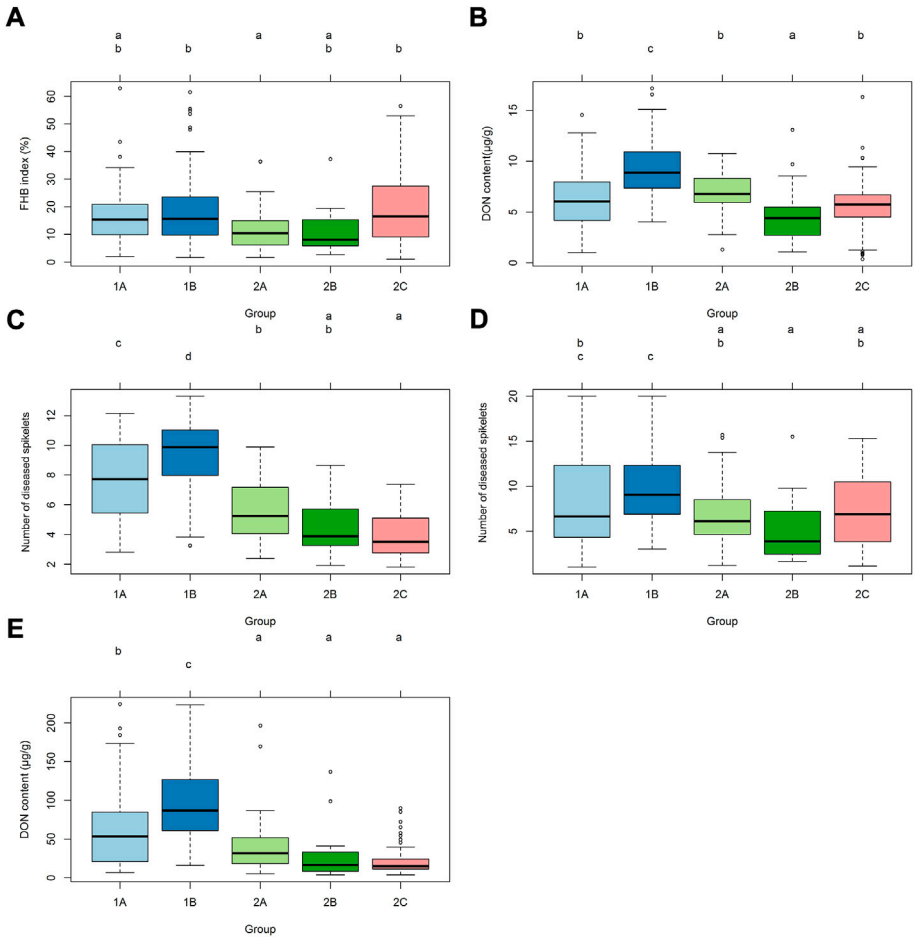



FIGURE 2
Distributions of FHB index, number of diseased spikelets and DON content among different groups. **(A)** Distribution of mean data of FHB index using spray inoculation among different groups in 2018 and 2019 in Mexico. **(B)** Distribution of mean data of DON content using spray inoculation among different groups in 2018 and 2019 in Mexico. **(C)** Distribution of mean data of FHB severity using point inoculation among different groups in 2018, 2019 and 2020 in China. **(D)** Distribution of mean data of FHB severity using spawn inoculation among different groups in 2018 and 2019 in China. **(E)** Distribution of mean data of DON content in grains using point inoculation among different groups in 2018, 2019 and 2020 in China.

1044062 on 4B chromosome was identified in association with FHB resistance in China and DON content in Mexico. Only one marker 1099971 was significantly associated with FHB resistance in both China and Mexico, overlapping a reported mQTL sMQTL-1A-5. Six markers on chromosomes 2B, 3A, 3D, 4D, 5A and 5B were associated with only DON content across China and Mexico (Table 3). Significant markers for FHB resistance or DON content in individual environments were listed in Supplementary Table S2, explaining phenotypic variation of 4.34–11.33%.

## Discussion

FHB of wheat is a serious disease in the temperate and humid regions around the world. Screening for resistant sources is a prerequisite for the improvement of FHB resistance, but many factors affect the development of locally adapted resistant varieties, including complex inheritance, multiple resistance types, difficulties on precise phenotyping, association of FHB resistance with undesirable traits (late and tall plant phenology etc.), and strong genotype-by-environment interaction. In the present study, 265 accessions from China, CIMMYT-Mexico and other countries were screened for FHB resistance in Mexico and China. FHB index and disease severity were used to evaluate FHB resistance of all the accessions in field trials, and then DON content in infected grains was detected. Spray inoculation simulates the process of disease development under natural conditions in Mexico, and the FHB index showed a combination of Types I and II resistance to FHB. Therefore, the data of FHB incidence does not refer to strictly Type I resistance due to the late scoring time. FHB

10.3389/fgene.2022.988264

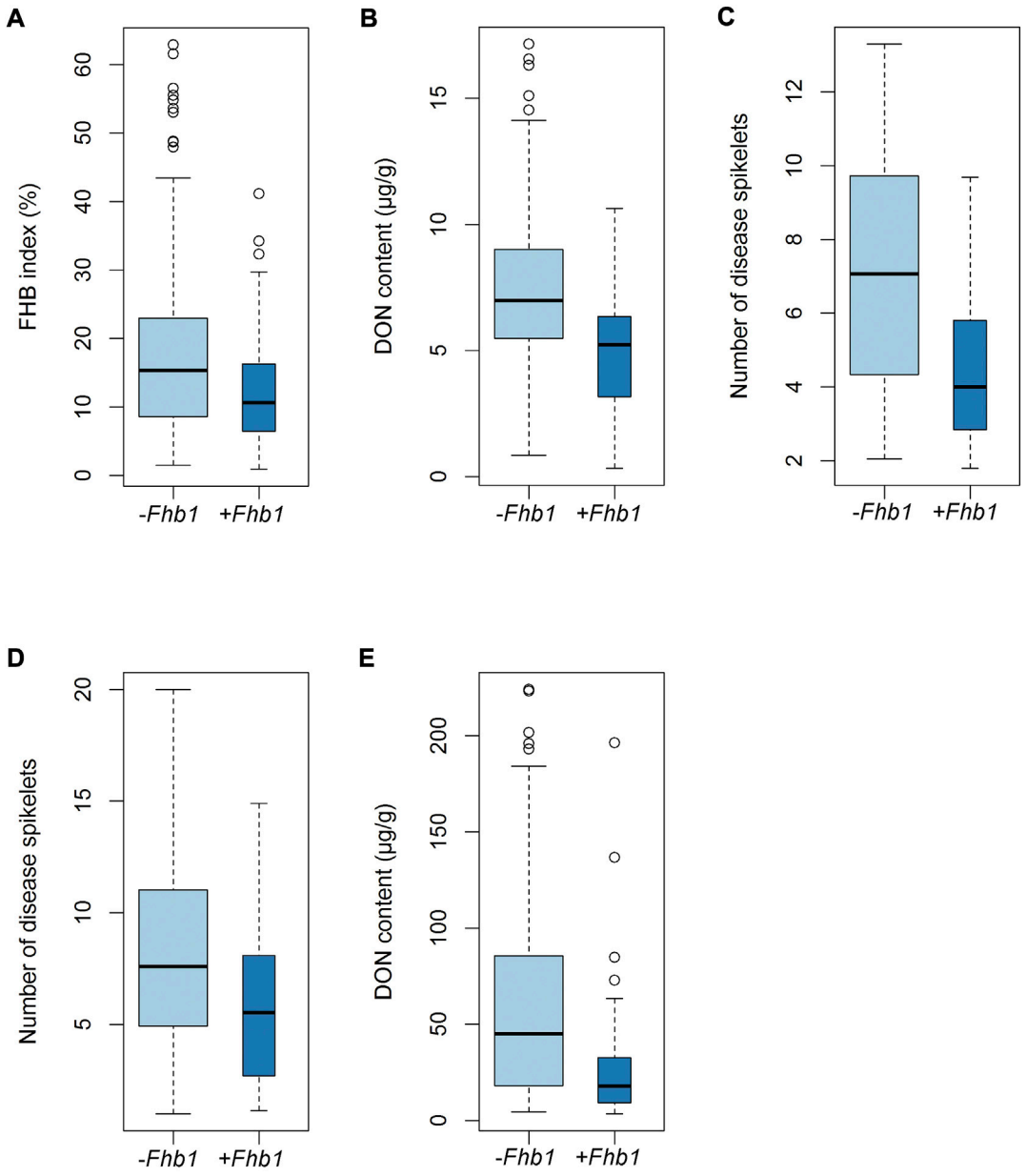



FIGURE 3
Differences in FHB index (A), number of diseased spikelets (C and D) and DON content (B and E) between *Fhb1* and non-*Fhb1* accessions.

severity is characterized only for Type II resistance by the point inoculation method in China, which is considered a stable and accurate screening method. There was no significant correlation between the data of FHB index in Mexico and FHB severity in China (Table 1), which was caused by accessions with contrasting resistance components in the two countries, e.g., several Chinese accessions showed good Type II resistance in China but poor Type I resistance in Mexico (Supplementary Table S1). The reason could be ascribed to the relatively high frequency of *Fhb1* (34.1%) conferring good Type II resistance in Chinese accessions. Uncontrollable factors such as environmental change and inoculum content make accurate phenotyping of Type I resistance difficult. Type II resistance evaluation is mandatory before the release of Chinese varieties, while other resistance components including Type I resistance are optional in breeding program. However, accessions with good Type II resistance but poor Type I resistance could still suffer high yield loss and DON contamination under natural conditions with high *Fusarium* pressure. Pyramiding of Types I and II resistance to FHB will be a promising breeding strategy, and introgression of *Fhb1*, *Fhb4* and *Fhb5* into five modern Chinese

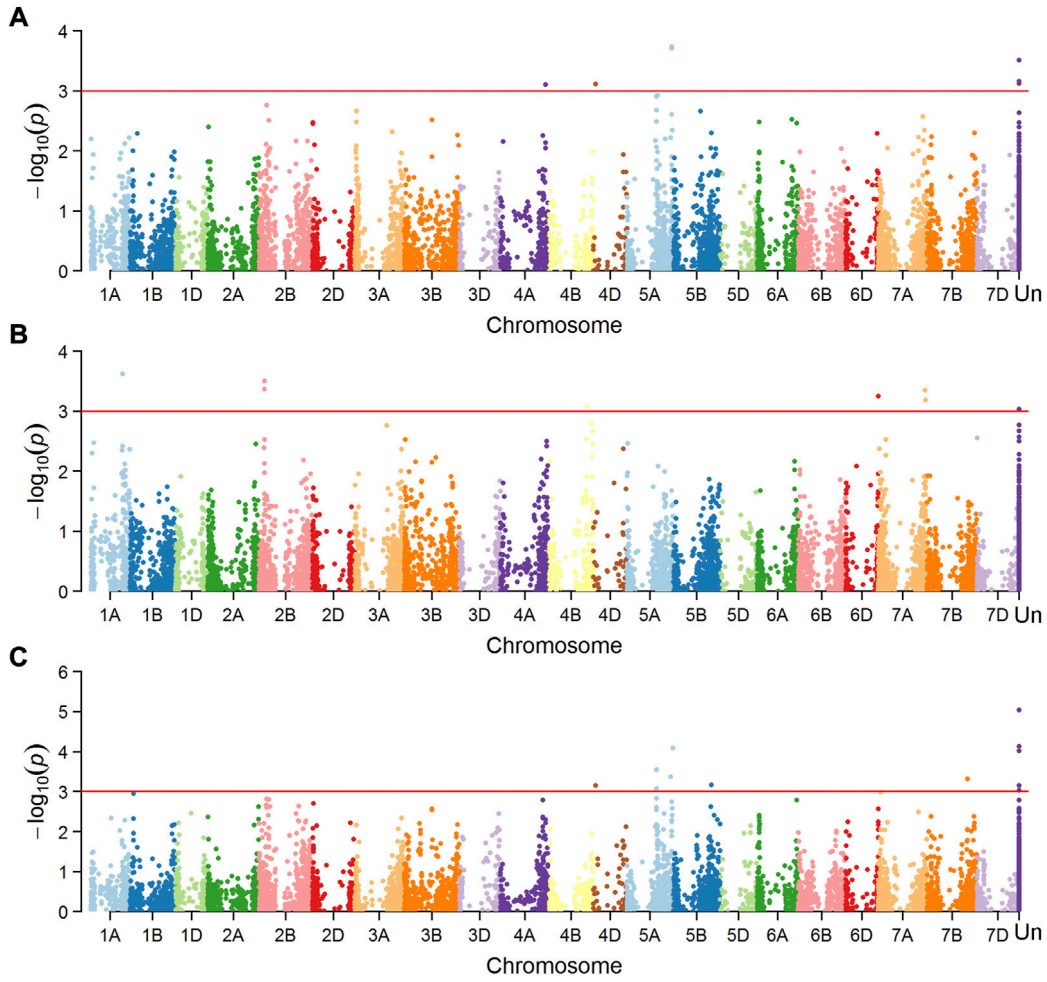



FIGURE 4
Manhattan plots showing SNPs associated with number of diseased spikelets using point inoculation **(A)** and spawn inoculation **(B)**, and those for DON content using point inoculation **(C)** in China.

wheat reduced FHB severity by 95% without penalty for agronomic traits and yield (Zhang et al., 2021). In Mexico, toxin content was moderately correlated with the FHB index, while in China, grain toxin levels were highly correlated with Type II resistance (Table 1). These data suggested that DON content in grains actually represented a combination of Type I/II and III resistance, which may explain why QTLs exclusively associated with Type III resistance are rarely mapped (He et al., 2019).

Most of the wheat varieties in the middle and lower reaches of the Yangtze river in China have *Fhb1* that confers moderate resistance to FHB (data not shown), otherwise it would not be released in National Wheat Production Trials. Recently, the development of wheat germplasm combining *Fhb1* and *Sr2* in CIMMYT backgrounds would be used in breeding for both FHB and stem rust resistance (He et al., 2020). Large quantities of QTL mapping and omics data for FHB resistance have been released in the previous reports, it is widely accepted that *Fhb1* on 3BS chromosome is a major and stable locus for Type II resistance to FHB (Cuthbert et al., 2006; Liu et al., 2006; Gunnaiah et al., 2012; Schweiger et al., 2013; Schweiger et al., 2016; Eldakak et al., 2018). Previous studies have shown that *Fhb1* did not effectively increase resistance to FHB in certain genotypes (Pumphrey et al., 2007), and a similar case has been observed in our data, that is, several accessions harboring *Fhb1* showed moderate susceptibility to FHB (Supplementary Table S1). The additive effect of minor loci might play an important role in FHB resistance, and non-*Fhb1* accessions with moderate susceptibility to resistance could still be used in breeding or genetic analysis (Supplementary Table S1). Totally 11 accessions exhibited stable moderate resistance and low DON content in grains (Table 2). It was worth noting that three accessions Klein Don Enrique, Chuko and Yumai34 did not contain *Fhb1*, but were resistant to multiple *Fusarium* species from Mexico and China. These non-*Fhb1* accessions with moderate resistance to

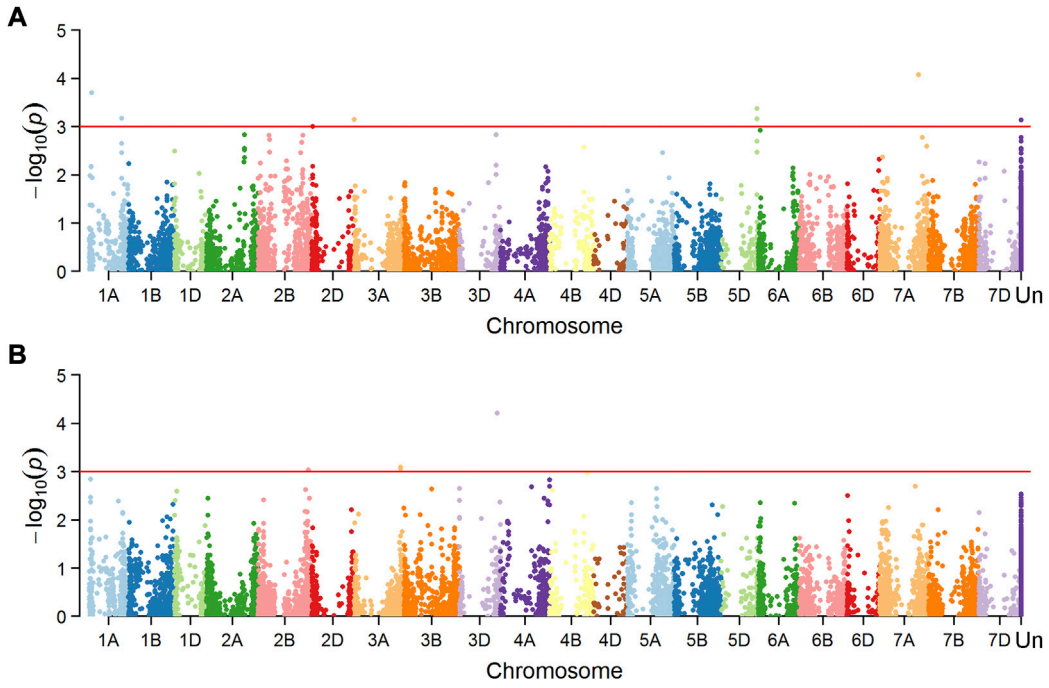



FIGURE 5
Manhattan plots showing SNPs associated with FHB index using spray inoculation **(A)** and those for DON content **(B)** using spray inoculation in Mexico.

FHB would facilitate their application in breeding and provide alternative options for FHB improvement.

The present study identified 16 genetic loci associated with FHB resistance and/or DON content in grains, 6 of which overlapped with reported metaQTL intervals (Zheng et al., 2021). Marker 4910975 is in the interval of previously reported smQTL-1A-3 (28–38 Mb) derived from "CJ9303" and two European winter wheats "History" and "Pirat" (Jiang et al., 2007; Holzapfel et al., 2008). Marker 1099971 is close to the flanking marker IWA7577 of smQTL-1A-5 (484–509 Mb) and overlapped with this QTL based on the LD data (Zheng et al., 2021). Marker 1099971 is found in the second exon of *TraesCS1A02G283400*, which is annotated as a copper ion-binding protein involved in lignin catalytic synthesis. Marker 979146 associated with Type II FHB resistance here overlaps with reported QTL interval of smQTL-5A-5 from "Wangshuibai" and CIMMYT wheat line "C615" (Ma et al., 2006; Yi et al., 2018). A cluster of SNPs tagged by marker 1157139 falls into the interval of smQTL-5A-8 (681–694 Mb), which is a minor QTL from Swiss winter wheat cultivar "Arina" (Paillard et al., 2004). A locus associated with DON content in grains tagged by marker 989900 is close to the flanking marker of smQTL-5A-7 (662-645 Mb) (Zheng et al., 2021), which has been associated with Types II and III resistance in different accessions by several independent research groups (Buerstmayr et al., 2011; Chu et al., 2011; Lu et al., 2013; Malihipour et al., 2017; Zhao et al., 2018). Similarly, marker 1091498 that was only associated with DON content in grains falls into a known QTL interval, too. Considering that these two loci were significant in point inoculation experiments, they are likely associated with Type II resistance, too (Zheng et al., 2021). The rest 10 loci are not tightly linked to the previously reported QTL through sequence alignments, implying that they may be novel for FHB resistance. Marker 1229379 associated with FHB resistance is located in the coding region of *TraesCS7A02G524200* on 7A, which could play a role in defense responses to pathogens by maintaining cell membrane integrity.

Our results, together with the large number of loci reported in previous studies (Arruda et al., 2016a; He et al., 2016; Buerstmayr et al., 2020; Ma et al., 2020; Zheng et al., 2021), suggest that multiple genes/loci are involved in FHB resistance under different environments, although their effects are mostly small and could not be easily used in MAS. Four MTAs (3026949, 979146, 1157139 and 1034520) were associated with both Type II resistance and DON content, which was expected due to the high correlation between grain toxin levels and Type II resistance, as described in the present study and previous reports (He et al., 2019; Zheng et al., 2021). Markers 1099971, 1044062 and 1091498 were identified in experiments in both China and Mexico, which may suggest their roles in response to different environmental factors and *Fusarium* species.

TABLE 3 Markers significantly (-log10(*p*) ≥3.0) associated with FHB resistance and DON content.

| Marker id | SNP | Chr | Position (bp) | $R^2$ (%) | RAF (%) | Overlapping gene | Overlapping mQTL | Model | Environments |
|---|---|---|---|---|---|---|---|---|---|
| FHB resistance | | | | | | | | | |
| 4910975 | A/G | 1A | 31918043 | 7.26 | 15.66 | NA | sMQTL-1A-3 | P+K; Q+K | MX-spray |
| 1099971a | C/T | 1A | 480934139 | 5.40 | 21.26 | TraesCS1A02G283400 | sMQTL-1A-5 | P+K; Q+K | CN-spawn; MX-spray |
| 1081753 | T/C | 1B | 258837217 | 7.04 | 5.63 | NA | NA | P+K; Q+K | MX-spray |
| 1055088 | T/G | 2B | 73002236 | 5.76 | 48.58 | NA | NA | P+K; Q+K | CN-spawn |
| 1044062 | A/G | 4B | 572558191 | 4.69 | 8.12 | NA | NA | P+K; Q+K | CN-spawn; MX-DON |
| 3026949 | A/G | 4D | 35405745 | 4.43 | 4.31 | NA | NA | P+K | CN-point; CN-pointDON |
| 979146 | G/A | 5A | 466024980 | 5.56 | 94.62 | NA | sMQTL-5A-5 | Q+K | CN-point; CN-pointDON |
| 1157139b | T/C | 5A | 697068377 | 8.27 | 6.05 | NA | sMQTL-5A-8 | P+K; Q+K | CN-point; CN-pointDON |
| 1229379 | G/C | 7A | 706764943 | 5.73 | 18.11 | TraesCS7A02G524200 | NA | P+K; Q+K | CN-spawn |
| 1034520 | C/A | 7B | 701319079 | 10.49 | 19.60 | NA | NA | P+K; Q+K | CN-point; CN-pointDON |
| DON content | | | | | | | | | |
| 3956613 | G/A | 2B | 753628598 | 7.00 | 21.20 | TraesCS2B02G559400 | NA | P+K; Q+K | MX-DON |
| 1285715 | T/C | 3A | 705290933 | 6.01 | 67.57 | NA | NA | P+K; Q+K | MX-DON |
| 1218288 | A/G | 3D | 571054888 | 7.21 | 68.25 | NA | NA | P+K; Q+K | MX-DON |
| 1091396 | T/C | 4D | 502708990 | 5.27 | 6.92 | TraesCS4D02G350500 | NA | P+K | CN-pointDON |
| 989900 | C/A | 5A | 666683794 | 4.73 | 90.77 | NA | sMQTL-5A-7 | Q+K | CN-pointDON |
| 1091498 | C/T | 5B | 571213517 | 6.92 | 65.32 | NA | sMQTL-5B-4 | Q+K | CN-pointDON; MX-DON |

1099971a, significant markers 1099971 and 3064923 located on the same locus based on the LD, analysis. 1157139b, significant markers 1157139, 1217190, 5050428, 2259167, 5371234, 982983, 1100295 and 1091475 located on the same locus based on LD analysis. Physical positions for the associated SNPs were based on Chinese Spring reference genome v1.0. Physical positions of the significant SNPs were compared to metaQTL intervals reported in Zheng et al., 2021. RAF refers to the resistance allele frequencies of the associated SNPs, and $R^2$ refers to the phenotypic variation explained by the associated SNPs.

Most significant MTAs reported here appeared as single SNP markers, possibly due to the low resolution of genotyping. Similar cases have been reported by (Arruda et al., 2016a) and (Miedaner et al., 2011). Based on MTAs, genomic selection has shown intermediate to high prediction accuracy for FHB resistance traits, including DON content in grains, Fusarium damaged kernels and FHB severity (Arruda et al., 2016b; Gaire et al., 2022a). The effects of the markers reported here could be re-estimated in larger breeding populations and used in genomic selection for FHB resistance improvement.

## Conclusion

FHB screening experiments on 265 wheat accessions in Mexico and China showed that 11 accessions had stable FHB resistance or low DON content in grains, of which three accessions Klein Don Enrique, Chuko and Yumai34 did not contain *Fhb1*. Sixteen loci associated with FHB resistance or DON content in grains were identified on chromosomes 1A, 1B, 2B, 3A, 3D, 4B, 4D, 5A, 5B, 7A, and 7B in multiple environments, 6 of which overlapped with reported metaQTL. The genetic sources could be used in the breeding programs for FHB improvement, and the associated loci could be further mapped to identify better markers for MAS.

## 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. Raw genotypic data is accessible via https://hdl.handle.net/11529/10548748.

## Author contributions

XZ and PS conceived and designed the experiments, XH and KX conducted the FHB screening in Mexico, LW conducted the FHB screening in China, YH and PJ assisted in genotypic analysis, LW and XH analyzed the data and wrote the first draft of the manuscript, and all authors reviewed and approved the final manuscript.

## Funding

This research has been funded by National Key R&D Program of China (2017YFE0126700), CRP WHEAT, National Natural Sciences Foundation of China (31561143004), Jiangsu Agricultural Science and Technology Innovation Fund (CX (21)3099), Seed Industry Revitalization Project of Jiangsu Province (JBGS2021006), and Collaborative

Innovation Center for Modern Crop Production co-sponsored by Province and Ministry.

## Acknowledgments

Technical support from Francisco Lopez in field trials and Nerida Lozano in inoculum preparation for trials in Mexico is highly acknowledged. Technical support from Chang Li and Lixuan Yu in field trials in China is also well appreciated.

## Conflict of interest

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

## Publisher's note

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

## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.988264/full#supplementary-material

## References

Arruda, M. P., Brown, P., Brown-Guedira, G., Krill, A. M., and Thurber, C. (2016a). Genome-wide association mapping of Fusarium head blight resistance in wheat using genotyping-by-sequencing. *Plant Genome* 9, 1–14. doi:10.3835/plantgenome2015.04.0028

Arruda, M. P., Lipka, A. E., Brown, P. J., Krill, A. M., Thurber, C., Brown-Guedira, G., et al. (2016b). Comparing genomic selection and marker-assisted selection for Fusarium head blight resistance in wheat (*Triticum aestivum* L.). *Mol. Breed.* 36, 84. doi:10.1007/s11032-016-0508-5

Backhouse, D. (2014). Global distribution of *Fusarium graminearum*, *F. asiaticum* and *F. boothii* from wheat in relation to climate. *Eur. J. Plant Pathol.* 139, 161–173. doi:10.1007/s10658-013-0374-5

Bai, G., and Shaner, G. (2004). Management and resistance in wheat and barley to Fusarium head blight. *Annu. Rev. Phytopathol.* 42, 135–161. doi:10.1146/annurev.phyto.42.040803.140340

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

Buerstmayr, M., Lemmens, M., Steiner, B., and Buerstmayr, H. (2011). Advanced backcross QTL mapping of resistance to Fusarium head blight and plant morphological traits in a *Triticum macha* × *T. aestivum* population. *Theor. Appl. Genet.* 123, 293–306. doi:10.1007/s00122-011-1584-x

Buerstmayr, M., Steiner, B., and Buerstmayr, H. (2020). Breeding for Fusarium head blight resistance in wheat—progress and challenges. *Plant Breed.* 139, 12797. doi:10.1111/pbr.12797

Buerstmayr, M., Steiner, B., Wagner, C., Schwarz, P., Brugger, K., Barabaschi, D., et al. (2018). High-resolution mapping of the pericentromeric region on wheat chromosome arm 5AS harbouring the Fusarium head blight resistance QTL Qfhs.ifa-5A. *Plant Biotechnol. J.* 16, 1046–1056. doi:10.1111/pbi.12850

Cainong, J. C., Bockus, W. W., Feng, Y., Chen, P., Qi, L., Sehgal, S. K., et al. (2015). Chromosome engineering, mapping, and transferring of resistance to Fusarium head blight disease from *Elymus tsukushiensis* into wheat. *Theor. Appl. Genet.* 128, 1019–1027. doi:10.1007/s00122-015-2485-1

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

Chu, C., Niu, Z., Zhong, S., Chao, S., Friesen, T. L., Halley, S., et al. (2011). Identification and molecular mapping of two QTLs with major effects for resistance to Fusarium head blight in wheat. *Theor. Appl. Genet.* 123, 1107–1119. doi:10.1007/s00122-011-1652-2

Cuthbert, P. A., Somers, D. J., and Brulé-Babel, A. (2007). Mapping of Fhb2 on chromosome 6BS: A gene controlling Fusarium head blight field resistance in bread wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 114, 429–437. doi:10.1007/s00122-006-0439-3

Cuthbert, P. A., Somers, D. J., Thomas, J., Cloutier, S., and Brulé-Babel, A. (2006). Fine mapping Fhb1, a major gene controlling Fusarium head blight resistance in bread wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 112, 1465–1472. doi:10.1007/s00122-006-0249-7

Eldakak, M., Das, A., Zhuang, Y., Rohila, J. S., Glover, K., and Yen, Y. (2018). A quantitative proteomics view on the function of Qfhb1, a major QTL for Fusarium head blight resistance in wheat. *Pathogens* 7, 7030058. doi:10.3390/pathogens7030058

Gaire, R., De Arruda, M. P., Mohammadi, M., Brown-Guedira, G., Kolb, F. L., and Rutkoski, J. (2022a). Multi-trait genomic selection can increase selection accuracy for deoxynivalenol accumulation resulting from fusarium head blight in wheat. *Plant Genome* 15, e20188. doi:10.1002/tpg2.20188

Gaire, R., Sneller, C., Brown-Guedira, G., Van Sanford, D. A., Mohammadi, M., Kolb, F. L., et al. (2022b). Genetic trends in Fusarium head blight resistance due to 20 years of winter wheat breeding and cooperative testing in the Northern US. *Plant Dis.* 106, 364–372. doi:10.1094/pdis-04-21-0891-sr

Ghimire, B., Sapkota, S., Bahri, B. A., Martinez-Espinoza, A. D., Buck, J. W., and Mergoum, M. (2020). Fusarium head blight and rust diseases in soft red winter wheat in the southeast United States: State of the art, challenges and future perspective for breeding. *Front. Plant Sci.* 11, 1080. doi:10.3389/fpls.2020.01080

Gunnaiah, R., Kushalappa, A. C., Duggavathi, R., Fox, S., and Somers, D. J. (2012). Integrated metabolo-proteomic approach to decipher the mechanisms by which wheat QTL (Fhb1) contributes to resistance against *Fusarium graminearum*. *PLoS One* 7, e40695. doi:10.1371/journal.pone.0040695

Guo, J., Zhang, X., Hou, Y., Cai, J., Shen, X., Zhou, T., et al. (2015). High-density mapping of the major FHB resistance gene Fhb7 derived from *Thinopyrum ponticum* and its pyramiding with Fhb1 by marker-assisted selection. *Theor. Appl. Genet.* 128, 2301–2316. doi:10.1007/s00122-015-2586-x

Hao, Y., Rasheed, A., Zhu, Z., Wulff, B. B. H., and He, Z. (2020). Harnessing wheat Fhb1 for Fusarium resistance. *Trends Plant Sci.* 25, 1–3. doi:10.1016/j.tplants.2019.10.006

He, X., Brar, G. S., Bonnett, D., Dreisigacker, S., Hyles, J., Spielmeyer, W., et al. (2020). Disease resistance evaluation of elite CIMMYT wheat lines containing the coupled Fhb1 and Sr2 genes. *Plant Dis.* 104, 2369–2376. doi:10.1094/pdis-02-20-0369-re

He, X., Dreisigacker, S., Singh, R. P., and Singh, P. K. (2019). Genetics for low correlation between Fusarium head blight disease and deoxynivalenol (DON) content in a bread wheat mapping population. *Theor. Appl. Genet.* 132, 2401–2411. doi:10.1007/s00122-019-03362-9

He, X., Lillemo, M., Shi, J., Wu, J., Bjørnstad, Å., Belova, T., et al. (2016). QTL characterization of Fusarium head blight resistance in CIMMYT bread wheat line Soru#1. *PLoS One* 11, e0158052. doi:10.1371/journal.pone.0158052

He, X., Singh, P. K., Duveiller, E., Schlang, N., Dreisigacker, S., and Singh, R. P. (2013). Identification and characterization of international Fusarium head blight screening nurseries of wheat at CIMMYT, Mexico. *Eur. J. Plant Pathol.* 136, 123–134. doi:10.1007/s10658-012-0146-7

He, Y., Zhang, X., Zhang, Y., Ahmad, D., Wu, L., Jiang, P., et al. (2018). Molecular characterization and expression of PFT, an FHB resistance gene at the Fhb1 QTL in wheat. *Phytopathology* 108, 730–736. doi:10.1094/phyto-11-17-0383-r

Holzapfel, J., Voss, H. H., Miedaner, T., Korzun, V., Häberle, J., Schweizer, G., et al. (2008). Inheritance of resistance to Fusarium head blight in three European winter wheat populations. *Theor. Appl. Genet.* 117, 1119–1128. doi:10.1007/s00122-008-0850-z

Jakobsson, M., and Rosenberg, N. A. (2007). Clumpp: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23, 1801–1806. doi:10.1093/bioinformatics/btm233

Jiang, G. L., Shi, J., and Ward, R. W. (2007). QTL analysis of resistance to Fusarium head blight in the novel wheat germplasm CJ 9306. I. Resistance to fungal spread. *Theor. Appl. Genet.* 116, 3–13. doi:10.1007/s00122-007-0641-y

Li, G., Zhou, J., Jia, H., Gao, Z., Fan, M., Luo, Y., et al. (2019). Mutation of a histidine-rich calcium-binding-protein gene in wheat confers resistance to Fusarium head blight. *Nat. Genet.* 51, 1106–1112. doi:10.1038/s41588-019-0426-7

Liu, S., Zhang, X., Pumphrey, M. O., Stack, R. W., Gill, B. S., and Anderson, J. A. (2006). Complex microcolinearity among wheat, rice, and barley revealed by fine mapping of the genomic region harboring a major QTL for resistance to Fusarium head blight in wheat. *Funct. Integr. Genomics* 6, 83–89. doi:10.1007/s10142-005-0007-y

Lu, Q., Lillemo, M., Skinnes, H., He, X., Shi, J., Ji, F., et al. (2013). Anther extrusion and plant height are associated with Type I resistance to Fusarium head blight in bread wheat line 'Shanghai-3/Catbird. *Theor. Appl. Genet.* 126, 317–334. doi:10.1007/s00122-012-1981-9

Ma, H. X., Zhang, K. M., Gao, L., Bai, G. H., Chen, H. G., Cai, Z. X., et al. (2006). Quantitative trait loci for resistance to fusarium head blight and deoxynivalenol accumulation in Wangshuibai wheat under field conditions. *Plant Pathol.* 55, 739–745. doi:10.1111/j.1365-3059.2006.01447.x

Ma, Z., Xie, Q., Li, G., Jia, H., Zhou, J., Kong, Z., et al. (2020). Germplasms, genetics and genomics for better control of disastrous wheat Fusarium head blight. *Theor. Appl. Genet.* 133, 1541–1568. doi:10.1007/s00122-019-03525-8

Malihipour, A., Gilbert, J., Fedak, G., Brûlé-Babel, A., and Cao, W. (2017). Mapping the A Genome for QTL conditioning resistance to Fusarium head blight in a wheat population with *Triticum timopheevii* background. *Plant Dis.* 101, 11–19. doi:10.1094/pdis-02-16-0144-re

Mesterhazy, A. (1995). Types and components of resistance to Fusarium head blight of wheat. *Plant Breed.* 114, 377–386. doi:10.1111/j.1439-0523.1995.tb00816.x

Mesterhazy, A. (2020). Updating the breeding philosophy of wheat to Fusarium head blight (FHB): Resistance components, QTL identification, and phenotyping-A review. *Plants (Basel)* 9, 1702. doi:10.3390/plants9121702

Mesterházy, T., Mirocha, C. G., and Komoróczy, R. (1999). Nature of wheat resistance to Fusarium head blight and the role of deoxynivalenol for breeding. *Plant Breed.* 118, 97–110. doi:10.1046/j.1439-0523.1999.118002097.x

Miedaner, T., Würschum, T., Maurer, H. P., Korzun, V., Ebmeyer, E., and Reif, J. C. (2011). Association mapping for Fusarium head blight resistance in European soft winter wheat. *Mol. Breed.* 28, 647–655. doi:10.1007/s11032-010-9516-z

Miller, J. D., and Amison, P. G. (1986). Degradation of deoxynivalenol by suspension cultures of the Fusarium head blight resistant wheat cultivar Frontana. *Can. J. Plant Pathology* 8, 147–150. doi:10.1080/07060668609501818

Paillard, S., Schnurbusch, T., Tiwari, R., Messmer, M., Winzeler, M., Keller, B., et al. (2004). QTL analysis of resistance to Fusarium head blight in Swiss winter wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 109, 323–332. doi:10.1007/s00122-004-1628-6

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959. doi:10.1093/genetics/155.2.945

Pumphrey, M. O., Bernardo, R., and Anderson, J. A. (2007). Validating the Fhb1 QTL for Fusarium head blight resistance in near-isogenic wheat lines developed from breeding populations. *Crop Sci.* 47, 200–206. doi:10.2135/cropsci2006.03.0206

Qi, L. L., Pumphrey, M. O., Friebe, B., Chen, P. D., and Gill, B. S. (2008). Molecular cytogenetic characterization of alien introgressions with gene Fhb3 for resistance to Fusarium head blight disease of wheat. *Theor. Appl. Genet.* 117, 1155–1166. doi:10.1007/s00122-008-0853-9

R Team (2011). *R: A language and environment for statistical computing*. Vienna, Austria: the R foundation for statistical computing. Available at: http://www.R-project.org/.

Rawat, N., Pumphrey, M. O., Liu, S., Zhang, X., Tiwari, V. K., Ando, K., et al. (2016). Wheat Fhb1 encodes a chimeric lectin with agglutinin domains and a pore-forming toxin-like domain conferring resistance to Fusarium head blight. *Nat. Genet.* 48, 1576–1580. doi:10.1038/ng.3706

Schroeder, H. W., and Christensen, J. J. (1963). Factors affecting resistance of wheat to scab caused by Gibberella zeae. *Phytopathology* 53, 831–838.

Schweiger, W., Steiner, B., Ametz, C., Siegwart, G., Wiesenberger, G., Berthiller, F., et al. (2013). Transcriptomic characterization of two major Fusarium resistance quantitative trait loci (QTLs), Fhb1 and Qfhs.ifa-5A, identifies novel candidate genes. *Mol. Plant Pathol.* 14, 772–785. doi:10.1111/mpp.12048

Schweiger, W., Steiner, B., Vautrin, S., Nussbaumer, T., Siegwart, G., Zamini, M., et al. (2016). Suppressed recombination and unique candidate genes in the divergent haplotype encoding Fhb1, a major Fusarium head blight resistance locus in wheat. *Theor. Appl. Genet.* 129, 1607–1623. doi:10.1007/s00122-016-2727-x

Singh, L., Anderson, J. A., Chen, J., Gill, B. S., Tiwari, V. K., and Rawat, N. (2019). Development and validation of a perfect KASP marker for Fusarium head blight resistance gene Fhb1 in wheat. *Plant Pathol. J.* 35, 200–207. doi:10.5423/ppj.oa.01.2019.0018

Steiner, B., Buerstmayr, M., Wagner, C., Danler, A., Eshonkulov, B., Ehn, M., et al. (2019). Fine-mapping of the Fusarium head blight resistance QTL Qfhs.ifa-5A identifies two resistance QTL associated with anther extrusion. *Theor. Appl. Genet.* 132, 2039–2053. doi:10.1007/s00122-019-03336-x

Su, Z., Bernardo, A., Tian, B., Chen, H., Wang, S., Ma, H., et al. (2019). A deletion mutation in TaHRC confers Fhb1 resistance to Fusarium head blight in wheat. *Nat. Genet.* 51, 1099–1105. doi:10.1038/s41588-019-0425-8

Su, Z., Jin, S., Zhang, D., and Bai, G. (2018). Development and validation of diagnostic markers for Fhb1 region, a major QTL for Fusarium head blight resistance in wheat. *Theor. Appl. Genet.* 131, 2371–2380. doi:10.1007/s00122-018-3159-6

Van Der Lee, T., Zhang, H., Van Diepeningen, A., and Waalwijk, C. (2015). Biogeography of *Fusarium graminearum* species complex and chemotypes: A review. *Food Addit. Contam. Part A Chem. Anal. Control Expo. Risk Assess.* 32, 453–460. doi:10.1080/19440049.2014.984244

Venske, E., Dos Santos, R. S., Farias, D. D. R., Rother, V., Da Maia, L. C., Pegoraro, C., et al. (2019). Meta-analysis of the QTLome of Fusarium head blight resistance in bread wheat: Refining the current puzzle. *Front. Plant Sci.* 10, 727. doi:10.3389/fpls.2019.00727

Wang, H., Sun, S., Ge, W., Zhao, L., Hou, B., Wang, K., et al. (2020). Horizontal gene transfer of Fhb7 from fungus underlies Fusarium head blight resistance in wheat. *Science* 368, eaba5435. doi:10.1126/science.aba5435

Wu, L., He, X., Kabir, M. R., Roy, K. K., Anwar, M. B., Marza, F., et al. (2021). Genetic sources and loci for wheat head blast resistance identified by genome-wide association analysis. *Crop J.* 10, 793–801. doi:10.1016/j.cj.2021.07.007

Xue, S., Li, G., Jia, H., Xu, F., Lin, F., Tang, M., et al. (2010). Fine mapping Fhb4, a major QTL conditioning resistance to Fusarium infection in bread wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 121, 147–156. doi:10.1007/s00122-010-1298-5

Xue, S., Xu, F., Tang, M., Zhou, Y., Li, G., An, X., et al. (2011). Precise mapping Fhb5, a major QTL conditioning resistance to Fusarium infection in bread wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 123, 1055–1063. doi:10.1007/s00122-011-1647-z

Yi, X., Cheng, J., Jiang, Z., Hu, W., Bie, T., Gao, D., et al. (2018). Genetic analysis of Fusarium head blight resistance in CIMMYT bread wheat line C615 using traditional and conditional QTL mapping. *Front. Plant Sci.* 9, 573. doi:10.3389/fpls.2018.00573

Zhang, H., Van Der Lee, T., Waalwijk, C., Chen, W., Xu, J., Xu, J., et al. (2012). Population analysis of the *Fusarium graminearum* species complex from wheat in China show a shift to more aggressive isolates. *PLoS One* 7, e31722. doi:10.1371/journal.pone.0031722

Zhang, Y., Yang, Z., Ma, H., Huang, L., Ding, F., Du, Y., et al. (2021). Pyramiding of Fusarium head blight resistance quantitative trait loci Fhb1,Fhb4, and Fhb5, in modern Chinese wheat cultivars. *Front. Plant Sci.* 12, 694023. doi:10.3389/fpls.2021.694023

Zhao, M., Leng, Y., Chao, S., Xu, S. S., and Zhong, S. (2018). Molecular mapping of QTL for Fusarium head blight resistance introgressed into durum wheat. *Theor. Appl. Genet.* 131, 1939–1951. doi:10.1007/s00122-018-3124-4

Zheng, T., Hua, C., Li, L., Sun, Z., Yuan, M., Bai, G., et al. (2021). Integration of meta-QTL discovery with omics: Towards a molcular breeding platform for improving wheat resistance to Fusarium head blight. *Crop J.* 9, 739–749. doi:10.1016/j.cj.2020.10.006

Zhu, Z., Hao, Y., Mergoum, M., Bai, G., Humphreys, G., Cloutier, S., et al. (2019). Breeding wheat for resistance to Fusarium head blight in the global North: China, USA, and Canada. *Crop J.* 7, 730–738. doi:10.1016/j.cj.2019.06.003

Zhu, Z., Xu, D., Cheng, S., Gao, C., Xia, X., Hao, Y., et al. (2018). Characterization of Fusarium head blight resistance gene Fhb1 and its putative ancestor in Chinese wheat germplasm. *Acta Agron. Sin.* 44, 473–482. doi:10.3724/sp.j.1006.2018.00473

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 21 October 2022
DOI 10.3389/fgene.2022.1034947



OPEN ACCESS

EDITED BY
Dwijesh Chandra Mishra,
Indian Agricultural Statistics Research Institute (ICAR) India

REVIEWED BY
Upendra Kumar,
Chaudhary Charan Singh Haryana Agricultural University, India
Hengyou Zhang,
Northeast Institute of Geography and Agroecology (CAS), China

*CORRESPONDENCE
Hari Krishna,
harikrishna.agri@gmail.com
P. K. Singh,
pksinghiari@gmail.com

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

RECEIVED 02 September 2022
ACCEPTED 07 October 2022
PUBLISHED 21 October 2022

CITATION
Devate NB, Krishna H, Sunilkumar VP, Manjunath KK, Mishra CN, Jain N, Singh GP and Singh PK (2022), Identification of genomic regions of wheat associated with grain Fe and Zn content under drought and heat stress using genome-wide association study. *Front. Genet.* 13:1034947. doi: 10.3389/fgene.2022.1034947



# Identification of genomic regions of wheat associated with grain Fe and Zn content under drought and heat stress using genome-wide association study

Narayana Bhat Devate[1], Hari Krishna[1]*, V. P. Sunilkumar[1], Karthik Kumar Manjunath[1], C. N. Mishra[2], Neelu Jain[1], G. P. Singh[2] and P. K. Singh[1]*

[1]Division of Genetics, ICAR-Indian Agricultural Research Institute, New Delhi, India, [2]ICAR- Indian Institute of Wheat and Barley Research, Karnal, India

Wheat is the staple food crop of global importance for its grain nutrient quality. Grain iron and zinc content of the wheat grain is an important quantitatively inherited trait that is influenced by the environmental factors such as drought and heat stress. Phenotypic evaluation of 295 advanced breeding lines from the wheat stress breeding program of IARI was carried out under timely sown irrigated (IR), restricted irrigated, and late-sown conditions at New Delhi during the cropping season of 2020–21, and grain iron (GFeC) and zinc (GZnC) contents were estimated from both control and treatments. A statistically significant increase in GFeC and GZnC was observed under stress conditions compared to that of the control. Genotyping was carried out with the SNPs from the 35K Axiom Breeder's array, and marker–trait association was identified by GWAS analysis. Of the 23 MTAs identified, seven were linked with GFeC and sixteen were linked with GZnC. *In silico* analysis revealed a few important transcripts involved in various plant metabolism, growth, and development activities such as *auxin response factor*, *root UVB sensitive proteins*, *potassium transporter*, *glycosyl transferase*, *COBRA*, and *F-box-like domain*. The identified MTAs can be used for molecular breeding after validation and also for rapid development of micronutrient-rich varieties of wheat to mitigate hidden hunger.

KEYWORDS

**wheat, grain iron and zinc content, GWAS, drought stress, heat stress**

## Introduction

Micronutrient deficiency, commonly known as "hidden hunger," is mostly brought on by diets that are frequently dominated by staple foods that are poor in minerals and vitamins (Liu et al., 2019). Two billion people are affected by it globally, and it accounts for nearly 45% of all yearly fatalities in children under the age of 5. A higher risk factor for

human health is the mineral deficiency of iron (Fe) and zinc (Zn), which affects around one-third of the population in underdeveloped nations (Rathan et al., 2022). In areas that are severely afflicted by micronutrient deficiencies, cereals make up the majority of daily dietary intake (Bouis et al., 2011). The micronutrient content of common cereals like wheat and rice, notably Fe and Zn, is below ideal levels.

According to Listman and Ordóez (2019), wheat is consumed by 2.5 billion people globally and is a staple food for 30% of the population, especially in developing nations (Lobell et al., 2011). Food fortification, supplementation, and diet diversity can all help solve the issue of micronutrient deficiency; however, these solutions are not long-lasting and are costly affairs, especially for rural poor people (Pfeiffer and McClafferty, 2007). Thus, increasing the nutritional content of crops by traditional and molecular methods, known as "biofortification," has been accepted as an effective and sustainable method to address the issues related to micronutrient deficiency (Krishnappa et al., 2022).

The majority of the wheat-growing region is impacted by scattered showers and regular rising heat stress followed by heat waves. Lower grain filling time, reduced starch accumulation, and smaller seeds are all effects of heat and drought stress. Apart from this, the grain sink capacity decreased the most under heat and drought stress (Zahra et al., 2021). Warmer temperatures brought on by climate change and decreased water availability in the majority of spring wheat production regions will impact food and nutritional security (Velu et al., 2016). Accumulation of mineral nutrients in the grain is a complex process, including several genes, and is greatly impacted by environmental factors. A thorough understanding of the genetic regulation of nutritional characteristics and their relationship with grain yield is required to breed stable grain nutrient-containing varieties under drought and heat stress (Samineni et al., 2022).

In order to breed cereals like wheat for biofortification using MAS, it is important to have information about the genomic regions that control grain Zn and Fe concentration. Numerous QTLs and genes influencing the amount of these micronutrients have been identified as a result of genetic studies in wheat (Srinivasa et al., 2014; Tiwari et al., 2016; Krishnappa et al., 2017; Gupta et al., 2021). Genome-wide association studies (GWAS) are now the most widely used method for determining the genetic basis of complex characteristics such as grain Fe (GFeC) and Zn (GZnC) content. The GWAS has the advantages of increased QTL resolution, allele coverage, and the capacity to employ huge collections of readily available materials such as natural populations or advanced breeding lines. The GWAS has been used to show the genomic regions of grain iron and zinc content in many studies (Alomari et al., 2018; Bhatta et al., 2018; Velu et al., 2018; El-Soda and Aljabri, 2022; Krishnappa et al., 2022; Rathan et al., 2022). However, grain Zn and Fe content under drought and heat stress is not being explored much. Genomic regions governing complex traits are mostly adaptive QTLs that were detected under specific conditions of the environment such as drought stress (Khaled et al., 2022). Hence, our objective is to dissect the genomic regions related to grain Fe and Zn content under drought and heat stress conditions using GWAS in bread wheat (*Triticum aestivum*) advanced breeding lines.

# Materials and methods

## Plant material and field layout

The GWAS mapping panel consists of 295 advanced breeding lines developed from popular Indian and exotic varieties as described in Devate et al. (2022). All the materials were evaluated at the IARI, New Delhi (Indian Agricultural Research Institute, research farm located at 28°38′30.5″N, 77°09′58.2″E, 228 m AMSL) with three conditions, viz., timely sown irrigated (IR), timely sown drought (RI), and late-sown condition (LS). A total of six irrigations were given for the irrigated and late-sown trials, whereas one irrigation before and one irrigation after the sowing were provided for restricted irrigated trials. IR and RI trials were carried out in the first fortnight of November, whereas LS trials were carried out in the second fortnight of December to expose plants to the natural heat in later growing periods. The experiment was conducted in an augmented RCBD design with four checks and six blocks with a plot size of 1 $m^2$ each.

## Phenotyping

Twenty random spikes from each line were collected in clean polyethylene bags and hand-threshed. Grain Fe and Zn contents (in mg/Kg) were estimated with 20-g seeds of each line through a high-throughput energy-dispersive X-ray fluorescence (ED-XRF) machine (model X-Supreme 8000; Oxford Instruments plc, Abingdon, United Kingdom) calibrated with glass bead-based values (Paltridge et al., 2012). Thousand kernel weight was measured by manual counting. The GWAS panel under all three conditions was phenotyped for GFeC, GZnC, and TGW. Phenotypic data were analyzed using the R package "augmentedRCBD" (Aravind et al., 2021) for ANOVA and adjusted means for each genotype under study. The z-test was used to compare the statistically significant differences between control and treatment means.

## Genotyping, population structure, and LD analysis

DNA isolation from leaf samples was carried out with the CTAB extraction method (Murray and Thompson 1980),

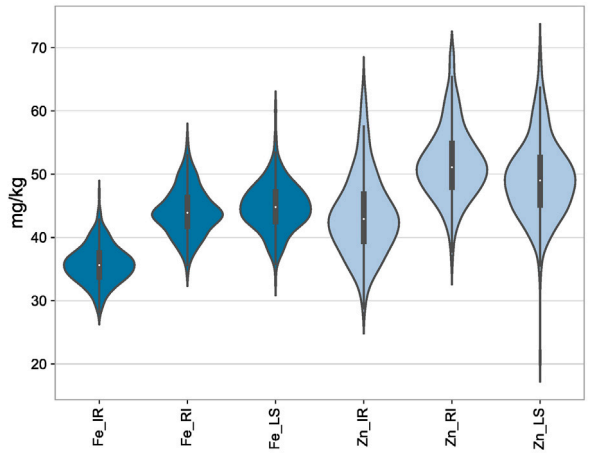



FIGURE 1
Depiction of the distribution of GFeC and GZnC under IR, RI, and LS conditions through violin plots.

followed by a DNA quality check through 0.8 percent agarose gel electrophoresis. Out of 295 genotypes, 282 DNA samples passed DNA quality thresholds and were genotyped using the Axiom Wheat Breeder's Genotyping Array (Affymetrix, Santa Clara, CA, United States) with 35,143 SNPs. The SNP filtering was carried out for minor allele frequency (MAF) of <5%, missing data of >20%, and heterozygote frequency of >25% before further analysis. The remaining 10,546 SNPs with phenotypic data from 282 genotypes were used for further analysis. Population structure based on STRUCTURE software and molecular marker-based PCA analysis was conducted. Also, pairwise $r^2$ values between markers and Linkage Disequilibrium decay plots were drawn as described in Devate et al. (2022).

## Analysis of data

Filtered 10,546 SNPs and adjusted means of GFeC and GZnC from each condition were used for genome-wide association analysis and were conducted using GAPIT v3 in R with the "BLINK" (Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway) (Huang et al., 2019) with PCA-based population structure as a fixed effect. Association model fitting was found using a Q–Q plot drawn with expected vs. observed $-log_{10}(p)$ value. Marker–trait associations (MTAs) in all three conditions for GFeC and GZnC were found with a significant $p$ value cut-off at 0.001. Chromosomal maps with identified MTAs at respective positions in Million Base-pairs (mb) on respective chromosomes with the name of associated traits were drawn with MapChart v2.32 (Voorrips, 2002).

Genotypes were grouped into two classes for each allele of MTAs, and their means were compared to identify trait-increasing and trait-decreasing alleles. Sequence information of significant markers was used for similarity search with the IWGSC reference genome with the basic local alignment search tool (BLAST) using the Ensembl Plants database (http://plants.ensembl.org/index.html) of the bread wheat genome [IWGSC (RefSeq v1.0)]. The gene coding regions located within and near the 100-kb flanking region of the MTAs were listed with transcript ID and described the protein (if any).

## Results and discussion

Analysis of variance in all the three conditions, viz., IR, RI, and LS, showed significant variation among the genotypes for GFeC, GZnC, and TGW. The mean values of GFeC and GZnC were increased under LS and RI conditions compared to those under IR (Figure 1), which was significant at both the one-tailed and two-tailed z tests at a $p$-value of 0.05. The average grain iron content under irrigated conditions was 35.7 mg/kg and was increased to 44.03 and 44.7 mg/kg under RI and LS conditions, respectively. Similarly, the average grain zinc content was increased from 43.52 to 51.86 mg/kg under RI and 49.29 mg/kg under LS conditions (Table 1). The decrease in TGW can be attributed to decreased assimilation of photosynthates and reduced grain sink capacity under drought and heat stress, ultimately leading to small and shriveled grains and reduced starch content in the endosperm (Zahra et al., 2021). It was observed that increased GFeC and GZnC under drought may be due to the shrinkage effect of grains, whereas under non-stressed conditions, due to high yield, the dilution effect of GFeC and GZnC results in a lower concentration. However, the nutrient yield per unit area decreases under stress (Velu et al., 2016).

A correlation study among the studied traits, viz., GFeC, GZnC, and TGW, showed a significant negative correlation between TGW and grain Fe and Zn content under RI conditions (at the $p$ values of 0.001 and 0.01, respectively) and grain Zn condition under LS conditions (at a $p$-value of 0.01), whereas no significant correlation was observed between TGW and grain Fe and Zn under control. GFeC and GZnC had a significant positive correlation among themselves under all the three treatment conditions, with a cut-off $p$-value of 0.001 (Table 1). Having a positive correlation between GFeC and GZnC can be efficiently utilized during a breeding program to achieve simultaneous improvement of both traits (Rathan et al., 2022). Broad sense heritability of GFeC, GZnC, and TGW was medium to high under all the three treatments, indicating the predominance of additive gene action. Traits governed by additive gene action and having positive correlation can be improved together efficiently in spite of environmental influences (Borah et al., 2018).

The GWAS panel under study had two subpopulations based on STRUCTURE and PCA scatter plots. Linkage-disequilibrium

TABLE 1 Summary and correlation of GFeC, GZnC, and TGW under IR, RI, and LS treatments in the GWAS panel evaluated at Delhi 2020–21.

| Treatment | Trait | Mean ± SD | Range | CV | Hbs | Percent increase/decrease | Z-test | Correlation | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | IR with RI and LS | IR with RI and LS | GFeC | GZnC | TGW |
| IR | GFeC | 35.7 ± 3.04 | 28.35–47.35 | 5.56 | 58.59 | — | — | 1 | 0.2*** | −0.09 |
| | GZnC | 43.52 ± 6.85 | 28.25–65.67 | 5.67 | 86.95 | — | — | | 1 | −0.04 |
| | TGW | 38.82 ± 4.48 | 24.92–57.29 | 6.2 | 75.72 | — | — | | | 1 |
| RI | GFeC | 44.03 ± 3.59 | 34.88–54.57 | 4.12 | 77.91 | 23.33% | 89.75** | 1 | 0.41*** | −0.17** |
| | GZnC | 51.86 ± 6.1 | 35.83–70.44 | 5.02 | 81.66 | 19.16% | 39.83** | | 1 | −0.21*** |
| | TGW | 32.98 ± 4.83 | 21.73–51.6 | 6.28 | 82.26 | −15.04% | $0.02^{NS}$ | | | 1 |
| LS | GFeC | 44.7 ± 4.01 | 32.8–59.85 | 4.19 | 78.08 | 25.21% | 85.41** | 1 | 0.36*** | 0.05 |
| | GZnC | 49.29 ± 6.51 | 21.85–71.71 | 4.73 | 87.14 | 13.25% | 26.89** | | 1 | −0.16** |
| | TGW | 36.28 ± 4.03 | 23.63–48.62 | 5.82 | 73.53 | −06.54% | $0.03^{NS}$ | | | 1 |

*$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

TABLE 2 Significant marker–trait associations for GFeC and GZnC along with the details of trait-increasing and trait-decreasing alleles.

| Trait | SNP | Chromosome | Position (MB) | *p*-value | *−log10(p)* | Decreasing allele | Increasing allele |
|---|---|---|---|---|---|---|---|
| Fe_IR | AX-94537892 | chr2B | 466.21 | 0.000506 | 3.295525 | A | G |
| Fe_IR | AX-95241551 | chr3A | 31.95 | 0.000651 | 3.186231 | C | T |
| Fe_LS | AX-95166268 | chr2D | 601.18 | 0.000241 | 3.617943 | T | C |
| Fe_LS | AX-94940971 | chr2B | 33.84 | 0.000383 | 3.417198 | C | T |
| Fe_LS | AX-95078562 | chr5D | 385.39 | 0.000491 | 3.309289 | C | T |
| Fe_RI | AX-94825289 | chr3D | 559.08 | 1.90E-07 | 6.720941 | A | T |
| Fe_RI | AX-94942225 | chr2A | 758.39 | 0.000587 | 3.231101 | C | T |
| Zn_IR | AX-94527403 | chr3B | 802.60 | 7.08E-05 | 4.149945 | C | A |
| Zn_IR | AX-95017992 | chr2D | 8.93 | 0.000236 | 3.627449 | A | C |
| Zn_IR | AX-94428612 | chr5A | 677.13 | 0.000255 | 3.592764 | G | A |
| Zn_IR | AX-94987428 | chr1B | 604.15 | 0.000308 | 3.510902 | C | T |
| Zn_IR | AX-94972272 | chr7B | 0.27 | 0.000342 | 3.466551 | A | G |
| Zn_IR | AX-95250489 | chr5A | 2.41 | 0.000528 | 3.276998 | A | G |
| Zn_IR | AX-94425009 | chr1B | 41.09 | 0.000548 | 3.26091 | A | G |
| Zn_IR | AX-94476475 | chr5B | 712.86 | 0.000685 | 3.164529 | T | C |
| Zn_IR | AX-94527853 | chr2B | 180.56 | 0.000721 | 3.14236 | G | A |
| Zn_IR | AX-94724239 | chr2D | 9.35 | 0.000902 | 3.044792 | G | A |
| Zn_LS | AX-95107190 | chr5D | 8.35 | 0.000812 | 3.090692 | A | C |
| Zn_RI | AX-94495098 | chr7B | 632.80 | 0.000164 | 3.785761 | A | G |
| Zn_RI | AX-94961429 | chr7B | 3.80 | 0.000518 | 3.285462 | T | A |
| Zn_RI | AX-94702925 | chr7D | 552.93 | 0.000625 | 3.203826 | C | T |
| Zn_RI | AX-94925767 | chr1B | 4.46 | 0.000893 | 3.048914 | G | C |
| Zn_RI | AX-94516789 | chr3B | 6.56 | 0.000901 | 3.045095 | C | T |

(LD) decay block size was shown to be 5.24, 5.26, and 9.22 MB for the A, B, and D genomes, respectively, and a block of 7.15 MB was observed for the whole genome (Devate et al., 2022). The effect of the population structure of the association panel on the GWAS analysis is addressed by taking them as a covariate to ensure true association. LD decay is used to determine the number of markers required to be used in the GWAS. The rate of LD decay varies depending on the rate of recombination

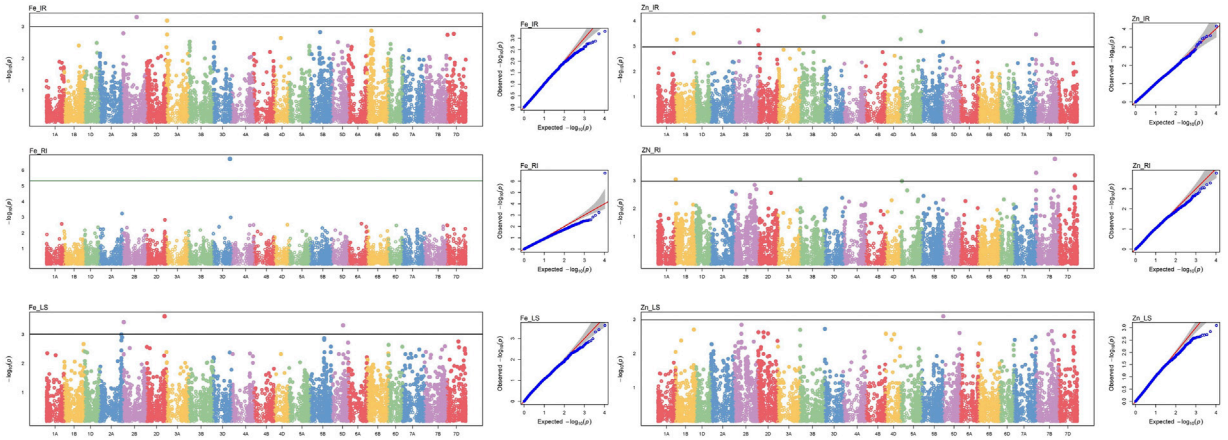



FIGURE 2
Manhattan and respective Q–Q plots of significant associations for GFeC and GZnC under IR, RI, and LS conditions.

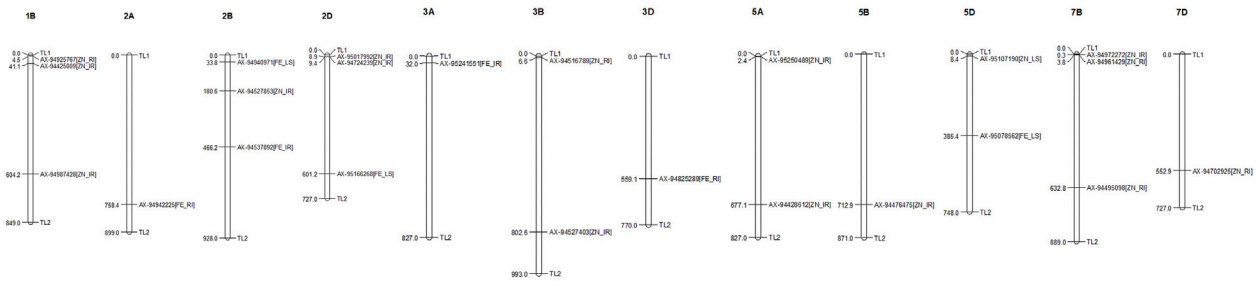



FIGURE 3
Distribution and position (in Mb) of identified MTAs at their respective chromosomes with associated traits.

between the marker pairs (Borah et al., 2018). In general, in outcrossing crop species such as maize, LD decays rapidly, in contrast to self-pollinated crops, which show slow decay, as in wheat (Yu et al., 2014; Roncallo et al., 2021).

Genome-wide 10,546 SNPs over 282 genotypes were used to identify the markers associated with grain Fe and Zn contents across control (IR), drought (RI), and heat stress (LS) treatments. Of the 23 MTAs identified, seven MTAs were detected for GFeC, and 16 for GZnC (Table 2) can be visualized by a Manhattan plot with the threshold at $p < 0.001$ (Figure 2) and their position on respective chromosomes in Figure 3. Trait-increasing allele and trait-decreasing allele for each associated marker are given in Table 2. The identified MTAs were treatment-specific, implicating differential expression of genes under different stress conditions. The putative candidate genes located within the 100-kb region of the linked MTAs were identified by a BLAST search against the IWGSC reference genome at the Ensembl Plants database.

The MTAs identified for grain Fe content under IR were located on chromosomes 2B and 3A at the positions of 466.2 and 31.9 mb, respectively, in concordance with those previously noted by Rathan et al. (2022) on chromosomes 2B and 3A and on 3A by Bhatta et al. (2018). A putative candidate gene (*TraesCS2B02G326400*) near marker AX-94537892 codes for the "*auxin response factor*," which has a role in various growth and development and response to abiotic stresses, such as drought, salt, or cold (Xu et al., 2020). The same region contains another protein-coding region (*TraesCS2B02G326300*) for "*cleft lip and palate transmembrane 1*," which is less studied in plants. Putative genes near the other SNP marker AX-95241551 were the "*WD40/YVTN repeat-like-containing domain superfamily*" and "*root UVB sensitive family*," the roles of which in supramolecular interactions (Guerriero et al., 2015) and sunlight-dependent

TABLE 3 Putative candidate genes in the 100-kb region of the linked marker with the protein produced from them.

| Trait | SNP | Position | Transcript ID | Protein |
|---|---|---|---|---|
| FE_IR | AX-94537892 | 2B: 474382601..474388204 (−strand) | TraesCS2B02G326400 | ≻ Auxin response factor |
| | | 2B: 474355042..474361352 (−strand) | TraesCS2B02G326300 | ≻ Cleft lip and palate transmembrane 1 |
| FE_IR | AX-95241551 | 3A: 32643085..32647110 (+strand) | TraesCS3A02G054700 | ≻ WD40/YVTN repeat-like-containing domain superfamily |
| | | 3A: 32647241..32651665 (−strand) | TraesCS3A02G054800 | ≻ Root UVB sensitive family |
| | | 3A: 32621252..32625702 (+strand) | TraesCS3A02G054600 | ≻ P-loop containing nucleoside triphosphate hydrolase |
| FE_LS | AX-95166268 | 2D: 604123946..604130138 (+strand) | TraesCS2D02G507500 | ≻ Staphylococcal nuclease (SNase-like), OB-fold |
| | | 2D: 604146523..604150893 (+strand) | TraesCS2D02G507600 | ≻ Peptidyl-tRNA hydrolase, PTH2 |
| | | 2D: 604151044..604155857 (−strand) | TraesCS2D02G507700 | ≻ Phospholipase C/P1 nuclease domain superfamily |
| FE_LS | AX-94940971 | 2B: 38339312..38342120 (−strand) | TraesCS2B02G067000 | ≻ Bulb-type lectin domain superfamily |
| | | 2B: 38348364..38351197 (+strand) | TraesCS2B02G067100 | ≻ FKBP-type peptidyl-prolyl cis-trans isomerase domain |
| FE_LS | AX-95078562 | 5D: 388220089..388223638 (+strand) | TraesCS5D02G284600 | ≻ Gnk2-homologous domain |
| | | 5D: 388268407..388270613 (−strand) | TraesCS5D02G284700 | ≻ Tetratricopeptide-like helical domain superfamily |
| FE_RI | AX-94825289 | 3D: 560443926..560449941 (+strand) | TraesCS3D02G450800 | ≻ Potassium transporter |
| | | 3D: 560449825..560456397 (−strand) | TraesCS3D02G450900 | ≻ Spermatogenesis-associated protein 20 |
| FE_RI | AX-94942225 | 2A: 762552292..762554307 (−strand) | TraesCS2A02G551800 | ≻ Tetratricopeptide-like helical domain superfamily |
| | | 2A: 762554483..762556597 (+strand) | TraesCS2A02G551900 | ≻ Ubiquinone biosynthesis protein Coq4 |
| | | 2A: 762556814..762564083 (−strand) | TraesCS2A02G552000 | ≻ Clathrin, heavy chain/VPS, 7-fold repeat |
| | | 2A: 762577383..762581944 (+strand) | TraesCS2A02G552100 | ≻ Tetratricopeptide-like helical domain superfamily |
| | | 2A: 762528939..762530243 (−strand) | TraesCS2A02G551700 | ≻ F-box associated interaction domain |
| ZN_IR | AX-94527403 | 3B: 820348909..820353167 (−strand) | TraesCS3B02G570600 | ≻ Protein kinase-like domain superfamily |
| | | 3B: 820363175..820364327 (−strand) | TraesCS3B02G570500 | ≻ Myb/SANT-like domain |
| ZN_IR | AX-95017992 | 2A: 9870676..9873085 (−strand) | TraesCS2A02G017600 | ≻ Domain of unknown function DUF1618 |
| | | 2A: 9864423..9870187 (+strand) | TraesCS2A02G017500 | ≻ 2-Oxoacid dehydrogenase acyltransferase, catalytic domain |
| | | 2A: 9838071..9840946 (+strand) | TraesCS2A02G017400 | ≻ FACT complex subunit Spt16 N-terminal lobe domain |
| | | 2A: 9831458..9833452 (−strand) | TraesCS2A02G017300 | ≻ FBD domain |
| ZN_IR | AX-94428612 | 5A: 678935942..678937957 (−strand) | TraesCS5A02G513300 | ≻ Ribosomal protein L21e |
| | | 5A: 678941500..678946425 (−strand) | TraesCS5A02G513400 | ≻ Methyltransferase type 12 |
| | | 5A: 678965411..678966822 (+strand) | TraesCS5A02G513500 | ≻ Ribosomal protein L5 domain superfamily |
| | | 5A: 678966202..678969731 (−strand) | TraesCS5A02G513600 | ≻ Pectin lyase fold/virulence factor |
| ZN_IR | AX-94987428 | 1B: 611299086..611301010 (−strand) | TraesCS1B02G373600 | ≻ 6-Phosphogluconate dehydrogenase, decarboxylating |
| ZN_IR | AX-94972272 | 7B: 265308..267721 (−strand) | TraesCS7B02G001200 | ≻ F-box-like domain superfamily |
| | | 7B: 261777..263347 (+strand) | TraesCS7B02G001100 | ≻ NTF2-like domain superfamily |
| ZN_IR | AX-95250489 | 5A: 3318157..3321929 (−strand) | TraesCS5A02G003200 | ≻ Glycosyl transferase, family 14 |
| | | 5A: 3315802..3317881 (−strand) | TraesCS5A02G003100 | ≻ Longin-like domain superfamily |
| | | 5A: 3312509..3313435 (+strand) | TraesCS5A02G003000 | ≻ Chaperone J-domain superfamily |
| | | 5A: 3306994..3308561 (−strand) | TraesCS5A02G002900 | ≻ Chalcone/stilbene synthase, N-terminal |
| ZN_IR | AX-94425009 | 1B: 45453621..45459411 (−strand) | TraesCS1B02G058700 | ≻ Beta-hexosaminidase, eukaryotic type, N-terminal |
| | | 1B: 45447127..45450887 (−strand) | TraesCS1B02G058600 | ≻ Protein-tyrosine phosphatase-like |
| ZN_IR | AX-94476475 | 5B: 714519498..714521198 (−strand) | TraesCS5B02G571900 | ≻ NTF2-like domain superfamily |
| | | 5B: 714565705..714567173 (−strand) | TraesCS5B02G572400 | ≻ Glutathione S-transferase, N-terminal |
| ZN_IR | AX-94527853 | 2B: 188627646..188629785 (+strand) | TraesCS2B02G201400 | ≻ COBRA and plant |
| | | 2B: 188674693..188676234 (+strand) | TraesCS2B02G201600 | ≻ WD40-repeat-containing domain superfamily |
| ZN_IR | AX-94724239 | 2D: 9539575..9541245 (+strand) | TraesCS2D02G019400 | ≻ P-loop containing nucleoside triphosphate hydrolase |
| | | 2D: 9560490..9570556 (+strand) | TraesCS2D02G019600 | ≻ Protein kinase-like domain superfamily |
| ZN_LS | AX-95107190 | 5D: 8234067..8262668 (+strand) | TraesCS5D02G014500 | ≻ Ankyrin repeat-containing domain superfamily |
| ZN_RI | AX-94495098 | 7B: 632793597..632795681 (−strand) | TraesCS7B02G368500 | ≻ Pentatricopeptide repeat |
| | | 7B: 632799627..632800394 (−strand) | TraesCS7B02G368600 | ≻ TFIIS/LEDGF domain superfamily |
| ZN_RI | AX-94961429 | 7B: 3789679..3794857 (+strand) | TraesCS7B02G006700 | ≻ Ubiquinone biosynthesis O-methyltransferase |

(Continued on following page)

TABLE 3 (*Continued*) Putative candidate genes in the 100-kb region of the linked marker with the protein produced from them.

| Trait | SNP | Position | Transcript ID | Protein |
|---|---|---|---|---|
| | | 7B: 3795056..3797564 (−strand) | TraesCS7B02G006800 | ➢ Transcription elongation factor 1 |
| ZN_RI | AX-94702925 | 7D: 552946865..552951091 (+strand) | TraesCS7D02G433000 | ➢ Putative S-adenosyl-L-methionine-dependent methyltransferase |
| | | 7D: 552961791..552966500 (+strand) | TraesCS7D02G433100 | ➢ GDP-fucose protein O-fucosyltransferase |
| ZN_RI | AX-94925767 | 1B: 4475908..4479190 (+strand) | TraesCS1B02G008200 | ➢ Snf7 family |
| | | 1B: 4480185..4481605 (−strand) | TraesCS1B02G008300 | ➢ Protein kinase-like domain superfamily |
| ZN_RI | AX-94516789 | 3B: 6554790..6558914 (+strand) | TraesCS3B02G016000 | ➢ Fatty acyl-coenzyme A reductase, NAD-binding domain |
| | | 3B: 6533011..6534111 (+strand) | TraesCS3B02G015800 | ➢Phosphatidylethanolamine-binding protein |

root development (Tong et al., 2008) were confirmed by previous studies, respectively. As protein interaction and root architecture are important criteria for nutrient uptake, they might be linked with the grain Fe content. Similarly, under drought stress conditions, AX-94825289 and AX-94942225 were identified on chromosomes 3D and 2A, respectively, and are linked with GFeC, present in the region having putative candidate genes for "*potassium transporter*" and "*ubiquinone biosynthesis protein.*" Potassium transporter has a role in concentration gradients of protons, homeostasis of monovalent cations (Gierth and Mäser, 2007), and symporter and antiporter transport of the ions, whereas ubiquinone protects the membrane from free radical-induced oxidative damage (Pobezhimova and Voinikov, 2000), which is commonly observed under stress. Three unique MTAs were identified under LS conditions for grain iron content on chromosomes 2D, 2B, and 5D. They were putatively linked to transcripts having a role in protein modification (*peptidyl-tRNA hydrolase* and *peptidyl-prolyl cis-trans isomerase*), carbohydrate binding (*bulb-type lectin domain*), and enzymes involved in lipid signaling pathways (*phospholipase C/P1 nuclease domain*) (Table 3).

GZnC under irrigated conditions was linked with 10 different SNPs which were located on 1B, 2B, 2D, 3B, 5A, 5B, and 7B. Previous studies reported stably expressing MTAs on 5B (Cu et al., 2020), 7B (Rathan et al., 2022), 3B, and 5A (Alomari et al., 2018). Genomic regions of GZnC-linked SNPs of the current study contained important genes coding for *glycosyl transferase (TraesCS5A02G003200)*, *COBRA*, *plant (TraesCS2B02G201400)*, *WD40-repeat-containing domain superfamily (TraesCS2B02G201600)*, and *P-loop containing nucleoside triphosphate hydrolase (TraesCS2D02G019400*), whose roles in flavonoid biosynthesis (Yao et al., 2019), orientation, and cell expansion in Arabidopsis root (Roudier et al., 2002), supramolecular interactions (Guerriero et al., 2015), and metallochaperones and metalloenzymes (Vaccaro and Drennan, 2022) were reported in previous studies. Flavonoids play a crucial role in the growth and development of plants, which might have an influence on the quality of the product and its nutrient status. Similarly, nucleoside triphosphate hydrolase plays a crucial role as a metallochaperone, which might have a direct or indirect role in the GZnC of wheat. Apart from these, *protein kinase*, *Myb/SANT-like domain*, *FACT complex*, *FBD domain*, *ribosomal protein L21e*, *methyltransferase type 12*, *ribosomal protein L5 domain*, and *F-box-like domain* are other important genes present in the candidate regions (Table 3).

Grain Zn content under drought stress was linked with SNP markers located on 1B, 3B, 7B, and 7D. A putative candidate gene near the linked SNP AX-94495098 was *TraesCS7B02G368500*, which codes for "*pentatricopeptide repeat*," which controls post-transcriptional regulation of many genes at the RNA level (Zhang et al., 2020). Other two MTAs, AX-94961429 and AX-94702925, were linked to the different genes that code for *methyltransferase*, which has a crucial role in biochemical reaction and plant metabolism (Moffatt and Weretilnyk, 2001). The *ubiquinone biosynthesis* gene was found to be linked in RI condition with GZnC like GFeC as mentioned earlier and has a crucial role in managing free radical-induced oxidative damage. Marker AX-94516789 is linked with the gene coding for "*phosphatidylethanolamine-binding protein*" and has a role in the morphological switch between shoot and flower structure and signal transduction (Banfield and Brady, 2000).

Under heat stress, i.e., LS condition, only one marker (AX-95107190) on 5D showed association with GZnC. A candidate gene at the genomic region of SNP was found to be coded for an *ankyrin repeat-containing domain*, whose critical role in plant growth and development, hormone response, and response to biotic and abiotic stresses was discovered in previous studies (Lopez-Ortiz, et al., 2020). Stress tolerance and nutrient assimilation in the plant involve many complex pathways governed by many related and superfamily genes. Putative candidate genes identified here have a direct or indirect influence on various plant growth and development processes and may play a role in nutrient uptake and grain nutrient content.

## Conclusion

The GWAS panel used in this study with 282 advanced breeding lines of wheat has shown that GFeC, GZnC, and TGW

were complex traits, inherited quantitatively, and their expression was highly influenced by abiotic stress factors such as drought and heat stress. A positive correlation between the GFeC and GZnC and their high heritability indicate that simultaneous improvement of both the traits can be possible. Out of 23 MTAs identified under IR, RI, and LS conditions, seven were linked to GFeC and sixteen were linked to GZnC. The identified MTAs were located near novel candidate genes and have a direct or indirect effect on traits. Several identified putative candidate genes encode important molecular functions such as metallochaperones, root architecture orientation, ionic homeostasis, and abiotic stress response. Further validation of identified MTAs can be carried out and is useful in marker-assisted selection programs to develop biofortified varieties.

## Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: https://doi.org/10.5061/dryad.qnk98sfkw.

## Author contributions

PS, NJ, GS, and HK conceptualized the investigation and edited the manuscript. PS supervised the conduction of the experiment. ND conducted the investigation and prepared the draft of the manuscript. ND, CM, and KM generated the phenotypic data. ND, HK, and SP contributed to the generation of genotyping data. ND and HK did the statistical and GWAS analysis. All authors contributed to the article and approved the submitted version.

## Funding

Part of the research was supported by a grant from the Bill & Melinda Gates Foundation (Grant number # OPP1215722) subgrant to India for the Zn mainstreaming project and the ICAR -BMGF project (Grant number: OPP1194767).

## Acknowledgments

NB acknowledges the Council of Scientific and Industrial Research (CSIR), New Delhi, and the ICAR-Indian Agricultural Research Institute (IARI), New Delhi, for scholarships to complete this work as part of PhD thesis.

## Conflict of interest

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

The handling editor DM declared a past co-authorship with the author GS.

## Publisher's note



## References

Alomari, D. Z., Eggert, K., Von Wiren, N., Alqudah, A. M., Polley, A., Plieske, J., et al. (2018). Identifying candidate genes for enhancing grain Zn concentration in wheat. *Front. Plant Sci.* 9, 1313. doi:10.3389/fpls.2018.01313

Aravind, J., Mukesh Sankar, S., Wankhede, D. P., and Kaur, V. (2021). augmentedRCBD: Analysis of augmented randomised complete block designs. Available at https://aravind-j.github.io/augmentedRCBD/https://cran.rproject.org/package=augmentedRCBD.

Banfield, M. J., and Brady, R. L. (2000). The structure of Antirrhinum centroradialis protein (CEN) suggests a role as a kinase regulator. *J. Mol. Biol.* 297 (5), 1159–1170. doi:10.1006/jmbi.2000.3619

Bhatta, M., Baenziger, P. S., Waters, B. M., Poudel, R., Belamkar, V., Poland, J., et al. (2018). Genome-wide association study reveals novel genomic regions associated with 10 grain minerals in synthetic hexaploid wheat. *Int. J. Mol. Sci.* 19 (10), 3237. doi:10.3390/ijms19103237

Borah, J., Singode, A., Talukdar, A., Yadav, R. R., and Sarma, R. N. (2018). Genome-wide association studies (GWAS) reveal candidate genes for plant height and number of primary branches in soybean [Glycine max (L.) Merrill]. *Indian J. Genet. Plant Breed.* 78 (04), 460–469.

Bouis, H. E., Hotz, C., McClafferty, B., Meenakshi, J. V., and Pfeiffer, W. H. (2011). Biofortification: A new tool to reduce micronutrient malnutrition. *Food Nutr. Bull.* 32, S31–S40. doi:10.1177/15648265110321S105

Cu, S. T., Guild, G., Nicolson, A., Velu, G., Singh, R., and Stangoulis, J. (2020). Genetic dissection of zinc, iron, copper, manganese and phosphorus in wheat (*Triticum aestivum* L.) grain and rachis at two developmental stages. *Plant Sci.* 291, 110338. doi:10.1016/j.plantsci.2019.110338

Devate, N. B., Krishna, H., Parmeshwarappa, S. K. V., Manjunath, K. K., Chauhan, D., Singh, S., et al. (2022). Genome wide association mapping for component traits of drought and heat tolerance in wheat. *Front. Plant Sci.* 13, 943033. doi:10.3389/fpls.2022.943033

El-Soda, M., and Aljabri, M. (2022). Genome-wide association mapping of grain metal accumulation in wheat. *Genes.* 13 (6), 1052. doi:10.3390/genes13061052

Gierth, M., and Mäser, P. (2007). Potassium transporters in plants–involvement in K+ acquisition, redistribution and homeostasis. *FEBS Lett.* 581 (12), 2348–2356. doi:10.1016/j.febslet.2007.03.035

Guerriero, G., Hausman, J. F., and Ezcurra, I. (2015). WD40-repeat proteins in plant cell wall formation: Current evidence and research prospects. *Front. Plant Sci.* 6, 1112. doi:10.3389/fpls.2015.01112

Gupta, P. K., Balyan, H. S., Sharma, S., and Kumar, R. (2021). Biofortification and bioavailability of Zn, Fe and Se in wheat: Present status and future prospects. *Theor. Appl. Genet.* 134 (1), 1–35. doi:10.1007/s00122-020-03709-7

Huang, M., Liu, X., Zhou, Y., Summers, R. M., and Zhang, Z. (2019). Blink: A package for the next level of genome wide association studies with both individuals

and markers in the millions. *Gigascience* 8 (2), giy154. doi:10.1093/gigascience/giy154

Khaled, K. A. M., Habiba, R. M. M., Bashasha, J. A., and El-Aziz, M. H. A. (2022). Identification and mapping of QTL associated with some traits related for drought tolerance in wheat using SSR markers. *Beni. Suef. Univ. J. Basic Appl. Sci.* 11 (1), 38–13. doi:10.1186/s43088-022-00212-4

Krishnappa, G., Khan, H., Krishna, H., Kumar, S., Mishra, C. N., Parkash, O., et al. (2022). Genetic dissection of grain iron and zinc, and thousand kernel weight in wheat (*Triticum aestivum* L.) using genome-wide association study. *Sci. Rep.* 12 (1), 12444–12514. doi:10.1038/s41598-022-15992-z

Krishnappa, G., Singh, A. M., Chaudhary, S., Ahlawat, A. K., Singh, S. K., Shukla, R. B., et al. (2017). Molecular mapping of the grain iron and zinc concentration, protein content and thousand kernel weight in wheat (*Triticum aestivum* L.). *PLoS One* 12 (4), e0174972. doi:10.1371/journal.pone.0174972

Listman, M., and Ordóñez, R. (2019). *Ten things you should know about maize and wheat*. Mexico: CIMMYT: Batan.

Liu, J., Wu, B., Singh, R. P., and Velu, G. (2019). QTL mapping for micronutrients concentration and yield component traits in a hexaploid wheat mapping population. *J. cereal Sci.* 88, 57–64. doi:10.1016/j.jcs.2019.05.008

Lobell, D. B., Schlenker, W., and Costa-Roberts, J. (2011). Climate trends and global crop production since 1980. *Science* 333 (6042), 616–620. doi:10.1126/science.1204531

Lopez-Ortiz, C., Peña-Garcia, Y., Natarajan, P., Bhandari, M., Abburi, V., Dutta, S. K., et al. (2020). The ankyrin repeat gene family in Capsicum spp: Genome-wide survey, characterization and gene expression profile. *Sci. Rep.* 10 (1), 4044–4116. doi:10.1038/s41598-020-61057-4

Moffatt, B. A., and Weretilnyk, E. A. (2001). Sustaining S-adenosyl-l-methionine-dependent methyltransferase activity in plant cells. *Physiol. Plant.* 113 (4), 435–442. doi:10.1034/j.1399-3054.2001.1130401.x

Murray, M. G., and Thompson, W. F. (1980). Rapid isolation of high molecularweight plant DNA. *Nucleic Acids Res.* 8, 4321–4325. doi:10.1093/nar/8.19.4321

Paltridge, N. G., Milham, P. J., Ortiz-Monasterio, J. I., Velu, G., Yasmin, Z., Palmer, L. J., et al. (2012). Energy-dispersive X-ray fluorescence spectrometry as a tool for zinc, iron and selenium analysis in whole grain wheat. *Plant Soil* 361 (1), 261–269. doi:10.1007/s11104-012-1423-0

Pfeiffer, W. H., and McClafferty, B. (2007). HarvestPlus: Breeding crops for better nutrition. *Crop Sci.* 47, S-88–S-105. doi:10.2135/cropsci2007.09.0020ipbs

Pobezhimova, T. P., and Voinikov, V. K. (2000). Biochemical and physiological aspects of ubiquinone function. *Membr. Cell. Biol.* 13 (5), 595–602.

Rathan, N. D., Krishna, H., Ellur, R. K., Sehgal, D., Govindan, V., Ahlawat, A. K., et al. (2022). Genome-wide association study identifies loci and candidate genes for grain micronutrients and quality traits in wheat (*Triticum aestivum* L.). *Sci. Rep.* 12 (1), 7037–7115. doi:10.1038/s41598-022-10618-w

Roncallo, P. F., Larsen, A. O., Achilli, A. L., Pierre, C. S., Gallo, C. A., Dreisigacker, S., et al. (2021). Linkage disequilibrium patterns, population structure and diversity analysis in a worldwide durum wheat collection including Argentinian genotypes. *BMC genomics* 22 (1), 233–317. doi:10.1186/s12864-021-07519-z

Roudier, F., Schindelman, G., DeSalle, R., and Benfey, P. N. (2002). The COBRA family of putative GPI-anchored proteins in Arabidopsis. A new fellowship in expansion. *Plant Physiol.* 130 (2), 538–548. doi:10.1104/pp.007468

Samineni, S., Mahendrakar, M. D., Hotti, A., Chand, U., Rathore, A., Gaur, P. M., et al. (2022). Impact of heat and drought stresses on grain nutrient content in chickpea: Genome-wide marker-trait associations for protein, Fe and Zn. *Environ. Exp. Bot.* 194, 104688. doi:10.1016/j.envexpbot.2021.104688

Srinivasa, J., Arun, B., Mishra, V. K., Singh, G. P., Velu, G., Babu, R., et al. (2014). Zinc and iron concentration QTL mapped in a Triticum spelta× *T. aestivum* cross. *Theor. Appl. Genet.* 127 (7), 1643–1651. doi:10.1007/s00122-014-2327-6

Tiwari, C., Wallwork, H., Arun, B., Mishra, V. K., Velu, G., Stangoulis, J., et al. (2016). Molecular mapping of quantitative trait loci for zinc, iron and protein content in the grains of hexaploid wheat. *Euphytica* 207 (3), 563–570. doi:10.1007/s10681-015-1544-7

Tong, H., Leasure, C. D., Hou, X., Yuen, G., Briggs, W., and He, Z. H. (2008). Role of root UV-B sensing in Arabidopsis early seedling development. *Proc. Natl. Acad. Sci. U. S. A.* 105 (52), 21039–21044. doi:10.1073/pnas.0809942106

Vaccaro, F. A., and Drennan, C. L. (2022). The role of nucleoside triphosphate hydrolase metallochaperones in making metalloenzymes. *Metallomics.* 14 (6), mfac030. doi:10.1093/mtomcs/mfac030

Velu, G., Guzman, C., Mondal, S., Autrique, J. E., Huerta, J., and Singh, R. P. (2016). Effect of drought and elevated temperature on grain zinc and iron concentrations in CIMMYT spring wheat. *J. Cereal Sci.* 69, 182–186. doi:10.1016/j.jcs.2016.03.006

Velu, G., Singh, R. P., Crespo-Herrera, L., Juliana, P., Dreisigacker, S., Valluru, R., et al. (2018). Genetic dissection of grain zinc concentration in spring wheat for mainstreaming biofortification in CIMMYT wheat breeding. *Sci. Rep.* 8 (1), 13526–13610. doi:10.1038/s41598-018-31951-z

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

Xu, L., Wang, D., Liu, S., Fang, Z., Su, S., Guo, C., et al. (2020). Comprehensive atlas of wheat (*Triticum aestivum* L.) AUXIN RESPONSE FACTOR expression during male reproductive development and abiotic stress. *Front. Plant Sci.* 11, 586144. doi:10.3389/fpls.2020.586144

Yao, P., Deng, R., Huang, Y., Stael, S., Shi, J., Shi, G., et al. (2019). Diverse biological effects of glycosyltransferase genes from Tartary buckwheat. *BMC Plant Biol.* 19 (1), 339–415. doi:10.1186/s12870-019-1955-z

Yu, H., Deng, Z., Xiang, C., and Tian, J. (2014). Analysis of diversity and linkage disequilibrium mapping of agronomic traits on B-genome of wheat. *J. Genomics* 2, 20–30. doi:10.7150/jgen.4089

Zahra, N., Wahid, A., Hafeez, M. B., Ullah, A., Siddique, K. H., and Farooq, M. (2021). Grain development in wheat under combined heat and drought stress: Plant responses and management. *Environ. Exp. Bot.* 188, 104517. doi:10.1016/j.envexpbot.2021.104517

Zhang, Q., Xu, Y., Huang, J., Zhang, K., Xiao, H., Qin, X., et al. (2020). The rice pentatricopeptide repeat protein PPR756 is involved in pollen development by affecting multiple RNA editing in mitochondria. *Front. Plant Sci.* 11, 749. doi:10.3389/fpls.2020.00749

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 11 November 2022
DOI 10.3389/fgene.2022.1004211



OPEN ACCESS

EDITED BY
Dwijesh Chandra Mishra,
Indian Agricultural Statistics Research Institute (ICAR) India

REVIEWED BY
Xiaoli Fan,
Chengdu Institute of Biology (CAS), China
Thanda Dhliwayo,
International Maize and Wheat Improvement Center, Mexico
Sudha K. Nair,
The International Maize and Wheat Improvement Center (CIMMYT), India

*CORRESPONDENCE
Xuehai Zhang,
xuehai85@126.com
Jihua Tang,
tangjihua1@163.com

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

RECEIVED 27 July 2022
ACCEPTED 28 October 2022
PUBLISHED 11 November 2022

CITATION
Duan H, Li J, Sun Y, Xiong X, Sun L, Li W, Gao J, Li N, Zhang J, Cui J, Fu Z, Zhang X and Tang J (2022), Candidate loci for leaf angle in maize revealed by a combination of genome-wide association study and meta-analysis. *Front. Genet.* 13:1004211. doi: 10.3389/fgene.2022.1004211



# Candidate loci for leaf angle in maize revealed by a combination of genome-wide association study and meta-analysis

Haiyang Duan[1], Jianxin Li[1], Yan Sun[1], Xuehang Xiong[1], Li Sun[1], Wenlong Li[1], Jionghao Gao[1], Na Li[1], Junli Zhang[1,2], Jiangkuan Cui[3], Zhiyuan Fu[1], Xuehai Zhang[1]* and Jihua Tang[1,4]*

[1]National Key Laboratory of Wheat and Maize Crop Science, College of Agronomy, Henan Agricultural University, Zhengzhou, China, [2]State Key Laboratory of Crop Stress Adaptation and Improvement, School of Life Sciences, Henan University, Kaifeng, China, [3]College of Plant Protection, Henan Agricultural University, Zhengzhou, China, [4]The Shennong Laboratory, Zhengzhou, China

Leaf angle (LA) is a key component of maize plant architecture that can simultaneously govern planting density and improve final yield. However, the genetic mechanisms underlying LA have not been fully addressed. To broaden our understanding of its genetic basis, we scored three LA-related traits on upper, middle, and low leaves of 492 maize inbred lines in five environments. Phenotypic data revealed that the three LA-related traits were normally distributed, and significant variation was observed among environments and genotypes. A genome-wide association study (GWAS) was then performed to dissect the genetic factors that control natural variation in maize LA. In total, 85 significant SNPs (involving 32 non-redundant QTLs) were detected ($p \leq 2.04 \times 10^{-6}$), and individual QTL explained 4.80%–24.09% of the phenotypic variation. Five co-located QTL were detected in at least two environments, and two QTLs were co-located with multiple LA-related traits. Forty-seven meta-QTLs were identified based on meta-analysis combing 294 LA-related QTLs extracted from 18 previously published studies, 816 genes were identified within these meta-QTLs, and seven co-located QTLs were jointly identified by both GWAS and meta-analysis. *ZmULA1* was located in one of the co-located QTLs, *qLA7*, and its haplotypes, hap1 and hap2, differed significantly in LA-related traits. Interestingly, the temperate materials with hap2 had smallest LA. Finally, we also performed haplotype analysis using the reported genes that regulate LA, and identified a lot of maize germplasms that aggregated favorable haplotypes. These results will be helpful for elucidating the genetic basis of LA and breeding new maize varieties with ideal plant architecture.

KEYWORDS

**maize, leaf angle, genome-wide association study, meta-analysis, ideal plant architecture**

## Introduction

Leaf angle (LA) is one of the key traits of maize plant architecture. Upright leaves can make maize plants more compact, enable them to adapt to high planting density, and reduce effects of shading between plants, thereby increasing photosynthetic efficiency and final grain yield (Stewart et al., 2003). Maize yield in the United States has increased eight-fold in the past 90 years. During this period, changes in maize LA have altered plant architecture, resulting in new varieties that are tolerant of high planting densities, allowing for more efficient light capture as planting density has increased (Tian et al., 2011). Compact plant architecture with upright leaves can increase the ability of crops to capture light energy and increase final yield (Saitoh et al., 2002).

LA is one of the important indicators used to describe maize canopy structure (Arkebauer et al., 2009) and is a quantitative trait. Thirty QTLs related to upper leaf angle were identified in a Nested Association Mapping Population (NAM) composed of 25 recombinant inbred line (RIL) families, including 4,892 lines (Tian et al., 2011). A maize $BC_2S_3$ population was constructed from teosinte and the maize inbred line W22, 12 QTLs underlying LA were identified. *UPA1*-NIL$^{W22}$ (Near Isogenic Line) has smaller LA (Upper, Middle and Low LA) than *UPA1*-NIL$^{8759}$. On the contrary, *UPA2*-NIL$^{W22}$ has larger LA (Upper, Middle and Low LA) than *UPA2*-NIL$^{8759}$. These results indicated that *UPA1* and *UPA2* have opposite regulatory patterns, and with the increase of planting density, *UPA2*-NIL$^{8759}$ has higher grain yield (kg/ha) than *UPA2*-NIL$^{W22}$ (Tian et al., 2019). Several genes within the quantitative trait loci (QTL) that govern LA have been cloned. *ZmILI1* was located within the *qLA2*, which was identified in a maize $F_{2:3}$ population constructed from the compact inbred line Yu82 and the expanded inbred line Yu87-1. *ZmILI1* (*Zm00001d002121*) binds to the promoter of *CYP90D1* (*Zm00001d039453*) and inhibits expression of *CYP90D1* through a *CYP90D1*-mediated cytochrome P450-catalyzed reaction, Brassinosteroids (BRs) can be delivered to the BR receptor *BAK1*, which binds to the nuclear transcription factors *BZR1*, *BZR2* and *BES1*, resulting in a smaller LA (Upper LA) (Tang et al., 2010; Ren et al., 2020). Li et al. found that the rice gene *oslazy1* regulates LA by affecting gravitropism (Li et al., 2007), its maize homolog *ZmCLA4* regulates LA by changing mRNA accumulation, leading to changes in gravitational properties and cell development (Zhang et al., 2014). Gao et al. used a *Gmilpa1* mutant with increased soybean petiole angle and isolated the gene *GmILPA1*, which encodes the *APC8* protein. It is expressed in leaf primordium cells and can increase petiole angle by promoting the growth and division of leaf occipital cells (Gao et al., 2017).

Genome-wide association study (GWAS) has been widely used in plant genetics research and has proven to be a powerful tool for mining QTLs or causal genes for complex quantitative traits. The first GWAS report in plants was published in 2008. Single nucleotide polymorphism (SNP) haplotypes at 8,590 loci across 10 maize chromosomes were tested for association with kernel oleic acid content in 553 maize inbred lines, and a putative gene responsible for the target trait (*fad2*) was identified (Belo et al., 2008). Recently, Wang et al. performed whole-genome sequencing on 350 elite maize inbred lines representing multiple eras of germplasms from China and the United States and measured 15 agronomic traits for GWAS. A number of key candidate genes, such as *ZmNAC16*, *ZmSBP18*, *ZmPIF4*, and *ZmPIF3.3*, that regulate maize density tolerance and ideal plant architecture were cloned, providing an important foundation for future genomics-enabled maize breeding (Wang et al., 2020). Besides, meta-analysis is an objective method for statistically re-analyzing existing empirical literature, enabling a more unbiased evaluation of the evidence than that provided by traditional narrative commentary (Egger et al., 1997). It has been widely used to summarize and further explore complex biological mechanisms (Makinde et al., 2021), and it has also been applied in genetic studies of crop heterosis, grain yield, and stress tolerance (Li et al., 2011; Thiemann et al., 2014; Sharma et al., 2018; Wang et al., 2021). As an important plant architectural trait, LA affects the ability of the maize canopy to capture light and the light energy utilization efficiency of the population, understanding natural variation in LA and identifying its key genes are very important for breeding maize with high photosynthetic efficiency (Tao et al., 2002).

Here, 492 diverse maize inbred lines were used to investigate LA-related traits at different leaf positions in multiple environments. A GWAS was performed with 1.25 M SNPs to explore natural allelic variations that influence LA; in addition, a large amount of LA QTL data from previously published studies was used for meta-analysis. The main purposes were: 1) to explore the phenotypic variation of LA 2) to identify natural variation of SNPs/loci and candidate genes significantly associated with LA 3) to select germplasms with the favorable haplotypes of LA to improve maize plant architecture. The results will enrich our understanding of the genetic basis of LA and enhance ideotype-based maize breeding.

## Materials and methods

### Plant materials and growth conditions

The association mapping panel (AMP) used in this study consists of 492 maize inbred lines, including 225 tropical/subtropical germplasms and 267 temperate germplasms, which is a subset of 527 inbred lines (Yang et al., 2011). In 2020, all 492 maize inbred lines were planted in Yuanyang Modern Agricultural Science and Technology Park of Henan Agricultural University (Yuanyang, N35°N, E113°E) in the end of April (defined as YYC) and in early July (defined as YYX). In addition, 269 inbreds (73 tropical/subtropical and 196 temperate) were randomly selected from the AMP and planted at the XunXian Experimental Station of Hebi Academy of Agricultural Sciences in Henan province (Hebi, N35°N, E114°E, defined as HB), the Cotton Seed Farm in

Yongcheng, Henan (YongCheng, N33°, E116°, defined as YC) and the Yuanyang Modern Agricultural Science and Technology Park of Henan Agricultural University (Yuanyang, N35°, E113°, defined as YY) in early June 2020. All inbred lines were planted in a randomized complete block design with two replications, a single row length of 3 m, a row spacing of 0.67 m, and a final planting density of 45,000 plants/ha in all environments.

## Measurement of LA

LA data from the AMP were collected using a digital angle ruler (digital display 360° angle ruler 0–200 mm, Wenzhou Weidu Electronics Co., Ltd.). To be consistent, we investigated the flowering date of each line, and LA were investigated at 15 days after pollination for each line at all environments. Three Leaf Angle (LA-related traits) was scored at three positions: the upper leaf (first leaf below the flag leaf, ULA), the middle leaf (first leaf above the first ear, MLA), and the lower leaf (second leaf below the first ear, LLA). Only ULA was investigated at YYC and YYX, whereas all three LA-related traits were recorded at HB, YC, and YY. The average values of ULA, MLA, and LLA were calculated from five uniformly growing plants in each row. The average of each LA-related traits for the two replications in each environment was calculated. Overall, ULA at all the five environments (YYC, YYX, YY, HB, and YC), MLA, and LLA at the three environments (YY, HB, and YC) were collected. These phenotypic data was used for general statistical analysis, Pearson correlation analysis, two-way ANOVA, broad-sense heritability calculation, Best Linear Unbiased Prediction (BLUP) and GWAS.

## Statistical analyses of LA

General statistical analyses (e.g. mean, variation range, standard deviation, kurtosis, skewness) were finished in SPSS Statistics V17.0 after removal of outliers. For the five values measured for one trait of a genotype, if a value is not within the range of the mean plus or minus 1.5 times of standard deviation, it will be regarded as an outlier. A repeated-measures two-way ANOVA, using the following formula: $V = G + E + G \times E + e$, here, V is total variance, G is variance of genotype, E is environmental variance, G × E is variance of genotype-environment interaction, e is error. It was also performed in SPSS Statistics V17.0. The mixed linear model of the lme4 package in R (version 4.1.1, R Foundation for Statistical Computing, http://www.r-project.org/) was used to calculate the Best Linear Unbiased Prediction (BLUP) value for each trait in the five environments (Eugster et al., 2011). In addition, the Best Linear Unbiased Prediction (BLUP) was calculated for GWAS, which can estimate the random effects (Coram et al., 2017). For MLA and LLA, broad-sense heritability was computed using the following formula: $H^2 = [\delta_G^2 / \delta_G^2 + (\delta_{GE}^2/n) + \delta_e^2/(nr)]$, where $\delta_G^2$ is the genotypic variance, $\delta_{GE}^2$ is the genotype × environment variance, $\delta_e^2$ is the error variance, $r$ is the number of replications, and $n$ is the number of environments. Considering the unbalanced data for ULA, the harmonic mean ($h$), $h = (492 + 269)/(492/2 + 269/3)$, was used to instead of the number of environments ($n$) for calculating the $H^2$ (Nyquist and Baker, 1991; Schmidt et al., 2019). Pearson correlation coefficients between paired traits were calculated using the corr function in R (version 4.1.1).

## GWAS

The genotype data used in this study was downloaded from the Maizego website (http://www.maizego.org/Resources.html) and have been deposited in the European Variation Archive (EVA) at EMBL-EBI under accession number PRJEB56161 (https://www.ebi.ac.uk/eva/?eva-study=PRJEB56161). This genotype data was inferred from the Illumina MaizeSNP50 array, reduced-representation genome sequencing (genotyping-by-sequencing, GBS), the high density Affymetrix Axiom Maize 600K array (600K), and previous deep RNA-sequencing data of whole kernels at 15 days after pollination obtained by Liu et al. (Liu et al., 2016). The genotype data consisted of 1.25 M SNPs (B73_RefGen_v2) covering the whole maize genome with a minimum allele frequency (MAF) ≥ 0.05 (Liu et al., 2016). Here, the GLM approach controlling population structure (Q) was adopted after comparing the performances of three linear models, that is, GLM (GLM + Q, only control population structure), MLM (GLM + K, only control relative kinship) and MLM (GLM + Q + K, correcting for population structure and relative kinship) models (Li et al., 2013) implemented in TASSEL 3.0 (Bradbury et al., 2007) was therefore used to perform GWAS. For GLM model, $y = X\alpha + Z\beta + e$. or $y = SNP + Q + e$. For MLM model, $y = X\alpha + Z\beta + W\mu + e$. or $y = SNP + Q + Kinship + e$. where, y is the trait value, Xα (population structure or Q matrix) is fixed effect, Zβ (SNP or marker effect) is fixed effect, Wμ (Kinship matrix) is random effect and e is residual (Yu et al., 2006). In addition, to control the type I (false positive) and type II (false negative) error rates, Quantile-Quantile (QQ) plots of the three statistical models for each LA-related trait were compared, if a model that has a distribution closer to the diagonal line indicates a better control for type I and II errors (Zhang et al., 2010). Thus, the more appropriate model was selected to interpret the GWAS results of LA-related traits.

Taking into account the linkage disequilibrium (LD) among SNP markers, the effective marker number (En) for the genotypic dataset was 490 548, as previously calculated using GEC software (Deng et al., 2017). The suggested $p$-value of $2.04 \times 10^{-6}$ (1/En) was used as the genome-wide threshold for significant SNP–trait associations, as commonly used in plant genome-wide association studies.

## Meta-analyses

Based on a published review (Cao et al., 2022), the QTL information have been reported for LA-related traits (Lu et al.,

TABLE 1 Descriptive statistics for leaf angle-related traits of maize in the association mapping panel in different environments.

| Trait | Environment | Range (°) | Mean (°) | sd. | Ske. | Kur. | $H^2$ |
|---|---|---|---|---|---|---|---|
| ULA | HB | 11.65–67.56 | 34.65 | 11.57 | 0.61 | 0.08 | 0.59 |
| | YC | 14.32–62.98 | 37.65 | 10.47 | 0.17 | −0.56 | |
| | YY | 9.90–86.50 | 31.93 | 14.19 | 1.17 | 1.60 | |
| | YYC | 11.48–72.65 | 33.11 | 11.49 | 0.69 | 0.59 | |
| | YYX | 6.71–80.08 | 35.11 | 12.56 | 0.74 | 0.60 | |
| | BLUP | 14.71–61.08 | 35.08 | 8.37 | 0.58 | 0.21 | |
| MLA | HB | 16.95–63.60 | 38.08 | 9.31 | 0.43 | −0.03 | 0.71 |
| | YC | 14.05–63.34 | 38.26 | 9.97 | 0.18 | −0.45 | |
| | YY | 18.68–79.68 | 39.51 | 11.67 | 0.85 | 0.74 | |
| | BLUP | 23.31–60.31 | 38.77 | 6.72 | 0.58 | 0.36 | |
| LLA | HB | 23.42–77.64 | 43.43 | 10.39 | 0.72 | 0.68 | 0.69 |
| | YC | 19.62–68.87 | 42.00 | 9.21 | 0.14 | −0.13 | |
| | YY | 19.48–74.00 | 43.87 | 9.95 | 0.37 | −0.04 | |
| | BLUP | 30.81–62.95 | 43.19 | 5.69 | 0.46 | 0.49 | |

sd., STDEVP, standard deviation calculated based on the given sample population.
Ske., SKEW, the degree of asymmetry used to represent the relative mean.
Kur., KURT, a peak value used to represent the dataset.
$H^2$, Broad-sense heritability.

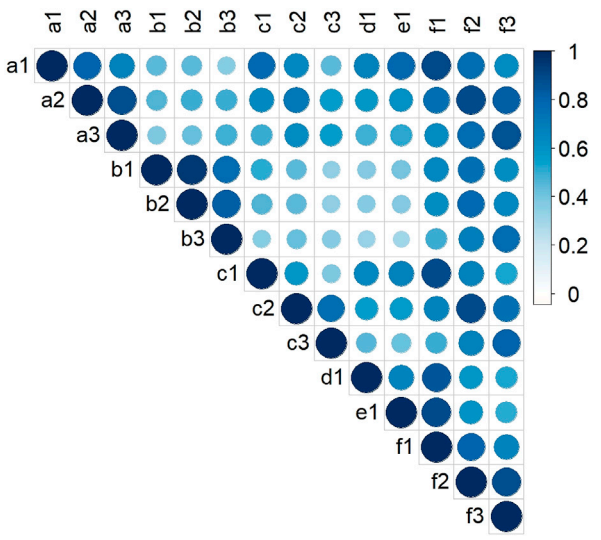



FIGURE 1
Pearson correlation coefficients for LA-related traits of maize in the association mapping panel in different environments. **(a–f)**: HB, YC, YY, YYC, YYX, and BLUP, respectively; (1–3): ULA, MLA, and LLA, respectively.

2007; Ku et al., 2010; Tian et al., 2011; Ku et al., 2012; Chen et al., 2015; Ding et al., 2015; Hu et al., 2015; Li et al., 2015; Yang et al., 2015; Ku et al., 2016; Pan et al., 2017; Shi et al., 2017; Wang et al., 2017; Dzievit et al., 2019; Liu et al., 2019; Zhang et al., 2019; Zhang et al., 2020; Tang et al., 2021). The QTL information (Chromosome, LOD, $R^2$, Confidence Interval and so on) were summarized from eighteen studies published in the lasted 15 years. For QTLs whose confidence interval (CI) was unknown, the following formulas were used to calculate CI:

$$\mathrm{CI} = 530/(N \times R^2) \quad (1)$$

$$\mathrm{CI} = 163/(N \times R^2) \quad (2)$$

where CI is the confidence interval, $N$ is the number of materials in the mapping population, and $R^2$ is the phenotypic variation. Eq. 1 was used for Backcross and $F_2$ mapping populations, and Eq. 2 was used for recombinant inbred line (RIL) mapping populations (Darvasi and Soller, 1997). If the LOD value was unknown, it was calculated using $R^2 = 1–10^{(-2\mathrm{LOD}/N)}$ (Nagelkerke, 1991; Liu, 1997). QTLs without $R^2$ information were discarded.

IBM2 2008 Neighbors (MaizeGDB, https://www.maizegdb.org/) (Sharopova et al., 2002) was used as a reference genetic map, and 19 051 high-density markers (SSR, RFLP, RAPD, and SNP) covering the whole maize genome were obtained. The markers were combined with the collected QTLs and genetic map information, and a meta-analysis was performed to analyze all QTLs and markers in the genome and obtain the most suitable number of QTLs (Veyrieras et al., 2007). Five different models (1-, 2-, 3-, 4-, or N-QTL) with different Akaike information criterion (AIC) values were proposed and used in BioMercator V4.2.3, a genetic map compilation and meta-analysis software to integrate QTL data with genome structural and functional annotation. The model with the lowest AIC-value was considered optimal (Arcade et al., 2004). Finally, the QTLs presented by the optimum model were regarded as the meta-QTLs, we named

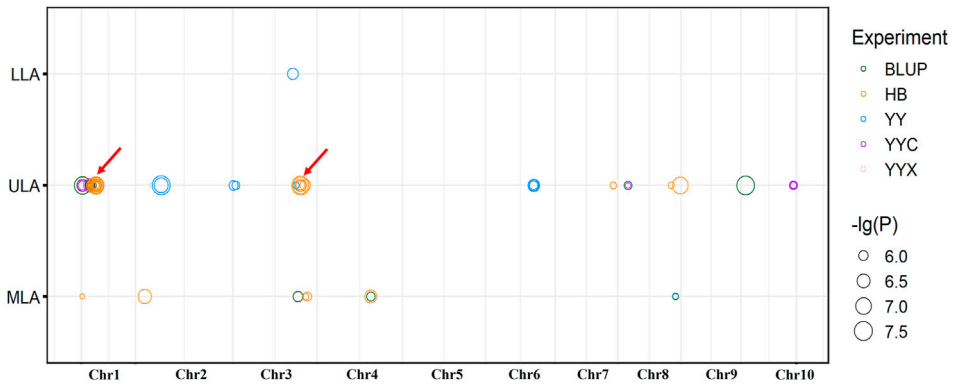



FIGURE 2
Distribution of significant loci detected by GWAS on maize chromosomes. Red arrows indicate co-located QTLs, for the same trait in different environments.

these meta-QTLs using the format M-q-trait-chromosome number-sequence number [for example, *MqLA1-1*, *MqLA* (Leaf Angle) *1* (chromosome number) *-1* (sequence number)].

## Analyses of candidate genes

The previously estimated decay distance of LD in this AMP (~30 kb, $R^2 = 0.1$) (Liu et al., 2016) was used to define a 60-kb QTL interval, i.e., the 30 kb upstream and downstream of each SNP. In each environment and for each trait, all QTLs with overlapping QTL intervals were categorized as non-redundant QTLs. If a non-redundant QTL was detected by different LA-related traits, in different environments, from previously published LA QTLs, or physical distance of less than 10 Mb by GWAS and meta-QTL analysis, it will be defined as a co-located QTL. All potential candidate genes within all non-redundant QTLs or co-located QTLs were identified based on the filtered working gene list from the reference genome of the maize inbred line B73 (RefGen_v2). These genes were downloaded from MaizeGDB and annotated using InterProScan (http://www.ebi.ac.uk/interpro/scan.html). Phenotypic variation explained (PVE) by each QTL was estimated based on the $R^2$ value of the most significant SNPs within the QTL. Candidate genes within the CIs of meta-QTLs were also obtained. In each non-redundant QTL or co-located QTL, the most likely candidate gene was selected based on its annotation or because it contained the peak SNP (the most significant SNP). If there was no gene in the interval, the neighboring gene of the peak SNP was considered to be the most likely candidate gene.

## Gene Ontology enrichment analyses

Gene Ontology (GO) enrichment analysis was performed using OmicShare tools (https://www.omicshare.com/tools) (Ding et al., 2019). Specifically, the candidate genes were mapped to the various sets of the GO database (http://www.geneontology.org/), the number of genes in each set was counted, and the list of genes with a specific GO function and the number of genes were obtained. The top 20 GO terms with minimum *p* values were selected for analysis and plotting (Lv et al., 2019).

## Linkage disequilibrium analyses

Linkage disequilibrium (LD) was estimated by the squared correlation of the paired SNPs, which was calculated with TASSEL 3.0 software. An LD plot was generated using the 'genetics' and 'LDheatmap' package in R (version 4.1.1).

## Haplotype analyses

All SNPs (MAF ≥0.05) in target genes (*ZmULA1*, *ZmCLA4*, *lg1*, *lg2* and *ZmTAC1*) were used for haplotype analysis by using 1.25M SNPs genotype data (Liu et al., 2016). The BLUP values of LA-related traits in 492 inbred lines were used as phenotypic data. Haplotypes contained in more than 10 inbred lines were used for comparative analysis.

# Results

## Phenotypic evaluation

LA at three positions, ULA, MLA and LLA (Supplementary Figure S1), were investigated in five environments. The LA-related traits showed the greatest variation in YY and lowest variation in YC. The maximum angle was 8.7-fold higher than the minimum angle (9.90°–86.50°) for ULA, 4.5-fold higher for

TABLE 2 List of meta-QTLs and accompanying details for leaf angle-related traits in maize.

| Meta QTL | Chr | Position (cM) | CI (cM) | CI(Mb) | Left marker | Right marker |
|---|---|---|---|---|---|---|
| *MqLA1-1* | 1 | 181.08 | 180.53–181.63 | 24.58–25.28 | sfp3 | cdo860a |
| *MqLA1-2* | 1 | 209.46 | 209.18–209.75 | 32.87–33.12 | bnlg176 | xyl5 |
| *MqLA2-1* | 2 | 169.93 | 169.92–169.94 | 15.57–16.04 | msl2 | hon101 |
| *MqLA3-1* | 3 | 50.37 | 50.27–50.47 | 3.56–3.82 | cdo511 | isu157 |
| *MqLA3-2* | 3 | 63.28 | 62.42–64.15 | 4.19–4.36 | IDP8355 | cl19880_1 |
| *MqLA3-3* | 3 | 92.58 | 90.70–94.46 | 6.94–7.41 | T3-9 (8447) (3) | eif3 |
| *MqLA3-4* | 3 | 107.90 | 107.66–108.14 | 8.59–8.89 | csu728c | gts1 (CBM 3.03) |
| *MqLA4-1* | 4 | 12.20 | 4.80–19.60 | 1.09–1.99 | bnlg1434 | IDP4473 |
| *MqLA4-2* | 4 | 61.87 | 59.77–63.97 | 3.75–4.09 | uaz52a | pd1 |
| *MqLA4-3* | 4 | 88.11 | 85.62–90.60 | 5.25–5.38 | TIDP2802 | uaz61a |
| *MqLA4-4* | 4 | 106.32 | 105.69–106.95 | 5.54–9.80 | uaz184 (hfi) | umc1288 |
| *MqLA4-5* | 4 | 108.76 | 108.07–109.45 | 6.44–7.57 | cle7 | mads25 |
| *MqLA4-6* | 4 | 122.55 | 120.36–124.75 | 10.02–10.90 | nbcs11 | IDP4286 |
| *MqLA4-7* | 4 | 133.29 | 131.80–134.79 | 10.46–11.92 | gpm574a | bnlg1126 |
| *MqLA4-8* | 4 | 183.50 | 181.95–185.05 | 17.50–17.86 | umc1902 | wrky36 |
| *MqLA5-1* | 5 | 10.10 | 4.20–16.00 | 0.54–0.88 | telomere5S | AY109758 |
| *MqLA5-2* | 5 | 46.73 | 46.29–47.17 | 2.14–2.59 | IDP7849 | umc1901 |
| *MqLA5-3* | 5 | 53.24 | 52.52–53.96 | 2.32–2.59 | IDP2557 | prh24 |
| *MqLA5-4* | 5 | 61.67 | 61.12–62.23 | 2.70–2.86 | umc2591 | phm5359 |
| *MqLA5-5* | 5 | 76.90 | 75.53–78.28 | 3.73–4.22 | bhlh159 | IDP1463 |
| *MqLA5-6* | 5 | 91.07 | 89.73–92.41 | 4.88–5.49 | IDP6013 | TIDP5654 |
| *MqLA5-7* | 5 | 108.57 | 107.26–109.59 | 6.14–6.67 | uaz163 | gpm921a |
| *MqLA5-8* | 5 | 132.51 | 132.26–132.76 | 7.13–7.36 | ucsd64a | npi305a |
| *MqLA7-1* | 7 | 91.36 | 90.65–92.08 | 6.60–7.45 | ao5 | bnlg2160a |
| *MqLA7-2* | 7 | 116.55 | 115.62–117.49 | 10.52–10.78 | gpm913a | magi108570 |
| *MqLA7-3* | 7 | 130.22 | 129.59–130.86 | 13.96–14.08 | csu794 | y8 |
| *MqLA8-1* | 8 | 34.72 | 30.97–38.47 | 3.66–4.24 | TIDP3564 | arf4 |
| *MqLA8-2* | 8 | 61.95 | 60.99–62.91 | 7.04–7.30 | nactf118 | nactf130 |
| *MqLA8-3* | 8 | 68.71 | 67.00–70.42 | 5.68–6.70 | gpm600 | TIDP5156 |
| *MqLA8-4* | 8 | 74.72 | 72.37–77.07 | 6.03–6.86 | TIDP5156 | IDP1629 |
| *MqLA8-5* | 8 | 79.01 | 78.34–79.69 | 6.86–7.98 | ncr (sod3b) | IDP1629 |
| *MqLA8-6* | 8 | 92.51 | 91.91–93.11 | 7.52–8.09 | TIDP3314 | isu1410a |
| *MqLA8-7* | 8 | 98.12 | 95.78–100.47 | 8.11–8.89 | gpm932d | IDP7980 |
| *MqLA8-8* | 8 | 156.01 | 153.42–158.61 | 17.24–18.37 | cle26 | cdo328 |
| *MqLA8-9* | 8 | 164.53 | 164.51–164.56 | 18.20–19.85 | wrky80 | wrky26 |
| *MqLA9-1* | 9 | 23.06 | 20.41–25.72 | 5.01–5.52 | rz144c | mads60 |
| *MqLA9-2* | 9 | 53.11 | 52.40–53.83 | 8.23–8.94 | bzip100 | TIDP4624 |
| *MqLA9-3* | 9 | 61.86 | 59.63–64.09 | 8.36–9.74 | mHbrMG162-Mo17 | c1 |
| *MqLA9-4* | 9 | 68.73 | 67.03–70.43 | 9.75–10.85 | umc113a | crs4a |
| *MqLA9-5* | 9 | 78.53 | 73.55–83.51 | 10.85–11.57 | isu1146 | ptf1 |
| *MqLA9-6* | 9 | 93.27 | 91.91–94.63 | 11.93–12.35 | mkk1 | TIDP5270 |
| *MqLA9-7* | 9 | 103.18 | 102.71–103.65 | 12.98–13.45 | umc1131 | chr113 |
| *MqLA10-1* | 10 | 41.83 | 40.53–43.13 | 2.78–3.06 | pza02221 | agrc561 |
| *MqLA10-2* | 10 | 58.78 | 57.86–59.71 | 4.07–4.17 | cl24029_1 | rpp9 |
| *MqLA10-3* | 10 | 101.90 | 95.60–108.20 | 5.38–5.92 | glk5 | TIDP3571 |
| *MqLA10-4* | 10 | 133.65 | 132.93–134.38 | 7.40–9.23 | umc2749 | oy1 |
| *MqLA10-5* | 10 | 187.8 | 187.45–188.15 | 58.30–59.51 | gpt1 | mHbrMC413-Mo17 |

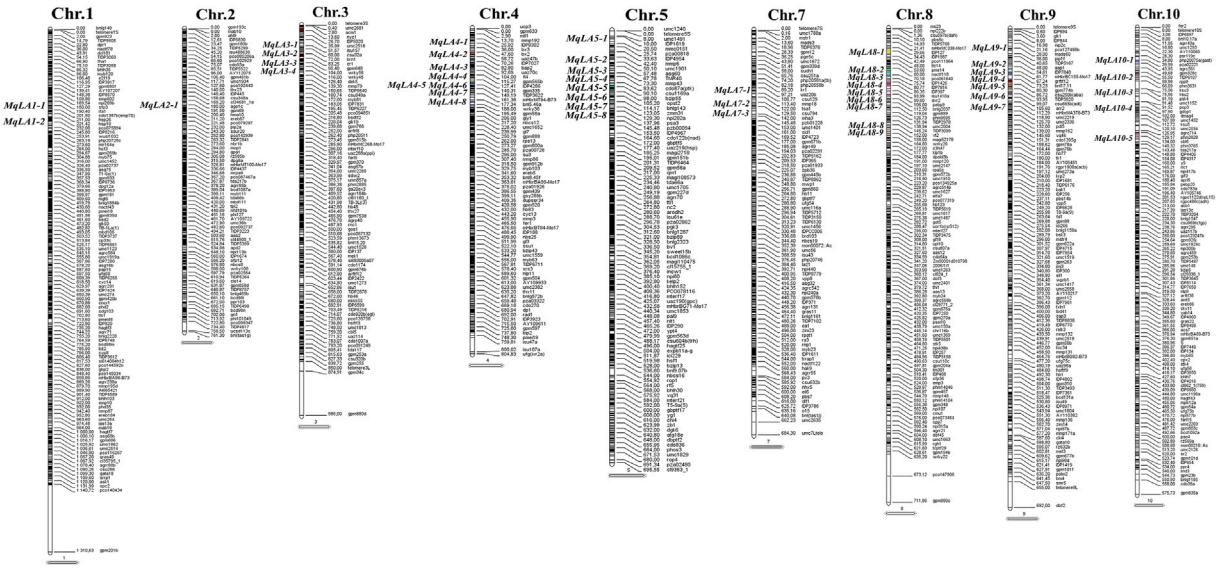



FIGURE 3
Meta-QTLs distributed on maize chromosomes with IBM2 2008 Neighbors as a reference.

TABLE 3 List of co-located QTLs and physical position revealed by GWAS and meta-analysis.

| Chr | GWAS | | Meta-analysis | |
|---|---|---|---|---|
| | QTL name | CI (Mb) | QTL name | CI (Mb) |
| 1 | *qLA3* | 23.53–23.59 | *MqLA1-1* | 24.58–25.28 |
| 1 | *qLA6* | 35.71–35.77 | *MqLA1-2* | 32.87–33.13 |
| 1 | *qLA7* | 40.11–40.17 | *MqLA1-2* | 32.87–33.13 |
| 1 | *qLA8* | 40.15–40.21 | *MqLA1-2* | 32.87–33.13 |
| 3 | *qLA12* | 1.95–2.01 | *MqLA3-1* | 3.56–3.82 |
| 3 | *qLA13* | 8.06–8.12 | *MqLA3-4* | 8.59–8.89 |
| 8 | *qLA27* | 28.27–28.33 | *MqLA8-9* | 19.83–19.85 |

MLA (14.05°–63.34°), and 3.8-fold higher for LLA (19.48°–74.00°) (Table 1). Thus, the panel exhibited rich genetic diversity in LA, which could be ranked LLA > MLA > ULA (Table 1 and Supplementary Figure S2). All phenotypes were typical quantitative traits, in that they were continuous variables that exhibited a normal distribution (Supplementary Figure S3).

Two-way ANOVA indicated that there were significant genetic (G) and environmental (E) effects on LA-related traits, but the effect of their interaction (G × E) was not significant (Supplementary Table S1). The results also showed that genetics had a greater effect on LA than environment. There were positive correlations between LA-related traits in all environments (Figure 1). The broad-sense heritabilities of ULA, MLA, and LLA were 0.59, 0.71, and 0.69, respectively (Table 1), again indicating that they were mainly affected by genetic factors.

## GWAS

LA was less sensitive to the K model than other two models, and this model could better control type I and type II errors (Supplementary Figure S4). In addition, we examined the distribution of the three LA traits in each sub-population across different locations, the results shown that population structure has a small effect on the leaf angle (Supplementary Figure S5). Therefore, the GWAS under the K model was analyzed further (Supplementary Figure S6).

In total, 85 significant SNPs were detected by ULA, MLA and LLA in all environments, and involving 32 non-redundant QTLs. The QTLs were distributed on all chromosomes except chromosome 5 and 9, and there was a QTL hot spot at chromosome 1 and 3 (Figure 2). Twenty-one major-effect QTLs explained more than 10% of the phenotypic variation ($R^2$ = 10.01–24.09%). Twenty-six QTLs were identified for ULA, with a mean $R^2$ of 9.34% (4.80%–15.16%), and seven QTLs were detected for MLA, with a mean $R^2$ of 15.14% (9.48%–24.09%). Only one QTL was associated with LLA, and it explained 11.63% of the phenotypic variation.

Fourteen non-redundant QTLs were detected in HB and explained 9.08%–21.93% of the phenotypic variation. Ten non-redundant QTLs were detected in YY (10.05%–15.16%), nine in BLUP (4.80%–24.09%), and only four and two in YYC and YYX, respectively. No significant QTLs were detected in YC (Figure 2).

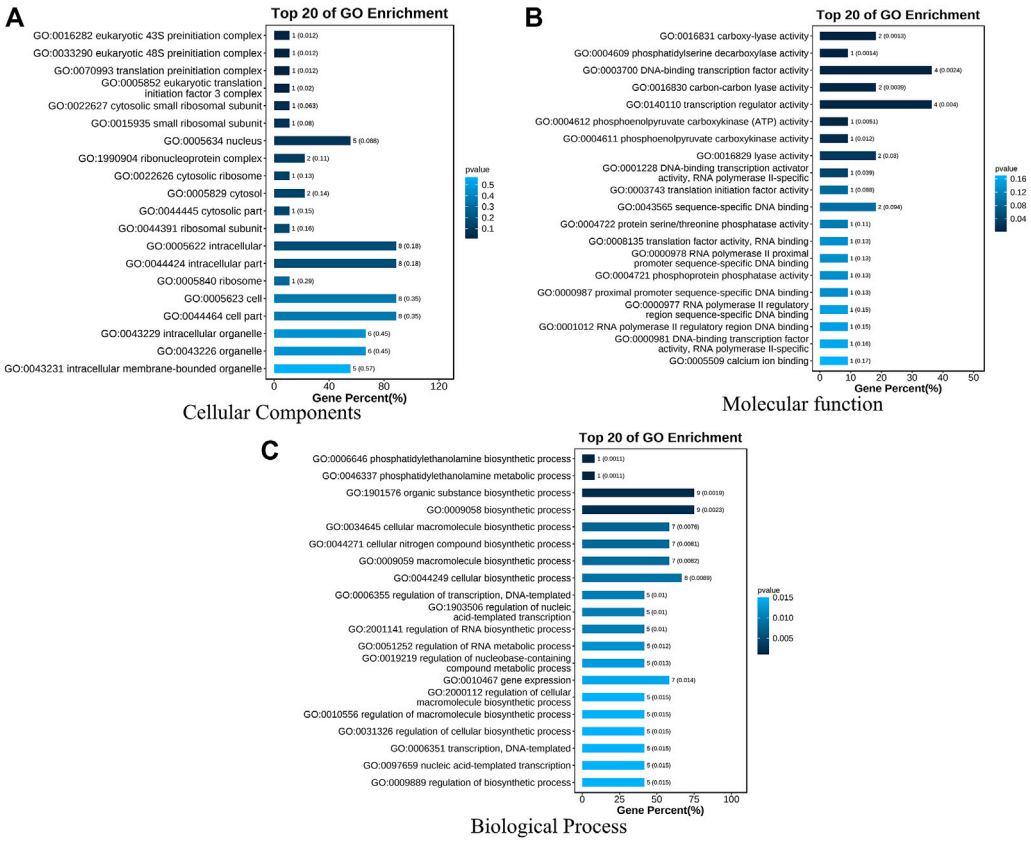



FIGURE 4
GO enrichment analysis of seventeen genes within seven co-localized QTLs. Three GO Ontologies are shown **(A)** Cellular Component; **(B)** Molecular Function, and **(C)** Biological Process.

Five QTLs were identified in multiple environments (Figure 2). For ULA, *qLA2* was detected in YYC, YYX, and BLUP, and *qLA3* and *qLA27* were detected in YYC and BLUP. For MLA, *qLA23* was identified in HB and BLUP, and *qLA29* was detected in BLUP and YY. These loci can be stably inherited in different environments and play an important role in the regulation of plant architecture. QTLs detected for different traits may have pleiotropic effects. *qLA1* was identified by MLA and ULA and may therefore regulate the size of both ULA and MLA. In addition, *qLA21* was also detected in different traits (ULA and MLA), suggesting that *qLA21* may have a similar role in the regulation of LA at different positions. Detailed information on the GWAS results, including the physical position, *p*-value, and $R^2$ for each QTL is provided in Supplementary Table S3.

## Candidate gene analysis

Based on the B73 RefGen_v2 reference genome (https://www.maizegdb.org/), 82 genes were found within 32 non-redundant QTLs. For example, *GRMZM2G049159* within *qLA7* was identified for ULA in HB and encodes a GRAS family transcription factor; its homolog *OsGRAS23* induces downstream stress-responsive genes and positively regulates drought tolerance in rice (Xu et al., 2015), and it may be involved in the regulation of rosette leaf development in *Arabidopsis* (Schmid et al., 2005). *qLA18* was also identified for ULA in HB, and four genes are located in this QTL. Among them, *GRMZM2G071705* encodes an F-box protein; this type of protein may be involved in the strigolactone (SL) biosynthetic pathway and further regulate plant architectural traits such as tiller number and LA (Ishikawa et al., 2005; Dong et al., 2016).

Twenty genes were identified in co-located QTLs and may be involved in protein ubiquitination (*GRMZM2G134176*), carbon metabolism (*GRMZM2G134256*), and the citric acid cycle (*GRMZM2G158378*) and play important roles in the regulation of plant growth and stress tolerance. *GRMZM2G311328* is located in *qLA21*, which was identified for MLA and ULA in HB. This gene may be pleiotropic; its homolog encodes a vesicle auxin mediated transporter, and

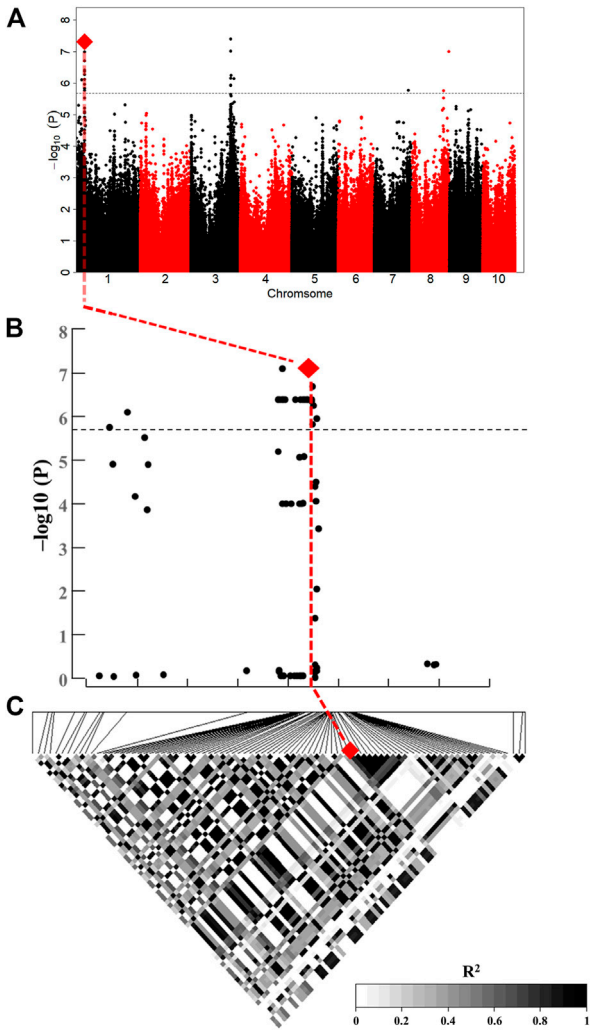



FIGURE 5
*ZmULA1* affected LA-related traits. **(A)** Manhattan plot of ULA at HB. The line represents the threshold $-\log_{10}(P) \geq 5.69$ ($p \leq 2.04 \times 10^{-6}$). **(B)** Enlarged Manhattan plot of the lead SNP and 85 SNPs associated with *ZmULA1*; red diamonds represent the lead SNP. **(C)** $R^2$ values of significant SNPs associated with *ZmULA1*; the lead SNP was located in 2.64 kb downstream of *ZmULA1*.

regulation of auxin polar transport is important for development and architecture of *Arabidopsis* (Kitakura et al., 2011). *GRMZM2G134073* in *qLA29* was simultaneously identified for MLA in BLUP and YY; it encodes a NAC family transcription factor that can promote resistance to heat damage (Li H et al., 2020). Its homolog *OsNAC10* plays a key role in rice drought and disease resistance (Jeong et al., 2010), and it may also participate in regulating plant architecture and abiotic stress tolerance.

Among the 82 genes, only *GRMZM2G071790* has been reported, it encodes a beta-6 tubulin and plays an important role in maize tolerance to *Ustilago maydis* (Ruan et al., 2021), as well as catalyzing auxin transport in *Arabidopsis* (Terasaka et al., 2005). Although most of these genes screened by GWAS have unknown functions, their homologous genes can provide valuable information. Interestingly, some genes may have pleiotropic effects, that may play important roles at both LA and stress resistance. Thus, our results provide new and valuable information for understanding the genetic mechanisms that underlie LA.

## Meta analysis

A total of 294 QTLs related to LA were obtained from studies published in the last 15 years (Supplementary Table S4) (Cao et al., 2022). 47 meta-QTLs were obtained by the meta-QTL analysis, containing 816 genes. These 47 QTLs were distributed on all chromosomes except chromosome 6 (Table 2 and Figure 3). The largest number of meta-QTLs (9) were located on chromosome 8, and only one was located on chromosome 2. Seven co-located QTLs were identified by both GWAS and meta-analysis (Table 3). Based on the GWAS results, 17 genes within the seven co-located QTLs were examined, but none had been characterized previously. Gene Ontology enrichment was then performed with the 17 genes, and they were mainly involved in the synthesis of intracellular parts (cellular component, GO: 0044424), transcriptional regulatory activity (molecular function, GO:0140110) and organic substance biosynthetic process (biological process, GO:1901576) (Figure 4). Interestingly, LA-related genes identified to date encode transcription factors or participate in hormonal signal transduction (Cao et al., 2022), and plant hormones are natural organic compounds (Adam 1999). These results suggest that the 17 genes in the co-localized QTLs may regulate LA by encoding transcription factors and/or participating in the synthesis of organic compounds, including plant hormones.

*qLA7*, which was related to ULA, was located on chromosome 1 and had an extremely strong association signal (Figure 5A and Figure 5B). There was only one gene (*GRMZM2G049159*) within this QTL, and it encodes a GRAS family transcription factor, involved in the synthesis of organic (GO:1901576) (Figure 4), plant hormones such as CK and BR, as natural organic compounds, play an important role in regulating LA. Therefore, *ZmULA1* may be involved in the metabolic process of plant hormones to regulate maize LA. Additionally, its homolog may be involved in the development of rosette leaves in *Arabidopsis* (Schmid et al., 2005). We speculate that it may affect the formation of LA in maize, and we named it *ZmULA1*.

## Haplotype analysis of *ZmULA1*

To analyze the haplotype of *ZmULA1*, we extracted all polymorphic sites within one LD decay distance near the lead

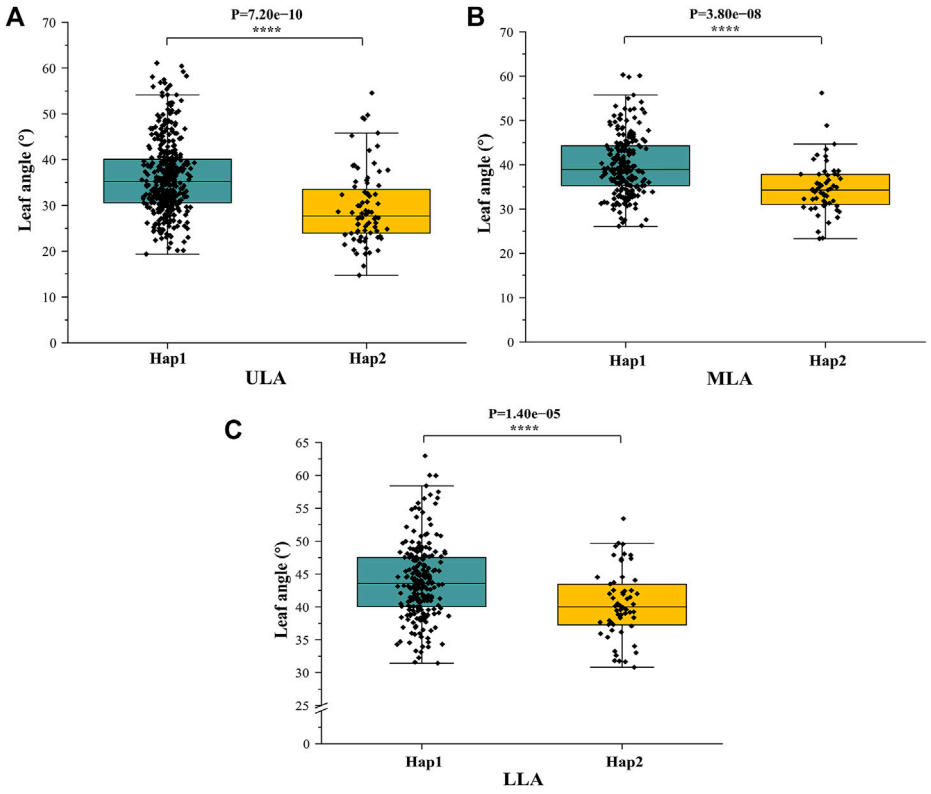



FIGURE 6
Differences in LA-related traits between Hap1 and Hap2. **(A)** ULA. **(B)** MLA. **(C)** (LLA). *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, ****: $p < 0.0001$.

SNP (chr1.S_40141659, $p$ = 7.57e−08) to perform LD analysis, there was a strong linkage relationship (the average pairwise $r^2$ value was 0.45) between the lead SNP and polymorphic sites (Figure 5C). Subsequent analysis of 17 SNPs in *ZmULA1* using BLUP values of 492 lines identified two haplotypes. All maize inbred lines belonged to hap1 (412) or hap2 (80), and the mean value of hap2 was smaller than that of hap1 for the three LA-related traits (Supplementary Table S5). Specifically, ULA ($p$ = 7.20e−10), MLA ($p$ = 3.80e−08), and LLA ($p$ = 1.40e−05) showed extremely significant differences between hap1 and hap2 (Figure 6). Temperate materials had smaller values of LA-related traits (ULA, MLA and LLA) than tropical lines, and the temperate materials with hap2 had the smallest LAs (Figure 7). Furthermore, 65% (52/80) of the maize inbred lines derived from China belonged to hap2 (Table 4), and these elite inbred lines can be used to improve the plant architecture of maize cultivars. In summary, these results suggest that natural variation in *ZmULA1* may affect the three LA-related traits, which may be influenced by constant selection during maize breeding. In addition, we conducted haplotype analysis of four known genes (*ZmCLA4*, *lg1*, *lg2* and *ZmTAC1*) that regulate LA in previous studies (Harper and Freeling, 1996; Moreno et al., 1997; Yu et al., 2007; Zhang et al., 2014), favorable haplotype for each gene which have smallest LA was identified, that are hap4 for *ZmCLA4*, hap4 for *lg1*, hap2 for *lg2* and hap2 for *ZmTAC1* (Figure 8). The germplasms with favorable haplotype combinations of 492 inbred lines were identified, and a trend was found that the more favorable haplotypes, the smaller the leaf angle (Supplementary Table S6).

## Discussion

Increasing planting density is an important way to improve maize yield, and LA is one of the key traits that determine whether maize can tolerate high planting density. Several genes that regulate maize LA have been cloned in previous studies. For example, *lg1* (liguleless1) encodes a protein containing an SBP domain, and its mutation causes the loss of tongues and ears, making plants more compact (Moreno et al., 1997). *lg2* (liguleless2) encodes a basic leucine zipper (bZIP) transcription factor. These two genes are located in the same developmental pathway, in which *lg2* plays an earlier role than *lg1* (Harper and Freeling, 1996; Walsh et al., 1998). *ZmTAC1* is also considered to regulate LA in maize. It encodes a protein composed of 263 amino acids that is most highly expressed in

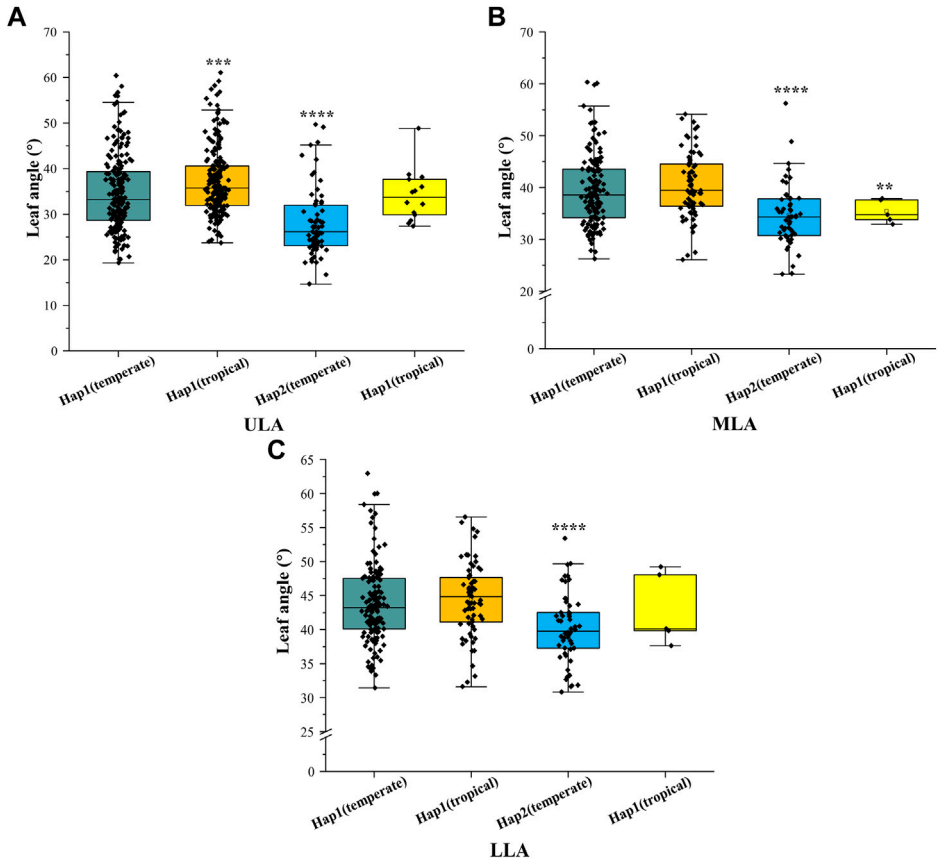



FIGURE 7
Differences in LA-related traits between Hap1 (temperate), Hap1 (tropical), Hap2 (temperate), and Hap2 (tropical). **(A)** ULA. **(B)** MLA. **(C)** (LLA). The significance level was obtained by comparing others haplotypes with Hap1(temperate), respectively. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, ****: $p < 0.0001$, not marked stands for not significant.

the maize leaf sheath, and the difference in LA traits between the compact leaf inbred line Yu82 and the extended leaf inbred line Shen137 is caused by nucleotide variation in its non-coding region (Ku et al., 2011). Recently, Li et al. cloned *ZmACS7* from the *Sdw3* mutant; its overexpression caused lower plant height and higher LA, and multi-omics analysis showed that it alters plant architecture by promoting growth of the auricle and inhibiting elongation of the internode cells (Li Z et al., 2020). The genetic and molecular mechanisms of LA have also been widely dissected in other crops. For example, *OsTAC1* in rice is a homolog of maize *ZmTAC1* and has an intron sequence AGGA in the 3′ non-coding region. A single SNP variation (A/G) that changes 'AGGA' to 'GGGA' can lead to reduced expression of *ZmTAC1* and an LA close to zero, making the rice plant more compact (Yu et al., 2007).

In this study, 85 significant SNPs in 32 QTLs were identified by GWAS and included 82 genes. Seven QTLs could be co-located in different environments or different LA-related traits (Figure 2), they are considered as genomic hot regions and worth to mining for valuable information. *qLA1* and *qLA21* were co-located by ULA and MLA and they explained 16.96% and 9.48% of phenotypic variation, respectively. Only one gene (*GRMZM2G158378*) was found within *qLA1*, and its rice homolog has been shown to promote the transport and absorption of silicon (Mitani et al., 2009). Five genes were located in *qLA21*, one of which encoded a glycosyltransferase (*GRMZM2G014770*) and another a vesicle-mediated substance transporter (*GRMZM2G311328*). Five co-located QTLs (*qLA2*, *qLA3*, *qLA23*, *qLA27*, and *qLA29*), explaining 6.13%–18.89% of phenotypic variation, were identified in at least two environments, indicating that they are genetically stable and less affected by the environment. Most QTLs (~84%) were detected in only one environment (27/32); a possible explanation is that LA is a complex quantitative trait regulated by a large number of minor genes that are greatly affected by the environment, leading QTL effects to vary in different environments. It is worth mentioning that 21 of the 32 QTLs (~66%) were major

TABLE 4 The number of haplotypes for *ZmULA1* in 492 maize inbred lines classified by origin and source, respectively.

| Haplotype | Total | Origin | | | Source | | |
|---|---|---|---|---|---|---|---|
| | | Temperate | Tropical/sub-tropical | China | United States | CIMMYT | Other |
| Hap1 | 412 | 202 | 210 | 166 | 46 | 195 | 5 |
| Hap2 | 80 | 65 | 15 | 52 | 14 | 14 | 0 |

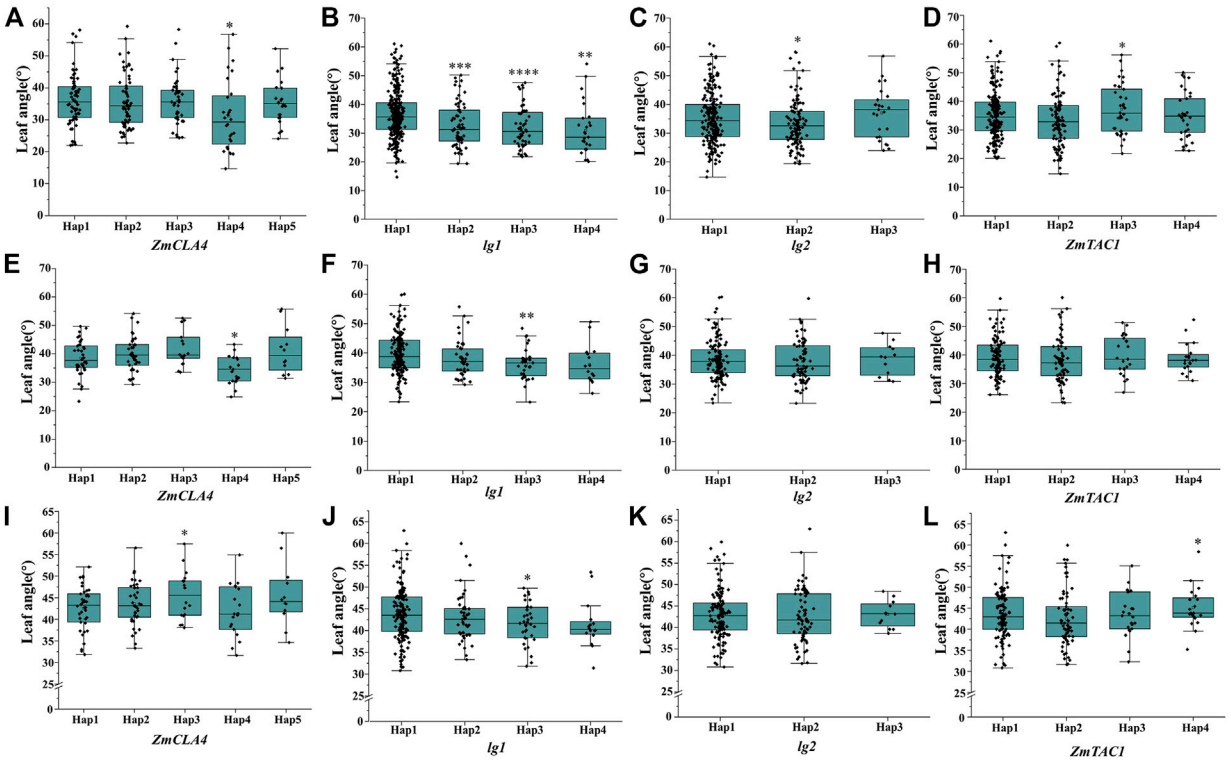



FIGURE 8
Differences of haplotypes in LA-related traits within four cloned genes that regulate maize LA. **(A–D)**, ULA, of *ZmCLA4*, *lg1*, *lg2*, *ZmTAC1*, respectively; **(E–H)**, MLA, of *ZmCLA4*, *lg1*, *lg2*, *ZmTAC1*, respectively; **(I–L)**, LLA, of *ZmCLA4*, *lg1*, *lg2*, *ZmTAC1*, respectively. The significance level was obtained by comparing others haplotype with hap1, respectively. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, ****: $p < 0.0001$, not marked stands for not significant.

QTLs with $R^2$ values greater than 10%. These results provide new information for understanding the genetic basis of the natural variation in LA.

In the current study, we collected QTLs for LA published over the past 15 years. In total, 294 QTLs were obtained and mapped to the IBM2 2008 Neighbors genetic map; 47 meta-QTLs were then identified by meta-analysis (Figure 3 and Table 2). Meta-analysis can effectively narrow the confidence interval of QTL and improve the prediction accuracy of candidate genes (Goffinet and Gerber, 2000; Arcade et al., 2004). Moreover, 816 genes were examined in 47 meta-QTLs; fifteen had been reported previously, and five had been functionally characterized. Two genes (*GRMZM2G005066*, *c1* and *GRMZM2G089713*, *sh1*) were involved in kernel colored aleurone and endosperm development (Tohge et al., 2017; Zhou et al., 2021). *GRMZM2G141399* (*du1*) encoded starch synthase III, which makes kernels glassy and inhibits glycogen accumulation (Cao et al., 1999). Two additional genes (*GRMZM2G051637*, *cr4* and *GRMZM2G419806*, *oy1*) were associated with plant height and senescence of maize seedlings (Peiffer et al., 2014; Khangura et al., 2020). Notably, *AC195340.3_FG001* (*tua1*), located in *MqLA5-6*, is associated with plant architecture at different planting densities and encodes an alpha tubulin

family protein; it has been annotated but not functionally characterized (Incognito et al., 2020). These results suggested that meta-analysis can provide more valuable information.

Seven QTLs were jointly identified by GWAS and meta-analysis (Table 3), and two, *qLA3* and *qLA7*, had significant association signals at chromosome one by GWAS (Figure 5A). Three genes (*AC205725.3_FG010*, *GRMZM2G471253*, and *GRMZM2G171073*) were located in *qLA3* and were expressed in leaves, potentially affecting leaf development. *GRMZM2G171073* encodes a C2H2-like zinc finger protein, and its homolog IDD1 is involved in gibberellin (GA) signal transduction and transport in *Arabidopsis* (Fukazawa et al., 2014). GA affects plant structure and interacts with BR, causing variation in LA (Tang et al., 2018). *ZmULA1* was located in *qLA7*, and hap1 and hap2 for this gene showed highly significant differences in LA-related traits (Figure 6). Hap2 was considered as a favorable haplotype because the hap2 had smaller LA (ULA, MLA and LLA) than other haplotypes, and temperate germplasms with hap2 had the smallest LA (Figure 7). About 80% (52/65) of the elite temperate germplasms from China (Table 4). Four genes (*ZmCLA4*, *lg1*, *lg2* and *ZmTAC1*) have been shown to regulate LA, and a lot of maize inbred lines with favorable haplotypes were selected by haplotype analysis (Figure 8). The results indicated that there was a trend that the more favorable alleles, the smaller the leaf angle. The germplasms carrying favorable haplotypes can be used to improve maize plant architecture to increase planting density and increase maize yield.

Although GWAS has been recognized as a powerful method for understanding the genetic basis of complex quantitative traits, meta-analysis is complementary to GWAS and can increase precision and accuracy of detected QTLs. Therefore, combining GWAS and meta-analysis, we screened a lot of potential targets/loci regulating LA in maize, and the candidate gene, *ZmULA1*, was predicted to play important roles in the regulation of LA. These results will provide reference for improving maize plant architecture.

## 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.ebi.ac.uk/eva/, PRJEB56161.

## Author contributions

XZ and JT designed the research. XZ supervised the research. HD, YS, XX, LS, JG, NL, JZ, JC, and ZF performed the experiment. HD, JL, and WL analyzed the data. HD and XZ wrote the paper. All authors read and approved the final manuscript.

## Funding

This research was supported by the National Natural Science Foundation of China (32171980), a project funded by the China Postdoctoral Science Foundation (2020M682295), a first-class postdoctoral research grant in Henan Province (202001032), the Research Start-up Fund for Youth Talents of Henan Agricultural University (30500563), the Open Project Funding of the State Key Laboratory of Crop Stress Adaptation and Improvement (2021KF07) and the HAU grant for collaborative crop science research (CCSR2022-1).

## Acknowledgments

We thank the research groups of Prof. Jianbing Yan at Huazhong Agricultural University and Prof. Xiaohong Yang at China Agricultural University for providing the maize association mapping panel and genotypic data.

## 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/fgene.2022.1004211/full#supplementary-material.

# References


Adam, C. (1999). Fatty acid composition of Mesorhizobium huakuii lipopolysaccharides. Identification of 27-oxooctacosanoic acid. *FEMS Microbiology Letters*. doi:10.1016/S0378-1097(99)00321-3

Arcade, A., Labourdette, A., Falque, M., Mangin, B., Chardon, F., Charcosset, A., et al. (2004). BioMercator: integrating genetic maps and QTL towards discovery of candidate genes. *Bioinformatics* 20, 2324–2326. doi:10.1093/bioinformatics/bth230

Arkebauer, T. J., Walter-Shea, E. A., Mesarch, M. A., Suyker, A. E., and Verma, S. B. (2009). Scaling up of Co2 fluxes from leaf to canopy in maize-based agroecosystems. *Agric. For. Meteorol.* 149, 2110–2119. doi:10.1016/j.agrformet.2009.04.013

Beló, A., Zheng, P., Luck, S., Shen, B., Meyer, D. J., Li, B., et al. (2008). Whole genome scan detects an allelic variant of fad2 associated with increased oleic acid levels in maize. *Mol. Genet. Genomics* 279, 1–10. doi:10.1007/s00438-007-0289-y

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

Cao, H., Imparl-Radosevich, J., Guan, H., Keeling, P. L., James, M. G., and Myers, A. M. (1999). Identification of the soluble starch synthase activities of maize endosperm. *Plant Physiol.* 120, 205–216. doi:10.1104/pp.120.1.205

Cao, Y., Zhong, Z., Wang, H., and Shen, R. (2022). Leaf angle: a target of genetic improvement in cereal crops tailored for high-density planting. *Plant Biotechnol. J.* 20, 426–436. doi:10.1111/pbi.13780

Chen, X., Xu, D., Liu, Z., Yu, T., Mei, X., and Cai, Y. (2015). Identification of qtl for leaf angle and leaf space above ear position across different environments and generations in maize (Zea mays L). *Euphytica* 204, 395–405. doi:10.1007/s10681-015-1351-1

Coram, M. A., Fang, H., Candille, S. I., Assimes, T. L., and Tang, H. (2017). Leveraging multi-ethnic evidence for risk assessment of quantitative traits in minority populations. *Am. J. Hum. Genet.* 101, 218–226. doi:10.1016/j.ajhg.2017.06.015

Darvasi, A., and Soller, M. (1997). A simple method to calculate resolving power and confidence interval of QTL map location. *Behav. Genet.* 27, 125–132. doi:10.1023/a:1025685324830

Deng, M., Li, D., Luo, J., Xiao, Y., Liu, H., Pan, Q., et al. (2017). The genetic architecture of amino acids dissection by association and linkage analysis in maize. *Plant Biotechnol. J.* 10, 1250–1263. doi:10.1111/pbi.12712

Ding, J., Zhang, L., Chen, J., Li, X., Li, Y., Cheng, H., et al. (2015). Genomic dissection of leaf angle in maize (Zea mays L.) using a four-way cross mapping population. *PLoS One* 10, e0141619. doi:10.1371/journal.pone.0141619

Ding, L., Zhao, K., Zhang, X., Song, A., Su, J., Hu, Y., et al. (2019). Comprehensive characterization of a floral mutant reveals the mechanism of hooked petal morphogenesis in Chrysanthemum morifolium. *Plant Biotechnol. J.* 17, 2325–2340. doi:10.1111/pbi.13143

Dong, H., Zhao, H., Xie, W., Han, Z., Li, G., Yao, W., et al. (2016). A novel tiller angle gene, TAC3, together with TAC1 and D2 largely determine the natural variation of tiller angle in rice cultivars. *PLoS Genet.* 12, e1006412. doi:10.1371/journal.pgen.1006412

Dzievit, M. J., Li, X., and Yu, J. (2019). Dissection of leaf angle variation in maize through genetic mapping and meta-analysis. *Plant Genome* 12, 180024. doi:10.3835/plantgenome2018.05.0024

Egger, M., Smith, G. D., and Phillips, A. N. (1997). Meta-analysis: principles and procedures. *BMJ* 315, 1533–1537. doi:10.1136/bmj.315.7121.1533

Eugster, M., Knaus, J., Porzelius, C., Schmidberger, M., and Vicedo, E. (2011). Hands-on tutorial for parallel computing with R. *Comput. Stat.* 26, 219–239. doi:10.1007/s00180-010-0206-4

Fukazawa, J., Teramura, H., Murakoshi, S., Nasuno, K., Nishida, N., Ito, T., et al. (2014). DELLAs function as coactivators of GAI-ASSOCIATED FACTOR1 in regulation of gibberellin homeostasis and signaling in Arabidopsis. *Plant Cell* 26, 2920–2938. doi:10.1105/tpc.114.125690

Gao, J., Yang, S., Cheng, W., Fu, Y., Leng, J., Yuan, X., et al. (2017). GmILPA1, encoding an APC8-like protein, controls leaf petiole angle in soybean. *Plant Physiol.* 174, 1167–1176. doi:10.1104/pp.16.00074

Goffinet, B., and Gerber, S. (2000). Quantitative trait loci: a meta-analysis. *Genetics* 155, 463–473. doi:10.1093/genetics/155.1.463

Harper, L., and Freeling, M. (1996). Interactions of liguleless1 and liguleless2 function during ligule induction in maize. *Genetics* 144, 1871–1882. doi:10.1093/genetics/144.4.1871

Hu, Y., Gu, Y., Xiao, Q., Hou, X., Zhang, X., et al. (2015). Genetic analysis for canopy architecture in an $F_{2:3}$ population derived from two-type foundation parents across multi-environments. *Euphytica* 205, 421–440. doi:10.1007/s10681-015-1401-8

Incognito, S., Maddonni, G., and López, C. (2020). Genetic control of maize plant architecture traits under contrasting plant densities. *Euphytica* 216, 20. doi:10.1007/s10681-019-2552-9

Ishikawa, S., Maekawa, M., Arite, T., Onishi, K., Takamure, I., and Kyozuka, J. (2005). Suppression of tiller bud activity in tillering dwarf mutants of rice. *Plant Cell Physiol.* 46, 79–86. doi:10.1093/pcp/pci022

Jeong, J. S., Kim, Y. S., Baek, K. H., Jung, H., Ha, S., Do Choi, Y., et al. (2010). Root-specific expression of OsNAC10 improves drought tolerance and grain yield in rice under field drought conditions. *Plant Physiol.* 153, 185–197. doi:10.1104/pp.110.154773

Khangura, R. S., Johal, G. S., and Dilkes, B. P. (2020). Variation in maize chlorophyll biosynthesis alters plant architecture. *Plant Physiol.* 184, 300–315. doi:10.1104/pp.20.00306

Kitakura, S., Vanneste, S., Robert, S., Löfke, C., Teichmann, T., Tanaka, H., et al. (2011). Clathrin mediates endocytosis and polar distribution of PIN auxin transporters in Arabidopsis. *Plant Cell* 23, 1920–1931. doi:10.1105/tpc.111.083030

Ku, L., Ren, Z., Chen, X., Shi, Y., Qi, J., Su, H., et al. (2016). Genetic analysis of leaf morphology underlying the plant density response by QTL mapping in maize (Zea mays L). *Mol. Breed.* 36, 63–16. doi:10.1007/s11032-016-0483-x

Ku, L., Wei, X., Zhang, S., Zhang, J., Guo, S., and Chen, Y. (2011). Cloning and characterization of a putative TAC1 ortholog associated with leaf angle in maize (Zea mays L). *PLoS One* 6, e20621. doi:10.1371/journal.pone.0020621

Ku, L., Zhang, J., Guo, S., Liu, H., Zhao, R., and Chen, Y. (2012). Integrated multiple population analysis of leaf architecture traits in maize (Zea mays L). *J. Exp. Bot.* 63, 261–274. doi:10.1093/jxb/err277

Ku, L., Zhao, W., Zhang, J., Wu, L., Wang, C., Wang, P., et al. (2010). Quantitative trait loci mapping of leaf angle and leaf orientation value in maize (Zea mays L). *Theor. Appl. Genet.* 121, 951–959. doi:10.1007/s00122-010-1364-z

Li, C., Li, Y., Shi, Y., Song, Y., Zhang, D., Buckler, E. S., et al. (2015). Genetic control of the leaf angle and leaf orientation value as revealed by ultra-high density maps in three connected maize populations. *PLoS One* 10, e0121624. doi:10.1371/journal.pone.0121624

Li, H., Peng, Z., Yang, X., Wang, W., Fu, J., Wang, J., et al. (2013). Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. *Nat. Genet.* 45, 43–50. doi:10.1038/ng.2484

Li, H., Wang, L., Liu, M., Dong, Z., Li, Q., Fei, S., et al. (2020). Maize plant architecture is regulated by the ethylene biosynthetic gene *ZmACS7*. *Plant Physiol.* 183, 1184–1199. doi:10.1104/pp.19.01421

Li, J., Zhang, Z., Li, Y., Wang, Q., and Zhou, Y. (2011). QTL consistency and meta-analysis for grain yield components in three generations in maize. *Theor. Appl. Genet.* 122, 771–782. doi:10.1007/s00122-010-1485-4

Li, P., Wang, Y., Qian, Q., Fu, Z., Wang, M., Zeng, D., et al. (2007). LAZY1 controls rice shoot gravitropism through regulating polar auxin transport. *Cell Res.* 17, 402–410. doi:10.1038/cr.2007.38

Li, Z., Tang, J., Srivastava, R., Bassham, D. C., and Howell, S. H. (2020). The transcription factor bZIP60 links the unfolded protein response to the heat stress response in maize. *Plant Cell* 32, 3559–3575. doi:10.1105/tpc.20.00260

Liu, B. (1998). *Statistical genomics: linkage, mapping and QTL analysis*. Boca Raton: CRC Press.

Liu, H., Luo, X., Niu, L., Xiao, Y., Chen, L., Liu, J., et al. (2016). Distant eQTLs and non-coding sequences play critical roles in regulating gene expression and quantitative trait variation in maize. *Mol. Plant* 10, 414–426. doi:10.1016/j.molp.2016.06.016

Liu, X., Hao, L., Kou, S., Su, E., Zhou, Y., and Wang, R. (2019). The combined effects of omitting confounders and measurement error on statistical inference of mediation and a new tool for sensitivity analysis. *Multivar. Behav. Res.* 39, 137–138. doi:10.1080/00273171.2019.1694478

Lu, M., Zhou, F., Xie, C., Li, M., Xu, Y., Marilyn, W., et al. (2007). Construction of a SSR linkage map and mapping of quantitative trait loci (QTL) for leaf angle and leaf orientation with an elite maize hybrid. *Yi Chuan* 29, 1131–1138. doi:10.1360/yc-007-1131

Lv, R., Li, Z., Li, M., Dogra, V., Lv, S., Liu, R., et al. (2019). Uncoupled expression of nuclear and plastid photosynthesis-associated genes contributes to cell death in a lesion mimic mutant. *Plant Cell* 31, 210–230. doi:10.1105/tpc.18.00813

Makinde, F. L., Tchamga, M. S. S., Jafali, J., Fatumo, S., Chimusa, E. R., Mulder, N., et al. (2021). Reviewing and assessing existing meta-analysis models and tools. *Brief. Bioinform.* 22, bbab324. doi:10.1093/bib/bbab324

Mitani, N., Chiba, Y., Yamaji, N., and Ma, J. (2009). Identification and characterization of maize and barley Lsi2-like silicon efflux transporters reveals a distinct silicon uptake system from that in rice. *Plant Cell* 21, 2133–2142. doi:10.1105/tpc.109.067884

Moreno, M. A., Harper, L. C., Krueger, R. W., Dellaporta, S. L., and Freeling, M. (1997). liguleless1 encodes a nuclear-localized protein required for induction of ligules and auricles during maize leaf organogenesis. *Genes Dev.* 11, 616–628. doi:10.1101/gad.11.5.616

Nagelkerke, N. (1991). A note on a general definition of the coefficient of determination. *Biometrika* 78, 691–692. doi:10.1093/biomet/78.3.691

Nyquist, E., and Baker, J. (1991). Estimation of heritability and prediction of selection response in Plant-populations. *Crit. Rev. Plant Sci.* 10, 235–322. doi:10.1080/07352689109382313

Pan, Q., Xu, Y., Li, K., Peng, Y., Zhan, W., Li, W., et al. (2017). The genetic basis of plant architecture in 10 maize recombinant inbred line populations. *Plant Physiol.* 175, 858–873. doi:10.1104/pp.17.00709

Peiffer, J. A., Romay, M. C., Gore, M. A., Flint-Garcia, S. A., Zhang, Z., Millard, M. J., et al. (2014). The genetic architecture of maize height. *Genetics* 196, 1337–1356. doi:10.1534/genetics.113.159152

Ren, Z., Wu, L., Ku, L., Wang, H., Zeng, H., Su, H., et al. (2020). ZmILI1 regulates leaf angle by directly affecting liguleless1 expression in maize. *Plant Biotechnol. J.* 18, 881–883. doi:10.1111/pbi.13255

Ruan, X., Ma, L., Zhang, Y., Wang, Q., and Gao, X. (2021). Dissection of the complex transcription and metabolism regulation networks associated with maize resistance to Ustilago maydis. *Genes (Basel)* 12, 1789. doi:10.3390/genes12111789

Saitoh, K., Yonetani, K., Murota, T., and Kuroda, T. (2002). Effects of flag leaves and panicles on light interception and canopy photosynthesis in high-yielding rice cultivars. *Plant Prod. Sci.* 5, 275–280. doi:10.1626/pps.5.275

Schmid, M., Davison, T. S., Henz, S. R., Pape, U. J., Demar, M., Vingron, M., et al. (2005). A gene expression map of *Arabidopsis thaliana* development. *Nat. Genet.* 37, 501–506. doi:10.1038/ng1543

Schmidt, P., Hartung, J., Rath, J., and Piepho, H. (2019). Estimating broad-sense heritability with unbalanced data from agricultural cultivar trials. *Crop Sci.* 59, 525–536. doi:10.2135/cropsci2018.06.0376

Sharma, R., Singh, G., Bhattacharya, S., and Singh, A. (2018). Comparative transcriptome meta-analysis of *Arabidopsis thaliana* under drought and cold stress. *PLoS One* 13, e0203266. doi:10.1371/journal.pone.0203266

Sharopova, N., McMullen, M. D., Schultz, L., Schroeder, S., Sanchez-Villeda, H., Gardiner, J., et al. (2002). Development and mapping of SSR markers for maize. *Plant Mol. Biol.* 48, 463–481. doi:10.1023/a:1014868625533

Shi, Y., Wang, X., Guo, S., Ren, Z., Ku, L., Zhu, Y., et al. (2017). Detection of epistatic and environmental interaction qtls for leaf orientation-related traits in maize. *Plant Breed.* 136, 33–40. doi:10.1111/pbr.12431

Stewart, D. W., Costa, C., Dwyer, L. M., Smith, D. L., Hamilton, R. I., and Ma, B. L. (2003). Canopy structure, light interception, and photosynthesis in maize. *Agron. J.* 95, 1465–1474. doi:10.2134/agronj2003.1465

Tang, D., Chen, Z., Ni, J., Jiang, Q., Liu, J., Wang, L., et al. (2021). Identification of qtl for leaf angle at canopy-wide levels in maize. *Euphytica* 217, 75. doi:10.1007/s10681-021-02781-4

Tang, W., Deng, Z., and Wang, Z. (2010). Proteomics shed light on the brassinosteroid signaling mechanisms. *Curr. Opin. Plant Biol.* 13, 27–33. doi:10.1016/j.pbi.2009.10.007

Tang, Y., Liu, H., Guo, S., Wang, B., Li, Z., Chong, K., et al. (2018). OsmiR396d affects gibberellin and brassinosteroid signaling to regulate plant architecture in rice. *Plant Physiol.* 176, 946–959. doi:10.1104/pp.17.00964

Tao, S., Chu, J., Liu, X., Zhang, R., Zhang, Z., and Luo, Z. (2002). High-resolution gene mapping using admixture linkage disequilibrium. *Chin. Sci. Bull.* 47, 1717–1719. doi:10.1007/bf03183315

Terasaka, K., Blakeslee, J. J., Titapiwatanakun, B., Peer, W. A., Bandyopadhyay, A., Makam, S. N., et al. (2005). PGP4, an ATP binding cassette P-glycoprotein, catalyzes auxin transport in *Arabidopsis thaliana* roots. *Plant Cell* 17, 2922–2939. doi:10.1105/tpc.105.035816

Thiemann, A., Fu, J., Seifert, F., Grant-Downton, R. T., Schrag, T. A., Pospisil, H., et al. (2014). Genome-wide meta-analysis of maize heterosis reveals the potential role of additive gene expression at pericentromeric loci. *BMC Plant Biol.* 14, 88. doi:10.1186/1471-2229-14-88

Tian, F., Bradbury, P. J., Brown, P. J., Hung, H., Sun, Q., Flint-Garcia, S., et al. (2011). Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nat. Genet.* 43, 159–162. doi:10.1038/ng.746

Tian, J., Wang, C., Xia, J., Wu, L., Xu, G., Wu, W., et al. (2019). Teosinte ligule allele narrows plant architecture and enhances high-density maize yields. *Science* 365, 658–664. doi:10.1126/science.aax5482

Tohge, T., de Souza, L. P., and Fernie, A. R. (2017). Current understanding of the pathways of flavonoid biosynthesis in model and crop plants. *J. Exp. Bot.* 68, 4013–4028. doi:10.1093/jxb/erx177

Veyrieras, J. B., Goffinet, B., and Charcosset, A. (2007). MetaQTL: a package of new computational methods for the meta-analysis of QTL mapping experiments. *BMC Bioinforma.* 8, 49. doi:10.1186/1471-2105-8-49

Walsh, J., Waters, C. A., and Freeling, M. (1998). The maize gene liguleless2 encodes a basic leucine zipper protein involved in the establishment of the leaf blade-sheath boundary. *Genes Dev.* 12, 208–218. doi:10.1101/gad.12.2.208

Wang, B., Lin, Z., Li, X., Zhao, Y., Zhao, B., Wu, G., et al. (2020). Genome-wide selection and genetic improvement during modern maize breeding. *Nat. Genet.* 52, 565–571. doi:10.1038/s41588-020-0616-3

Wang, H., Liang, K., Li, X., Hu, Y., Wu, H., Wang, Z., et al. (2017). QTL analysis of ear leaf traits in maize (Zea mays L.) under different planting densities. *Crop J.* 5, 387–395. doi:10.1016/j.cj.2017.05.001

Wang, J., Yan, C., Shi, D., Zhao, X., Yuan, C., Sun, Q., et al. (2021). The genetic base for peanut height-related traits revealed by a meta-analysis. *Plants (Basel)* 10, 1058. doi:10.3390/plants10061058

Xu, K., Chen, S., Li, T., Ma, X., Liang, X., Ding, X., et al. (2015). OsGRAS23, a rice GRAS transcription factor gene, is involved in drought stress response through regulating expression of stress-responsive genes. *BMC Plant Biol.* 15, 141. doi:10.1186/s12870-015-0532-3

Yang, G., Dong, Y., Li, Y., Wang, Q., Shi, Q., Zhou, Q., et al. (2015). Integrative detection and verification of QTL for plant traits in two connected RIL populations of high-oil maize. *Euphytica* 206, 203–223. doi:10.1007/s10681-015-1502-4

Yang, X., Gao, S., Xu, S., Zhang, Z., Prasanna, B. M., Li, L., et al. (2011). Characterization of a global germplasm collection and its potential utilization for analysis of complex quantitative traits in maize. *Mol. Breed.* 28, 511–526. doi:10.1007/s11032-010-9500-7

Yu, B., Lin, Z., Li, H., Li, X., Li, J., Wang, Y., et al. (2007). TAC1, a major quantitative trait locus controlling tiller angle in rice. *Plant J.* 52, 891–898. doi:10.1111/j.1365-313X.2007.03284.x

Yu, J., Pressoir, G., Briggs, W. H., Vroh Bi, I., Yamasaki, M., Doebley, J. F., et al. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* 8, 203–208. doi:10.1038/ng1702

Zhang, J., Ku, L., Han, Z., Guo, S., Liu, H., Zhang, Z., et al. (2014). The ZmCLA4 gene in the qLA4-1 QTL controls leaf angle in maize (Zea mays L). *J. Exp. Bot.* 65, 5063–5076. doi:10.1093/jxb/eru271

Zhang, K., Lv, F., Li, J., Wang, H., Yu, J., Li, W., et al. (2019). Genetic mapping of quantitative trait locus for the leaf morphological traits in a recombinant inbred line population by ultra-high–density maps across multi-environments of maize (Zea mays). *Plant Breed.* 139, 107–118. doi:10.1111/pbr.12749

Zhang, M., Jin, Y., Ma, Y., Zhang, Q., Wang, P., Jiang, N., et al. (2020). Identification of QTLs and candidate genes associated with leaf angle and leaf orientation value in maize (zea mays L.) based on GBS. *Trop. Plant Biol.* 1, 34–49. doi:10.1007/s12042-020-09270-3

Zhang, Z., Ersoz, E., Lai, C., Todhunter, R., Tiwari, H., Gore, M., et al. (2010). Mixed linear model approach adapted for genome-wide association studies. *Nat. Genet.* 42, 355–360. doi:10.1038/ng.546

Zhou, Q., Fu, Z., Liu, H., Guo, Z., Zhang, X., Gore, M., et al. (2021). Mining novel kernel size-related genes by pQTL mapping and multi-omics integrative analysis in developing maize kernels. *Plant Biotechnol J* 19 (8), 1489–1491. doi:10.1111/pbi.13634

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 20 June 2023
DOI 10.3389/fgene.2023.1180941



OPEN ACCESS

EDITED BY
Sundeep Kumar,
National Bureau of Plant Genetic Resources, Indian Council of Agricultural Research (ICAR), India

REVIEWED BY
Alsamman M. Alsamman,
Agriculture Research Center (ARC), Egypt
Jai Prakash Jaiswal,
G. B. Pant University of Agriculture and Technology, India

*CORRESPONDENCE
A. Arora,
romiarora@yahoo.com
Hari Krishna,
harikrishna.agri@gmail.com
K. K. Hazra,
kalikrishna123@gmail.com

RECEIVED 06 March 2023
ACCEPTED 23 May 2023
PUBLISHED 20 June 2023

CITATION
Gurumurthy S, Arora A, Krishna H, Chinnusamy V and Hazra KK (2023), Genotypic capacity of post-anthesis stem reserve mobilization in wheat for yield sustainability under drought and heat stress in the subtropical region. *Front. Genet.* 14:1180941. doi: 10.3389/fgene.2023.1180941



# Genotypic capacity of post-anthesis stem reserve mobilization in wheat for yield sustainability under drought and heat stress in the subtropical region

S. Gurumurthy[1,2], A. Arora[1]*, Hari Krishna[3]*, V. Chinnusamy[1] and K. K. Hazra[4]*

[1]Division of Plant Physiology, ICAR–Indian Agricultural Research Institute, New Delhi, India, [2]School of Water Stress Management, ICAR–National Institute of Abiotic Stress Management, Baramati, Maharashtra, India, [3]Division of Genetics, ICAR–Indian Agricultural Research Institute, New Delhi, India, [4]Crop Production Division, ICAR–Indian Institute of Pulses Research, Kanpur, Uttar Pradesh, India

Wheat productivity is severely affected by drought and heat stress conditions worldwide. Currently, stem reserve mobilization (SRM) is receiving increased attention as a trait that can sustain wheat yields under adverse environments. However, the significance of SRM in sustaining wheat yields under drought and heat stress conditions remains uncertain in the tropical climate of Indo-Gangetic Plain region. Therefore, this study aimed to investigate genotypic variations in SRM in wheat and their influence on yield sustainability under drought and heat stress environments. The experiment was designed in an alpha-lattice layout, accommodating 43 genotypes under four simulated environments [timely sown and well irrigated (non-stress); timely sown and water-deficit/drought stress; late-sown and well-irrigated crop facing terminally high temperature; and late-sown and water-deficit stress (both water-deficit and heat stress)]. The water-deficit stress significantly increased SRM (16%–68%, $p < 0.01$) compared to the non-stress environment, while the heat stress conditions reduced SRM (12%–18%). Both SRM and stem reserve mobilization efficiency exhibited positive correlations with grain weight (grain weight spike$^{-1}$) under all three different stress treatments ($p < 0.05$). Strong positive correlations between stem weight (at 12 days after anthesis) and grain weight were observed across the environments ($p < 0.001$); however, a significant positive correlation between stem weight and SRM was observed only with stress treatments. Results revealed that the SRM trait could effectively alleviate the impacts of water-deficit stress on yields. However, the SRM-mediated yield protection was uncertain under heat stress and combined water-deficit and heat stress treatments, possibly due to sink inefficiencies caused by high temperature during the reproductive period. Defoliated plants exhibited higher SRM than non-defoliated plants, with the highest increment observed in the non-stress treatment compared to all the stress treatments. Results revealed that wider genetic variability exists for the SRM trait, which could be used to improve wheat yield under drought stress conditions.

KEYWORDS

**stem reserve mobilization, drought stress, terminal heat stress, defoliation, grain yield, wheat**

# Introduction

Wheat is the most widely cultivated food crop globally, contributing to ~ approximately 20% of the daily dietary energy total calories and protein for 4.5 billion people (Gooding and Shewry, 2022). The Indo-Gangetic Plain (IGP) is one of the main wheat-producing regions in India and worldwide (Daloz et al., 2021). India is the second largest wheat producer globally, contributing 13% of the wheat supply and exporting 0.2 million tons annually (Zaveri, and Lobell, 2019). In the IGP region, wheat productivity is affected by drought and terminal heat stresses (Zaveri, and Lobell, 2019; Chowdhury et al., 2020). In tropical climates, heat stress at the terminal growth stage causes significant yield losses (21%–30%) in wheat under late-sown conditions (Wang H. et al., 2016). Similarly, midseason or terminal water-deficit stress or drought conditions threaten wheat productivity worldwide (Farooq et al., 2015). The primary yield-limiting factors under drought and heat stress environments include source limitations due to reduced photosynthate assimilation, oxidative damage to cells, membrane disruption, forced maturity, sink inefficiencies, and poor grain filling (Chowdhury et al., 2019). Consequently, there is an urgent need to develop stress-tolerant/adaptive cultivars by deploying stress-inducive functional traits to sustain wheat yields.

Stem-assimilate reserves serve as a carbon source for developing grains in wheat, and they are prominent when photosynthate assimilation during grain filling is affected by abiotic stress factors (Shirdelmoghanloo et al., 2016). The scale of stem reserve mobilization (SRM) in wheat varies from 10% to 20% of the total grain weight under non-stressed conditions, and it can increase up to 30%–50% under drought and heat stress conditions (Shearman et al., 2005; Ehdaie et al., 2008). Therefore, deploying the genetic potential of SRM could be one way of improving crops under adverse environments of tropical agro-regions, where wheat growth is more source-limited than in temperate regions (Cruz-Aguado et al., 1999). Our current understanding of the impact of traits such as stem-assimilate reserves, sink capacity and efficiencies, grain-filling duration, and senescence dynamics on SRM in wheat under different stress conditions in tropical regions is limited. Although the contribution of stem reserve assimilate toward grain development in wheat has been studied and reported by several researchers (Ehdaie et al., 2008; Zhu et al., 2009; Talukder et al., 2013; Khoshro et al., 2014), the comparative assessment of SRM under drought and heat stress (or combined drought and heat stress) on the same platform is presently lacking.

Plant functional traits, such as stem-assimilate reserve capacity, SRM, and stem reserve mobilization efficiency (SRE), have a significant impact on crop yield depending on genotypes and the intensity of stress (Ehdaie et al., 2006; Ruuska et al., 2006). To strategically deploy the SRM trait for crop improvement, it is necessary to have an improved understanding of its agronomic, physiological, and genetic basis under diverse stress conditions. Secondary functional traits, including SRM, membrane stability, photosynthetic rate, and grain weight stability, should be considered in wheat pre-breeding and breeding to alleviate abiotic stresses (Farooq et al., 2011). Thus, evaluating genotypic variations for SRM and SRE under stress conditions is crucial to identify the genotype(s)/breeding line(s) with higher levels of SRM-mediated stress-tolerant capacity and to comprehend the primary determinants of these processes.

Therefore, a field experiment was conducted in regards to the tropical climate of India (New Delhi) to assess the potential of SRM (stem reserve mobilization) in sustaining wheat yields under water-deficit and high-temperature stress conditions and to assess the genotypic variability for the trait. The sensitivity of traits of SRM and SRE under non-stress and stressful environments (water-deficit stress, heat stress, and combined water-deficit and heat stress) were evaluated using a panel of 43 wheat genotypes and their associations with grain yield and yield parameters, plant growth, and physiological attributes were determined. The major objectives of the study were to i) quantify the genotypic variations for the SRM trait and contribution of SRM toward grain development and sustaining wheat yield under drought and heat stress environments in tropics, ii) understand the physiological basis of SRM in wheat under drought and heat stress environments, iii) evaluate the interactive effect of genotype × environment on the SRM trait in wheat, and iv) quantify the scale of SRM under drought and heat stress environments in the absence of a primary photosynthesis organ during the grain-filling period (defoliation study).

# Materials and methods

## Site characteristics

The field experiment was carried out during the winter season of 2016–2017 and 2017–2018 at the ICAR–Indian Agricultural Research Institute (ICAR–IARI) research farm New Delhi, India (28°41′ N and 77°13′ E, 228 m above the sea level). The site has a sub-humid tropical climate, and weather conditions during the cropping period are shown in Supplementary Figure S1. The weather variables were recorded from a meteorological observatory located at the ICAR–IARI research farm. Maximum and minimum temperatures were recorded from the maximum and minimum temperature thermometers installed inside the Stevenson screen, while the evaporation rate and rainfall were recorded using a USWB Class-A pan evaporimeter and a rain gauge, respectively. The experimental soil belongs to the *Fluvisol* order (World Reference Base soil classification). The soil had pH 8.05 and organic carbon 4.3 g $kg^{-1}$, and the available nitrogen, phosphorus, and potassium were 117.6, 7.4, and 132.4 mg $kg^{-1}$, respectively.

## Treatment description and experimental design

In this study, a panel of 43 wheat genotypes with contrasting traits was selected (Supplementary Table 1). The genotypes were chosen based on their contrasting heat tolerance and susceptibility (6 + 6 genotypes), drought tolerance and susceptibility (7 + 7 genotypes), plant height (four tall and five dwarf genotypes), maturity duration (four early- and four late-maturing genotypes), RILs (9 + 2 parents), current best standard genotypes or best-adapted varieties (six genotypes), and popular ruling varieties from the last century (10 genotypes). The list of genotypes under different groups is shown in the Supplementary Material. The genotypes were evaluated under four different growing conditions: timely sown well-irrigated condition (non-stress), timely sown deficit irrigation condition (water-deficit stress), late-sown and well-irrigated condition

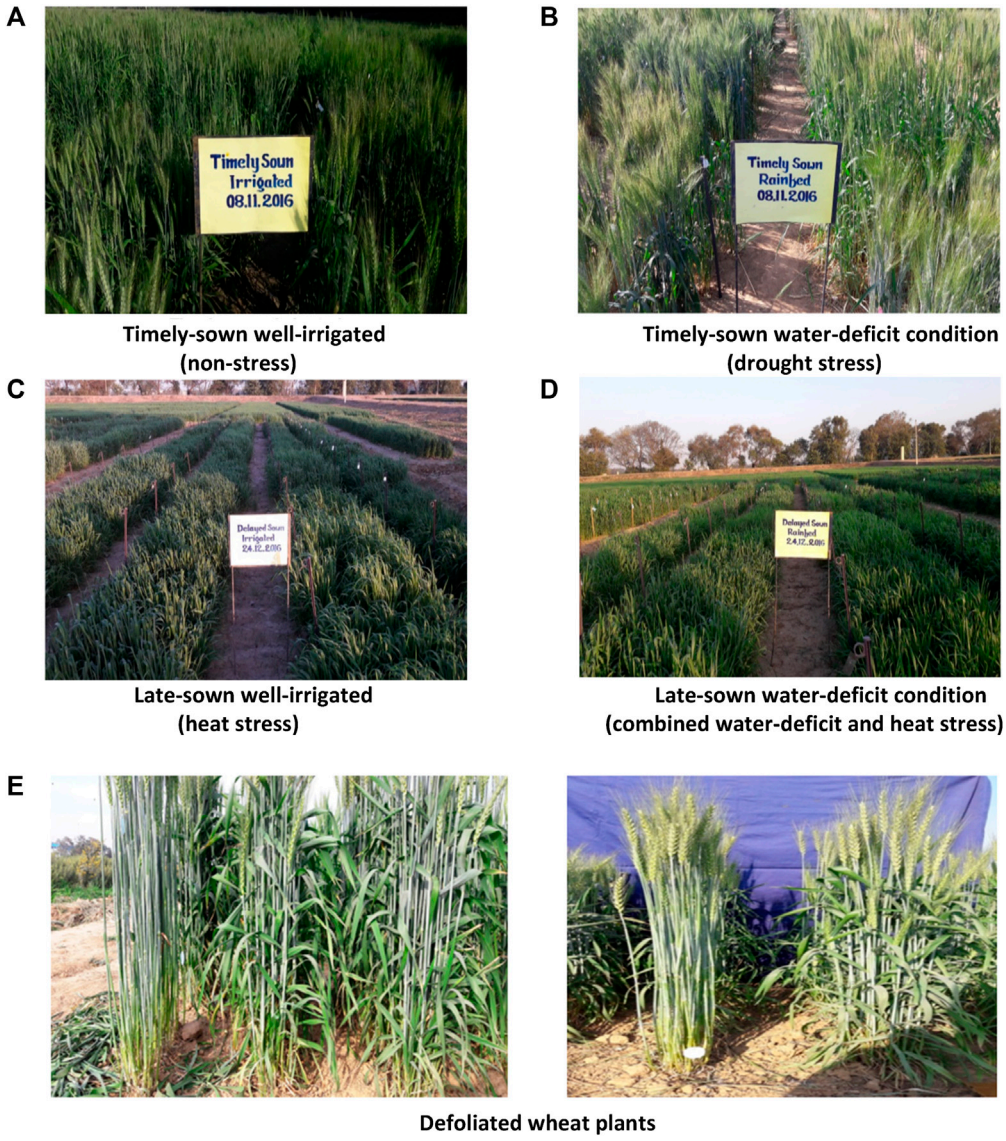



FIGURE 1
Field view of wheat crop under different environments **(A–D)** and defoliated wheat plants **(E)**.

(terminal heat stress), and late-sown deficit irrigation condition (combined terminal heat and drought stress) (Figure 1). The experiment was laid out following an alpha-lattice design with two replications (Patterson and Williams, 1976). Each genotype was sown manually in a gross plot size of 4 m × 1 m with an inter-row spacing of 23.0 cm. The timely sown and late-sown crops were sown on 8 November and 23 December, respectively, in both years. The late-sown crops were sown 45 days later than the timely sown crops to expose them to a terminal heat stress environment. Water-deficit stress was imposed by scheduling/withholding irrigation. In treatments with well-irrigated conditions, irrigation was applied at critical growth stages, including crown root initiation, active tillering, jointing, flowering, and dough stages (8 ha–cm each irrigation). In treatments with deficit-water conditions, life-saving irrigation (4 ha–cm each irrigation) was applied at crown root initiation, active tillering, jointing, and flowering stages, but no irrigation was applied at the dough stage.

In the year 2016–2017, from each plot of the experimental field, a complete row of 4 m length was defoliated by cutting off all the leaf blades at 12 days after anthesis (Figure 1). The objective of the study was to assess the changes in SRM in defoliated plants as compared to the non-defoliated (intact plants). This study was a part of the main experiment conducted during the year 2016–2017, and therefore, all the crop-growing environment treatments, panel of genotypes, and crop management practices were similar to that of the main experiment as described previously.

## Measurements

### Soil moisture content

Soil moisture content (w/w) was determined by the gravimetric method. Soil samples from each experimental unit were collected using a post-hole auger, and soils were placed in aluminum boxes

with secure lids. The samples were weighed immediately and then oven-dried at 105°C for 72 h, and soil moisture content (%) was calculated using the following formula:

$$\text{Soil moisture (\%)} = \frac{Weight\ of\ moist\ soil\ (g) - Weight\ of\ oven\ dry\ soil\ (g)}{Weight\ of\ oven\ dry\ soil\ (g)} \times 100. \quad (1)$$

## Crop phenology and growth attributes

Crop phenological events such as days to anthesis and days to maturity were recorded for each genotype/breeding line. The anthesis stage was determined visually for each genotype when anther extrusion occurs in 50% of the ears. The days taken from the sowing to anthesis stage was denoted as days to anthesis. The maturity (physiological maturity) of each genotype was determined both by visual (when all the leaves and spike turned to a complete yellowish color) and sensor-based observations (SPAD 502 chlorophyll meter reading and normalized differential vegetation index were shown in the range of <5 and <0.20, respectively) (reference). SPAD meter reading and normalized differential vegetation index were measured with the instrument SPAD 502 m (Konica–Minolta, Japan) and handheld Ntech "Greenseeker" (field portable NDVI sensor), respectively. The days taken from anthesis to physiological maturity was calculated and denoted as the reproductive period (days).

Five primary tillers (main stem) of each genotype were collected 12 days after anthesis. Then, the stem, leaf, and ear parts of the tillers were separated. The length of every single stem was recorded, and the stem part was oven-dried (65°C for 72 h) and dry-weight was recorded. The stem-specific weight (g $cm^{-1}$) was calculated as the ratio of stem dry weight (g) to its length (cm). Likewise, at the physiological maturity, again five primary tillers were sampled in each genotype and stem weight and length were recorded. The dry-weight of the grain in a spike from the five selected primary tillers was recorded at the time of crop maturity. The ratio of grain weight per spike to the total weight of the main tiller (including spike) was calculated to determine the harvest index (HI) and it is expressed as percentage.

For the defoliation experiment (2016–2017), similar observations (main stem weight) at 12 days after anthesis and physiological maturity were recorded from defoliated plants. Grain weight per spike was recorded and grain weight percentage at physiological maturity was calculated like non-defoliated plants mentioned previously.

## Temperature intensity and growing degree day

The average ambient maximum temperature (°C) during flowering (*i.e.*, at anthesis ±5 days) was calculated for each genotype to determine the temperature intensity during the flowering stage. Likewise, for each genotype, the average ambient maximum temperature (°C) during the reproductive period, *i.e.*, anthesis to physiological maturity was calculated using the daily maximum temperature data. The mean ambient maximum temperature (°C) during the anthesis and reproductive period of an environment represents the average of all genotypes. Growing degree days from anthesis to physiological maturity were calculated using the following formula (Santiveri et al., 2002):

$$Growing\ degree\ day\ (degree\ C) = \sum\left(\frac{Tmax + Tmin}{2} - base\ temperature\right), \quad (2)$$

where *T*max and *T*min represent the maximum and minimum temperatures, respectively. Here, the base–temperature (0°C) defines the minimum threshold temperature below which the crop development is ceased.

## Calculation of stem reserve mobilization and stem reserve mobilization efficiency

Ten uniform plants of each genotype (those plants were selected randomly where flowers appear on the same days) within a plot were tagged for sampling. Five tagged plants were sampled at 12 days after anthesis, and the remaining five plants were collected at crop maturity. According to reports, the maximum stem weight and carbohydrate concentration occur at 10–14 days after anthesis, and after that, there is a rapid decline in carbohydrate concentration in stems (Lopatecki et al., 1962; Ford et al., 1979). Stem reserve mobilization represents the amount of dry matter mobilized from stem to grain. The SRM value was calculated by the difference of stem dry weight at 12 days after anthesis and maturity (Eq. 3). Stem reserve mobilization efficiency was calculated as the proportional contribution percentage (%) of mobilized stem reserve to stem weight at 12 days after anthesis (Ehdaie et al., 2006) (Eq. 4).

$$\text{SRM (g)} = [Stem\ weight\ at\ twelve\ days\ after\ anthesis\ (g) - Stem\ weight\ at\ maturity\ (g)], \quad (3)$$

$$\text{SRE (\%)} = \frac{Stem\ weight\ at\ twelve\ days\ after\ anthesis\ (g) - Stem\ weight\ at\ maturity\ (g)}{Stem\ weight\ at\ twelve\ days\ after\ anthesis\ (g)} \times 100. \quad (4)$$

In Eq. 3, the contribution of SRM to grain weight was calculated as the grain weight percentage by using the following formula:

$$\text{Grain weight percentage (\%)} = \frac{Stem\ reserve\ mobilation\ (g)}{Weight\ of\ grains\ per\ spike\ (g)} \times 100. \quad (5)$$

For the defoliation study (2016–2017), SRM and SRE were estimated following the same procedure as mentioned previously.

## Grain yield estimation

A net area of 1.35 $m^2$ (3 m row length × 0.45 m) was harvested from each genotype plot to estimate the grain yield. A sub-sample from the harvested grains was analyzed for moisture content (%, w/w), and the grain yield was adjusted to 12% moisture content (w/w) and expressed as kg $ha^{-1}$. The grain yield loss of each genotype under different stress conditions was calculated by the difference of yield of timely sown well-

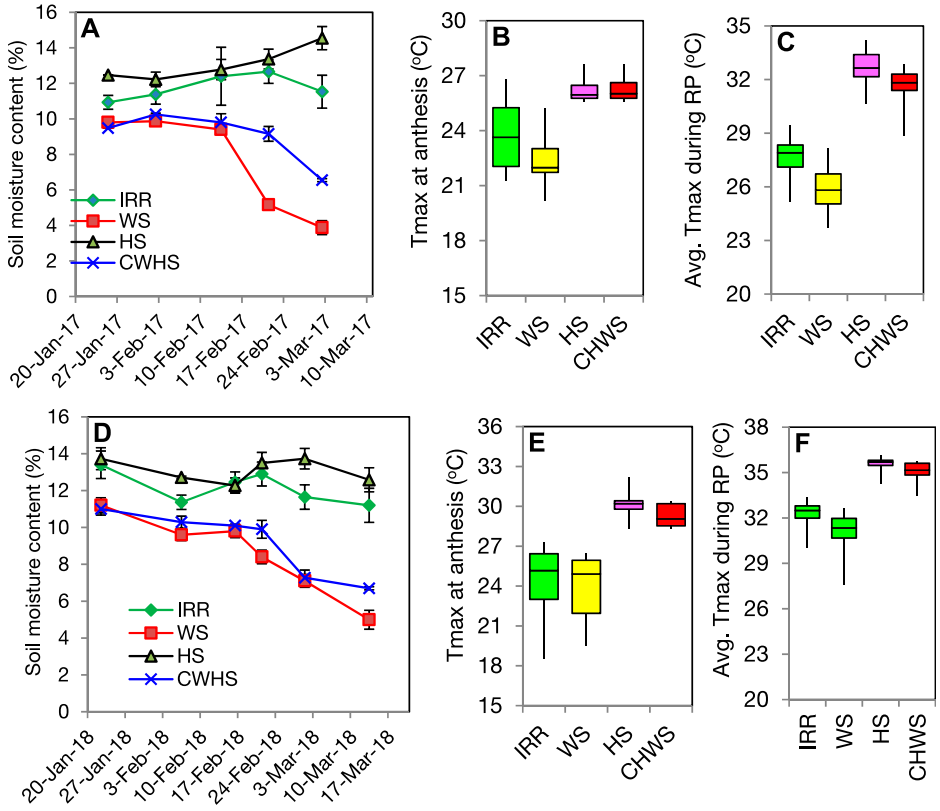



FIGURE 2
Temporal dynamics of soil moisture (w/w) during crop season of 2016–2017 **(A)** and 2017–2018 **(D)**. Box plot presentation of ambient maximum temperature (Tmax, oC) at anthesis stage of wheat as influenced by different crop growing environments in the year 2016–2017 **(B)** and 2017–2018 **(E)**. Box plot presentation of ambient maximum temperature (Tmax, oC) during the reproductive period (anthesis to maturity) of wheat crop as influenced by different crop growing environments in the year 2016-2017 **(C)** and 2017–2018 **(F)**. IRR, timely sown well-irrigated (non-stress); WS, water-deficit stress; HS, heat stress; and CWHS, combined water-deficit and heat stress.

irrigated condition (non-stress) treatment and stress treatments (drought stress, heat stress, and combined drought and heat stress).

## Statistical analysis

The pooled analysis of variance was analyzed considering the year ($n$ = 2), environments ($n$ = 4), and genotypes ($n$ = 43) as the main, sub, and sub–sub factors, respectively. The META–R program (multi environment trail analysis with R for Windows) was used for computing by fitting mixed and fixed linear models from the alpha-lattice experimental design (Alvarado et al., 2020). For the alpha-lattice design the model is explained as

$$Y_{ijkl} = \mu + Loc_i + Rep_j\ (Loc_i) + Block_k\left(Loc_i Rep_j\right) + Gen_l + Loc_i \times Gen_l + Cov + \varepsilon_{ijkl}, \quad (5a)$$

where $Y_{ijkl}$ is the trait of interest, $\mu$ is the overall mean effect, $Loc_i$ is effects of the *ith* environment, $Loc_i \times Gen_k$ are the effects environment × genotype (G × E) interaction, $Rep_j$ is the effect of the *jth* replicate, $Block_k$ the effect of the *kth* incomplete block, $Gen_l$ is the effect of the *lth* genotype, Cov is the effect of the covariate, and $\varepsilon_{ijkl}$ is the effect of the error associated with the *ith* environment, *jth* replication, *kth* incomplete block, and *lth* genotype.

The mean comparison of the different environments was performed following Tukey's test using SPSS statistical software (version 20.0.6). The linear regression analysis was performed using the R statistical package (version 3.5.2). Box plots of SRM and SRE parameters for non-defoliated crops were graphically presented using PAST statistical software (version 3.14). The mean comparison of plant traits and SRM of defoliated and non-defoliated plants was conducted by a paired *t*-test. Genotype + genotype-by-environment (GGE) biplots were developed using the R studio platform using the "GGEbiplotGUI" package (Yan and Tinker, 2006). The GGEbiplots were constructed by plotting the first two principal components derived by subjecting the mean values to singular-value decomposition (Yan and Tinker, 2006). To display the mean performance and stability of a genotype, the biplots were framed with the mean vs. stability function by adopting no scaling (scale = 0), tester–centred G + GE (centring = 2) with genotype focused (row metric preserving) singular-value partitioning (SVP = 1) (Parihar et al., 2018), (Eqs 5, 5a).

**TABLE 1 Growth and yield attributing parameters of wheat as influenced by non-stressed and stressed environments.**

| Year | Parameter | Environment | | | |
|---|---|---|---|---|---|
| | | Irrigated (non-stress) | Water-deficit stress | Heat stress | Combined water-deficit and heat stress |
| 2016–2017 | Plant height (cm) | 93.8 ± 3.5[a] | 87.0 ± 4.1[b] | 86.3 ± 3.15[b] | 76.6 ± 2.68[c] |
| | Single stem weight at 12 DAA (g) | 2.10 ± 0.12[a] | 1.87 ± 0.11[b] | 1.52 ± 0.08[c] | 1.48 ± 0.07[c] |
| | Single stem weight at maturity (g) | 1.80 ± 0.11[a] | 1.46 ± 0.07[b] | 1.27 ± 0.07[c] | 1.15 ± 0.06[c] |
| | Stem-specific weight (g $cm^{-1}$) | 22.03 ± 1.01[a] | 22.48 ± 0.99[a] | 17.39 ± 0.73[c] | 19.12 ± 0.80[b] |
| | Reproductive period (day) | 42.2 ± 1.7[a] | 38.9 ± 1.1[b] | 29.8 ± 0.9[c] | 28.5 ± 0.7[c] |
| | Grain weight per spike (g) | 2.68 ± 0.11[a] | 2.25 ± 0.09[b] | 1.55 ± 0.09[c] | 1.66 ± 0.07[c] |
| | Grain weight percentage (%) | 11.5 ± 2.7[c] | 22.4 ± 2.4[a] | 15.8 ± 2.3[bc] | 20.2 ± 2.3[ab] |
| | Harvest index (%) | 0.38 ± 0.02[a] | 0.37 ± 0.02[a] | 0.23 ± 0.02[c] | 0.32 ± 0.01[b] |
| | Growing degree days (°C-day)[#] | 843.4 ± 28.1[a] | 739.8 ± 18.1[bc] | 763.1 ± 14.8[b] | 705.8 ± 16.8[c] |
| 2017–2018 | Plant height (cm) | 84.8 ± 2.6[a] | 70.9 ± 1.8[b] | 61.9 ± 3.7[c] | 59.7 ± 4.0[c] |
| | Single stem weight at 12 DAA (g) | 2.30 ± 0.09[a] | 2.03 ± 0.08[b] | 1.32 ± 0.06[c] | 1.19 ± 0.05[d] |
| | Single stem weight at maturity (g) | 1.91 ± 0.09[a] | 1.57 ± 0.06[b] | 0.95 ± 0.04[c] | 0.78 ± 0.03[d] |
| | Stem-specific weight (g $cm^{-1}$) | 26.11 ± 0.91[a] | 26.48 ± 0.73[a] | 20.39 ± 0.71[b] | 19.72 ± 0.78[b] |
| | Reproductive period (day) | 43.14 ± 1.62[a] | 39.00 ± 1.03[b] | 30.11 ± 0.89[c] | 28.73 ± 0.65[c] |
| | Grain weight per spike (g) | 2.64 ± 0.09[a] | 2.32 ± 0.09[b] | 2.12 ± 0.07[c] | 1.48 ± 0.05[d] |
| | Grain weight percentage (%) | 15.39 ± 2.46[c] | 19.63 ± 1.98[b] | 17.49 ± 1.26[bc] | 27.70 ± 1.68[a] |
| | Harvest index (%) | 0.37 ± 0.02[a] | 0.36 ± 0.02[a] | 0.20 ± 0.03[d] | 0.29 ± 0.02[c] |
| | Growing degree days (°C-day) | 889.1 ± 22.1[a] | 769.8 ± 14.4[b] | 768.1 ± 16.9[b] | 713.2 ± 11.20[c] |

*DAA*, days after anthesis; the values represent the mean ±95% confidence interval.
#Growing degree-days value represents to the cumulative degree days during reproductive period (anthesis to physiological maturity);
[a]Different uppercase letters within the row values are the significant difference at $p = 0.05$ according to Tukey's test.
[b]Different uppercase letters within the row values are the significant difference at $p = 0.05$ according to Tukey's test.
[c]Different uppercase letters within the row values are the significant difference at $p = 0.05$ according to Tukey's test.
[d]Different uppercase letters within the row values are the significant difference at $p = 0.05$ according to Tukey's test.

**TABLE 2 Best linear unbiased estimators of parameters stem reserve mobilization, stem reserve mobilization efficiency, stem weight, grain weight spike$^{-1}$, and grain yield based on 2-year experimental data.**

| Environment | Statistic | SRM | SRE | Stem weight | Grain weight spike$^{-1}$ | Grain yield |
|---|---|---|---|---|---|---|
| OVERALL | Heritability | 0.497215 | 0.564039 | 0.73249 | 0.568307 | 0.363365 |
| | Genotype variance | 0.002719 | 6.867062 | 0.020052 | 0.010692 | 0.007774 |
| | Genotype × local variance | 0.008539 | 7.98844 | 0.038534 | 0.013855 | 0.159308 |
| | Residual variance | 0.026915 | 68.94685 | 0.0401 | 0.102242 | 0.318399 |
| | Grand mean | 0.376819 | 22.19486 | 1.738171 | 2.083643 | 2.512193 |
| | Least significant difference | 0.103215 | 4.848527 | 0.20395 | 0.189365 | 0.225201 |
| | Coefficient of variation | 43.53793 | 37.41148 | 11.52075 | 15.34589 | 22.46121 |
| | Replicates ($n$) | 2 | 2 | 2 | 2 | 2 |
| | Environments ($n$) | 8 | 8 | 8 | 8 | 8 |
| | Genotype significance ($p$-value) | 0.00124 | $7.94 \times 10^{-05}$ | $2.69 \times 10^{-11}$ | $6.28 \times 10^{-05}$ | 0.0339698 |
| | Genotype × environment significance ($p$-value) | $1.24 \times 10^{-05}$ | 0.067198 | $1.53 \times 10^{-21}$ | 0.032653 | $2.61 \times 10^{-06}$ |

**TABLE 3 Effect of different crop-growing environments on stem and grain attributes of defoliated wheat plants.**

| Parameter | Environment | | | |
|---|---|---|---|---|
| | Irrigated (non-stress) | Water-deficit stress | Heat stress | Combined water-deficit and heat stress |
| Single stem weight at 12 DAA (g) | 2.10 ± 0.12[a] | 1.87 ± 0.11[b] | 1.52 ± 0.08[c] | 1.48 ± 0.07[c] |
| Single stem weight at maturity (g) | 1.63 ± 0.10[a] | 1.43 ± 0.08[b] | 1.25 ± 0.07[c] | 1.14 ± 0.06[c] |
| Grain weight per spike (g) | 1.71 ± 0.04[a] | 1.55 ± 0.04[b] | 1.16 ± 0.04[d] | 1.32 ± 0.03[c] |
| Grain weight percentage (%) | 28.6 ± 8.1[ab] | 35.7 ± 8.3[a] | 23.6 ± 5.4[b] | 27.1 ± 6.5[b] |

[a]Different uppercase letters within the row values are the significant difference at $p = 0.05$ according to Tukey's test. *DAA*, days after anthesis.
[b]Different uppercase letters within the row values are the significant difference at $p = 0.05$ according to Tukey's test. *DAA*, days after anthesis.
[c]Different uppercase letters within the row values are the significant difference at $p = 0.05$ according to Tukey's test. *DAA*, days after anthesis.
[d]Different uppercase letters within the row values are the significant difference at $p = 0.05$ according to Tukey's test. *DAA*, days after anthesis.

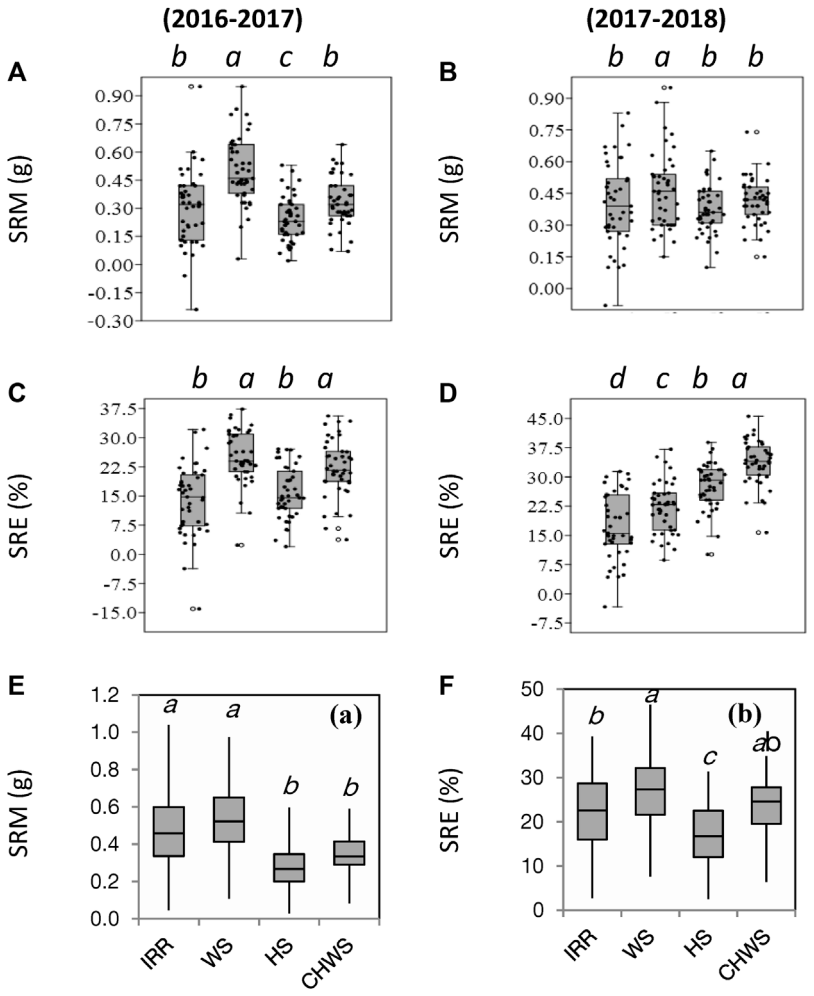



FIGURE 3
Box plot presentation of parameters stem reserve mobilization (SRM) **(A, C)** and stem reserve mobilization efficiency (SRE) **(B, D)** as influenced by different crop growing environments (2016–2017 and 2017–2018). Box plot presentation of stem reserve mobilization (SRM) **(E)** and stem reserve efficiency (SRE) **(F)** of defoliated wheat plants as influenced by different crop growing environments in the year 2016–2017 (coloured graphs). IRR, irrigated (non-stress); WS, water-deficit stress; HS, heat stress; and CWHS, combined water-deficit and heat stress. Each data point represents a single genotype. a–d, different lowercase letters corresponding to the treatments represent significantly different at $p < 0.05$ as per Tukey's test.

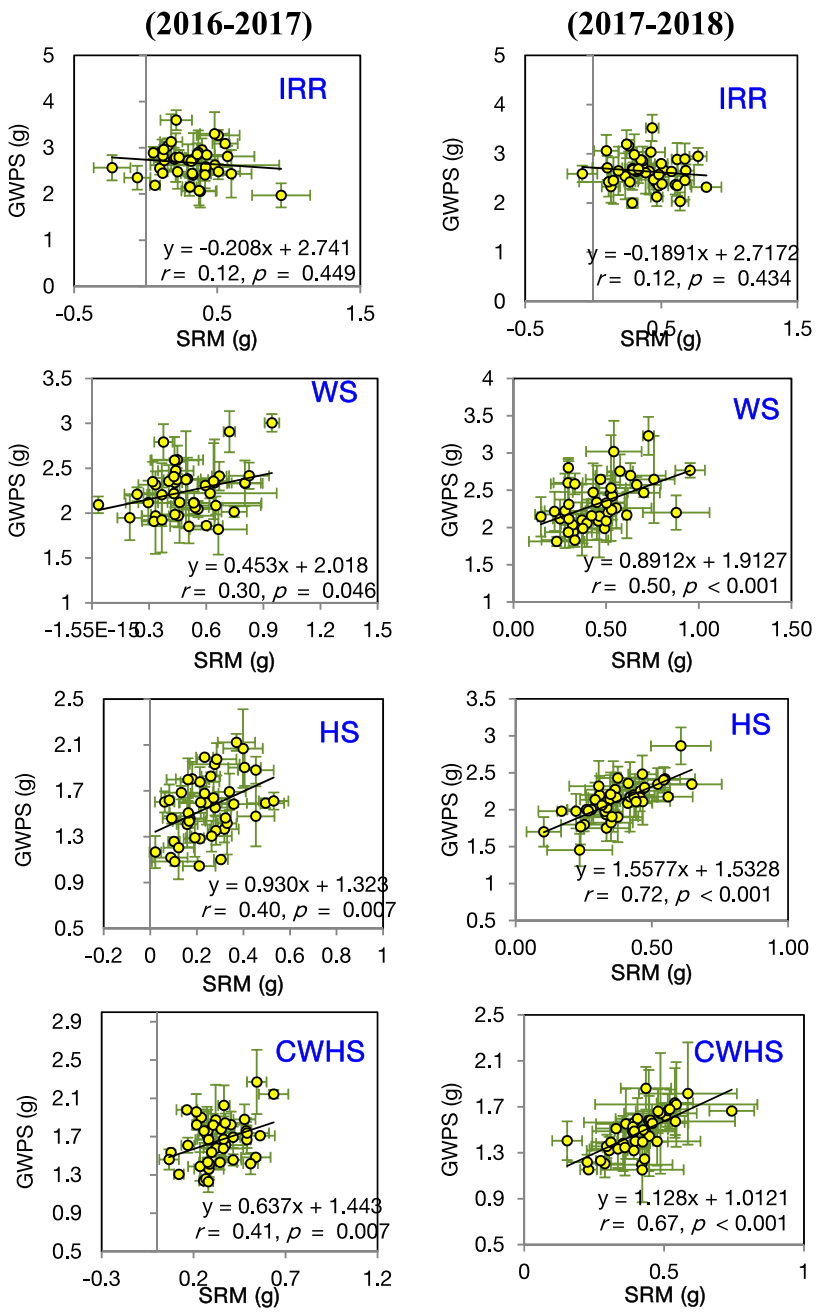



FIGURE 4
Linear relationship between stem reserve mobilization (SRM) and grain weight spike$^{-1}$ (GWPS) in the year 2016–2017 and 2017–2018 under different crop growing environments. IRR, irrigated (non-stress); WS, water-deficit stress; HS, heat stress; and CWHS, combined water-deficit and heat stress. The vertical and horizontal error bars represents the corresponding standard error of means.

# Results

## Weather condition and soil moisture

In the cropping year (2016–2017), soil moisture content during the grain-filling period (anthesis to maturity) was recorded to be the highest in the heat-stress treatment (13.1%), followed by timely sown irrigated (11.4%), combined water and heat stress (8.8%) and water-deficit stress (6.3%) treatments (Figure 2). Likewise, in the year 2017–2018, the soil moisture during the grain-filling period followed the same treatment order, *i.e.*, heat stress (12.9%) ≥ timely sown irrigated (12.3%) > combined water and heat stress (8.3%) > water-deficit stress (5.8%). Soil moisture gradient within the treatments was prominent after the anthesis stage and recorded as the highest at the end of the crop season (or maturity). In both cropping years, the ambient maximum temperature during flowering was higher in the heat stress and combined water-deficit and heat stress treatments (Figure 2), while the mean maximum temperatures during the reproductive period were significantly higher in the heat stress and combined water and

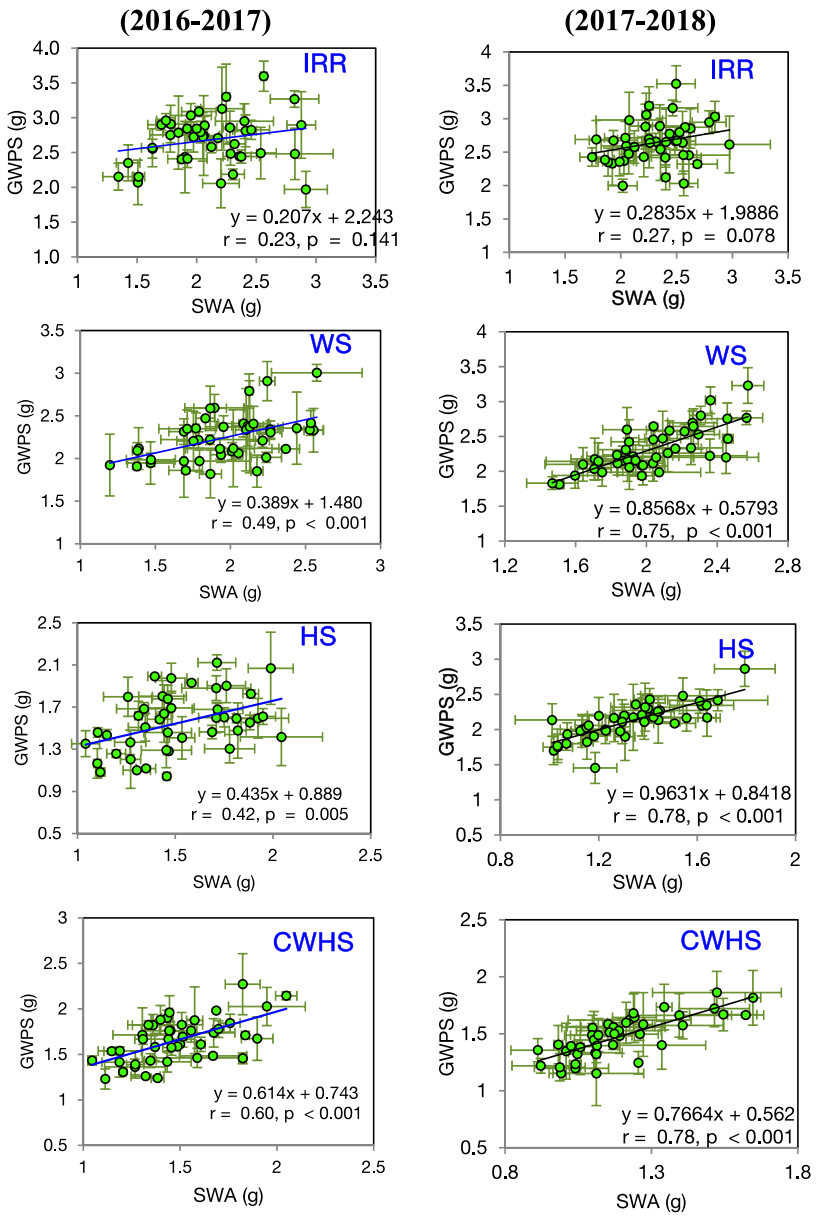



FIGURE 5
Linear correlations between single stem (main stem) weight at 12 days after anthesis (SWA) and grain weight spike$^{-1}$ (GWPS) under different crop growing environments in the year 2016–2017 and 2017–2018. IRR, irrigated (non-stress); WS, water-deficit stress; HS, heat stress; and CWHS, combined water-deficit and heat stress. The vertical and horizontal error bars represents the corresponding standard error of means.

heat stress treatments over the water-deficit stress, and non-stress (timely sown irrigated) treatments.

## Impact of drought and terminal heat stress on wheat crop

Water-deficit stress had reduced plant height (7%–16%, $p < 0.01$), stem weight (anthesis) (11%–12%, $p < 0.01$), reproductive period (8%, $p < 0.05$), and grain weight spike$^{-1}$ (12%–16%, $p < 0.01$) as compared to the non-stress environment (Table 1). The heat-stress treatment reduced the plant height, stem weight, stem-specific weight, reproductive period, and grain weight spike$^{-1}$ by 8%–27%, 28%–42%, 21%–22%, 29%–32%, and 20%–42%, respectively ($p < 0.05$), compared to the non-stress environment. Likewise, the corresponding reductions with the combined water-deficit and heat stress environment with reference to the non-stress environment were 18%–30%, 30%–48%, 13%–25%, 32%–33%, and 38%–39% ($p < 0.01$). Compared to the non-stress environment, the reductions in grain yield with water-deficit stress, heat stress, and combined water-deficit and heat stress environments were 26%–28%, 56%–61%, and 73%–79%, respectively ($p < 0.01$). The best linear unbiased estimator (BLUE) model explained that the genotype × environment

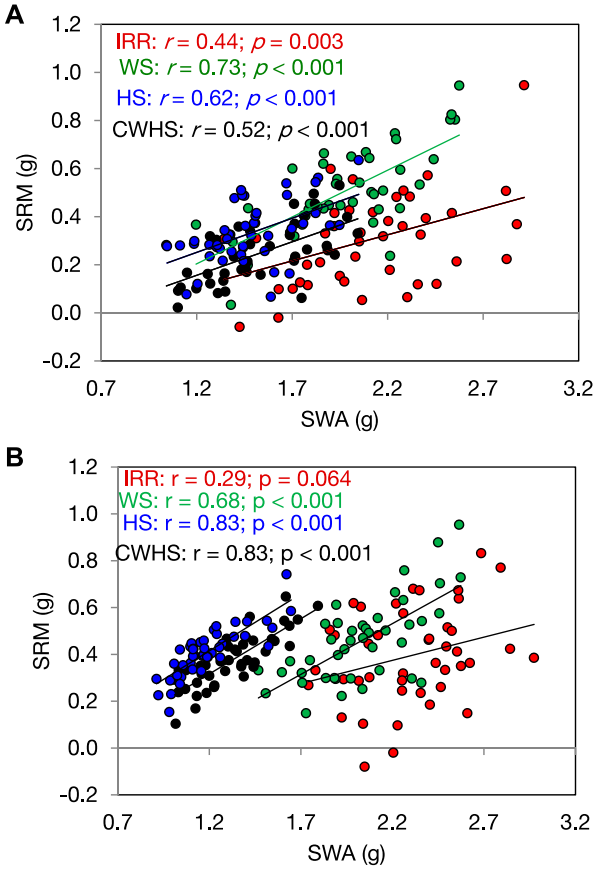



FIGURE 6
Relationship between stem weight [main stem weight at 12 days after anthesis (SWA)] and stem reserve mobilization (SRM) under different crop growing environments in the year 2016–2017 **(A)** and 2017–2018 **(B)**. IRR, irrigated (non-stress); WS, water-deficit stress; HS, heat stress; and CWHS, combined water-deficit and heat stress.

interaction was highly significant on stem weight, grain weight spike$^{-1}$, and grain yield ($p < 0.001$) (Table 2).

The grain weight spike$^{-1}$ in defoliated plants was lower than in the non-defoliated plants (paired *t*-test, $p < 0.05$), and the highest reduction was recorded in the non-stress environment (25%) followed by water-deficit (19%) and heat stress (13%) and the lowest in the combined water-deficit and heat stress environment (7%) ($p < 0.05$) (Table 3).

## Stem reserve mobilization and its relation with yield parameters

Water-deficit stress increased SRM by 68% and 16% over the non-stress environment (timely sown well irrigated) in the year 2016–2017 and 2017–2018, respectively ($p < 0.05$) (Figure 3). The heat stress treatment reduced SRM (12%–18%) compared to the non-stress environment, but the reduction was significant only in the year 2016–2017 ($p < 0.05$). SRM values were comparable within the non-stress, heat stress, and combined water-deficit and heat stress treatments in the year 2017–2018. The water-deficit stress, heat stress, and combined water-deficit and heat stress environments increased SRE by 28%–80%, 13%–61%, and 60%–95%, respectively, over the non-stressed environment ($p < 0.05$). The BLUE model explained that the genotype × environment interaction was highly significant on SRM ($p < 0.001$) (Table 2). As compared to the non-defoliated plants, the mean SRM values were higher in the defoliated plants, and the incremental change in SRM over the non-defoliated plants was found in the order non-stress (59%) heat stress (10%), water-deficit stress (7%), and combined water-deficit and heat stress (4%) environments (Figure 3E).

In both years, SRM and SRE had a non-significant correlation with grain weight spike$^{-1}$ in the non-stress environment (Figure 4). Under the stress environments (water-deficit stress, heat stress, and combined water-deficit and heat stress), SRM exhibited a significant positive correlation with grain weight spike$^{-1}$ ($p < 0.05$). The correlation coefficient (*r*) between SRM and grain weight spike$^{-1}$ was the highest under the heat stress environment. Stem weight (at 12 DAA) and grain weight spike$^{-1}$ were positively correlated under the stress environments, but not under the non-stress environment (Figure 5). Meanwhile, the significant correlations between stem weight (anthesis) and SRM were observed under all stress environments, and the results were consistent in both years (Figure 6). Under the water-deficit stress conditions, with an increase in SRM there was a significant reduction in the yield loss (yield loss calculated comparing with non-stress environment) (Figure 7). Under the heat stress environment, the SRM trait exhibited a negative relation with the yield loss in the year 2016–2017 ($p < 0.05$), but not in 2017–2018. The relationship between yield loss and SRM under the combined water-deficit and heat stress environments was non-significant in both years.

## Genotypic variability and genotype × environment interaction

Noticeable genotypic variations were observed for the SRM trait. Particularly, under the water-deficit stress environment, SRM varied between 0.15 and 0.95 g (2016–2017) and 0.03–0.95 g (2017–2018); SRE values ranged between and 3%–37% and 9%–37%. Under the non-stress environment, genotypes PBW 343, WH 542, GCP 30, and HD 2864 consistently had a higher SRM (>0.5 g stem$^{-1}$) and SRE (>30%) over the remaining genotypes; under the water-deficit stress condition, the genotypes HI 8627, C 306, WH 542, HD 4728, RAJ 3765, and GCP 30 were found to have a higher SRM (>0.6 g stem$^{-1}$) and SRE (>30%) over the remaining genotypes. Considering the parameter yield loss (<15%), SRM (>0.4 g stem$^{-1}$) and SRE (>25%) across the year, the genotypes C 306, GCP 29, and GCP 1 were found to be the best genotypes under the water-deficit stress environment. GGEbiplot analysis revealed that there was a differential sensitivity of the selected genotypes for the trait SRM with a strong G × E interaction (Figure 8). As shown in the mean stability graph, notable variation in the SRM trait was apparent within the experimental year (Figure 8).

## Other important correlations

Stem-specific weight exhibited positive associations with SRE and SRM across the environment (Table 4). A non-significant

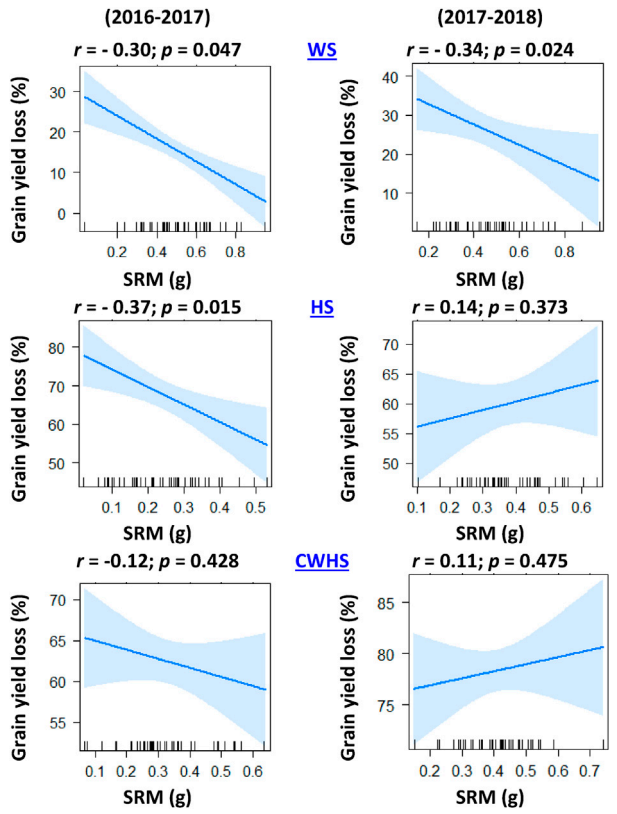



FIGURE 7
Linear regression trend line explaining the association between grain yield loss (loss as compared to irrigated non-stress treatment) and stem reserve mobilization (SRM) under different stressed conditions in the year 2016–2017 and 2017–2018. IRR, irrigated (non-stress); WS, water-deficit stress; HS, heat stress; and CWHS, combined water-deficit and heat stress.

association was observed between SRM and grain weight spike$^{-1}$ in defoliated plants; however, stem-specific weight had a strong positive correlation with grain weight spike$^{-1}$ but not with SRM. Similar to the non-defoliated plant, stem weight at anthesis had a positive association with SRM in defoliated plants across the environment (Supplementary Table S2).

## Discussion

The soil moisture data indicate that the crop experienced severe soil moisture-deficit stress (<10% w/w) during the post-anthesis stages in both water-deficit stress and combined water-deficit and heat stress treatments. Furthermore, during the reproductive stages, the ambient temperatures in the heat stress and combined water-deficit and heat stress treatments were significantly higher (23.9–29.7°C) than the optimal temperature for wheat (12–22°C) (Dwivedi et al., 2017). Thus, the results confirm that the terminal drought and heat stress conditions were effectively induced in this study.

Our findings demonstrate that water-deficit stress can enhance SRM and SRE in wheat. This aligns with previous studies that have also reported increased SRM in wheat under water-deficit stress when compared to well-irrigated non-stress conditions (Asseng and Van Herwaarden, 2003; Ruuska et al., 2006; Thapa et al., 2022). Under both drought and high-temperature environments, post-anthesis photosynthesis is affected, resulting in an increased reliance on pre-anthesis soluble carbohydrate reserves, especially from the stem (Ruuska et al., 2006; Tatar et al., 2016). Meanwhile, drought conditions have been shown to accelerate the process of reserve mobilization to grain (~10 days earlier) when compared to well-irrigated conditions (Ehdaie et al., 2008). In our study, the wider genotypic variations observed in both SRM and SRE under water-stress environments, as compared to other environment, suggest that the inherent capacity of plants to mobilize reserves is heavily influenced by crop-growing environments. Conversely, heat stress conditions resulted in reduced SRM, which may be attributed to reduced stem biomass at anthesis and significant reduction in the grain-filling duration (Ovenden et al., 2017; Islam et al., 2021). Our results indicate that wheat is more sensitive to terminal heat stress in a subtropical climate than to drought stress. Furthermore, the negative effects of heat stress on sink capacity may also be an important contributing factor to the lower SRM under heat stress environment.

The correlation analysis reveals that stem-assimilate reserve capacity (or stem biomass), stem-specific weight, and grain-filling duration are influential factors for SRM in wheat. Notably, a significant correlation between SRM and main stem grain weight was only observed under stress treatments. This suggests that the significance of the SRM trait for grain development is limited to stress conditions, and despite an increased scale of SRM under non-stress conditions, its relative contribution toward grain development is minimal. In the absence of abiotic stress factors, normal plant functions facilitate higher current photosynthate assimilation during the post-anthesis period, largely contributing toward developing grain; subsequently, the contribution of SRM is reduced. All such individual growth factors (fertilization, crop rotation, and plant protection measures) that directly or indirectly influence plant growth and development determine SRM. Therefore, the SRM trait may serve as a conditional functional trait, and its relative significance for grain development is dependent on several factors, such as the nature and intensity of stress factors, stem reserve capacity, phenological stability (especially the grain-filling duration), and sink capacity.

The significant negative correlation between SRM and grain yield loss (compared to the non-stress condition) highlights the potential of SRM as a trait for improving yield sustainability, particularly under water-deficit stress conditions. However, the contribution of SRM toward yield sustainability under heat stress conditions remains uncertain from the study results. Despite the strong positive relationship between SRM and grain weight spike$^{-1}$ under heat stress environments (*i.e.*, heat stress and combined water-deficit and heat stress), the influence of SRM on yield may have been impacted by the oversized influence of heat stress on sink inefficiencies and effective tiller production (data not presented). Previous studies in the IGP regions have reported inefficiencies at the sink development in late-sown crops, which are attributed to reduced functionality of reproductive organs such as pollen viability and embryo abortion (Bheemanahalli et al., 2019; Chowdhury et al., 2020).

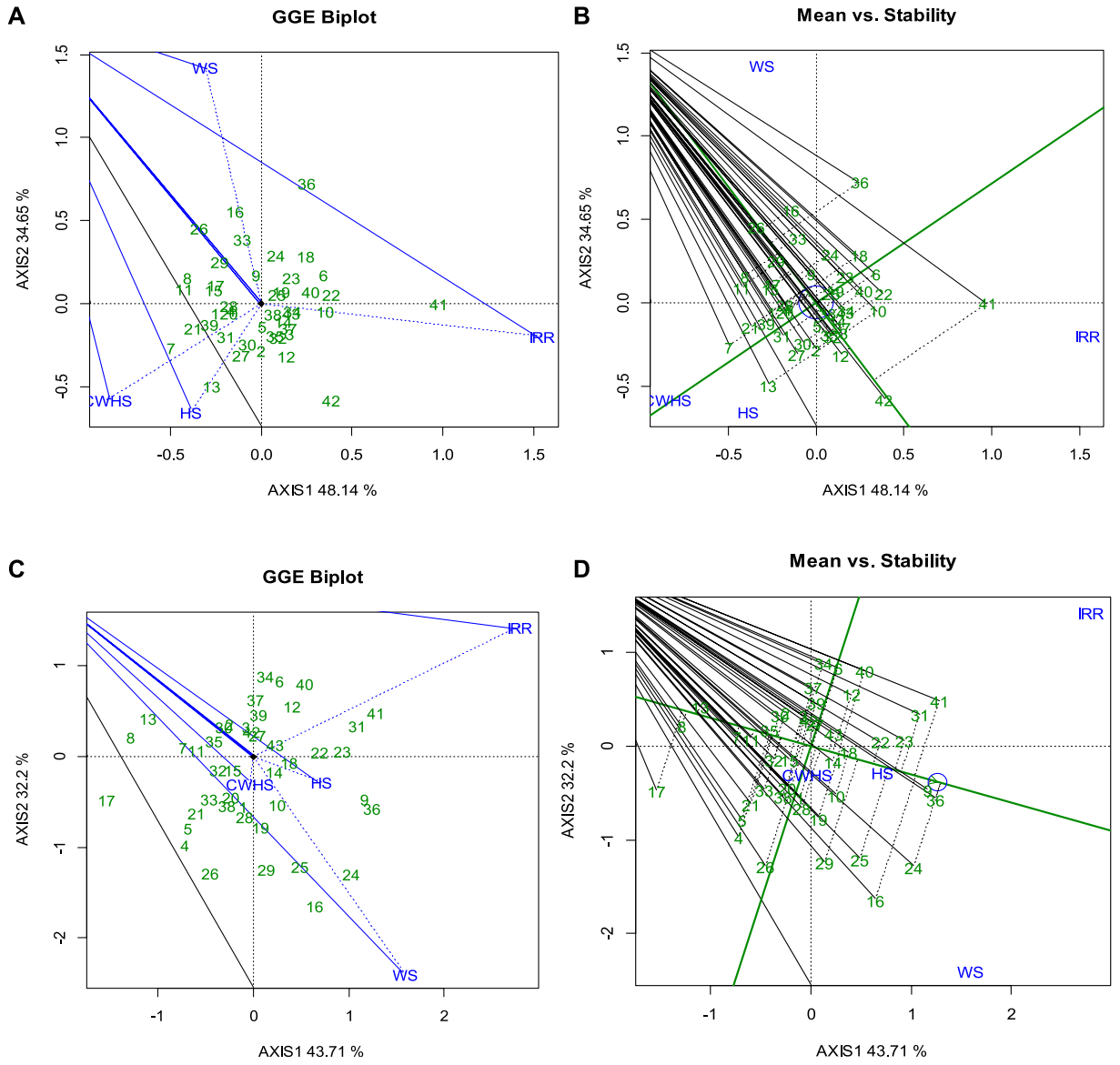



FIGURE 8
GGEBiplot presentation to explain the genotype × environment interactions for the trait SRM in the year 2016–2017 **(A)** and 2017–2018 **(C)**; Mean stability of wheat genotypes for SRM trait under different environments in the year 2016–2017 **(B)** and 2017–2018 **(D)**. IRR, irrigated (non-stress); WS, water-deficit stress; HS, heat stress; and CWHS, combined water-deficit and heat stress.

The imbalance in source-sink relations, particularly the reduction in the sink, may be a reason for the noticeable yield loss under heat stress environment. Studies have reasoned that an optimal balance in source and sink has a greater role in the mobilization of stored carbohydrates (Yu et al., 2015; Sehgal et al., 2018; Rivera-Amado et al., 2020). It is possible that under heat stress and combined water-deficit and heat stress environments, the impact of the environment on yield was much greater than that of the genotype, which might have confounded the correlation results.

The results of the defoliation study suggest that limiting current photosynthate assimilation during post-anthesis stages may trigger the reserve mobilization process (Liu et al., 2018); however, according to results, an increase in SRM does not essentially translate to grain development. Defoliation of wheat plant after anthesis may have caused major inefficiencies in physiological functions that have a direct impact on stem reserves mobilization efficiency. Notably, the maximum reduction in grain weight spike$^{-1}$ in defoliated plants was 36%, with grain weight percentages ranging between 24% and 36%, indicating that active photosynthesis by the stem, awn, and other photosynthetically active plant parts may have contributed to grain development. Saeidi et al. (2012) and Zhang et al. (2020) suggested that spike photosynthesis has a greater role in yield formation under stressed conditions than carbohydrate remobilization and leaf photosynthesis. Therefore, the actual values of SRM may be higher than measured, and the observed SRM may be somewhat underestimated.

**TABLE 4 Pearson correlation coefficients (*r*) of stem reserve mobilization parameters with growth and yield parameters under different crop growing environments.**

| Year | Environment | Parameter | SWM | STHA | SSWA | GDD | GWP | HI | [GW/SW] |
|---|---|---|---|---|---|---|---|---|---|
| 2016–2017 | Irrigated (non-stress) | SRM | −0.09 | −0.10 | 0.69*** | −0.27 | 0.97*** | 0.11 | 0.02 |
| | | SRE | −0.35* | −0.27 | 0.54*** | −0.18 | 0.90*** | 0.30* | 0.26 |
| | Water-deficit stress | SRM | 0.28 | 0.20 | 0.69*** | −0.19 | 0.93*** | −0.25 | −0.12 |
| | | SRE | −0.16 | −0.04 | 0.54*** | −0.22 | 0.90*** | −0.04 | 0.22 |
| | Heat stress | SRM | 0.21 | 0.12 | 0.65*** | 0.10 | 0.92*** | 0.41** | 0.18 |
| | | SRE | −0.12 | −0.11 | 0.50*** | 0.09 | 0.91*** | 0.48** | 0.38* |
| | Combined water-deficit and heat stress | SRM | −0.04 | −0.06 | 0.67*** | −0.01 | 0.91*** | 0.29 | 0.33* |
| | | SRE | −0.44** | −0.33* | 0.48** | −0.18 | 0.92*** | 0.35* | 0.56*** |
| 2017–2018 | Irrigated (non-stress) | SRM | −0.38* | −0.24 | 0.58*** | −0.20 | 0.96*** | 0.09 | 0.30* |
| | | SRE | −0.61*** | −0.36* | 0.50*** | −0.20 | 0.94*** | 0.12 | 0.48** |
| | Water-deficit stress | SRM | 0.05 | −0.10 | 0.39** | −0.26 | 0.93*** | −0.36* | 0.50*** |
| | | SRE | −0.25 | −0.26 | 0.34* | −0.30 | 0.95*** | −0.26 | 0.60*** |
| | Heat stress | SRM | 0.41** | −0.19 | 0.36* | −0.01 | 0.93*** | −0.08 | 0.26 |
| | | SRE | −0.05 | −0.43** | 0.24 | −0.02 | 0.90*** | −0.01 | 0.59*** |
| | Combined water-deficit and heat stress | SRM | 0.40** | −0.02 | 0.67*** | −0.08 | 0.88*** | 0.28 | 0.25 |
| | | SRE | −0.21 | −0.07 | 0.49*** | −0.08 | 0.85*** | 0.30* | 0.59*** |

*SRM,* stem reserve mobilization; *SRE,* stem reserve mobilization efficiency; *SWM,* single stem weight at maturity; *STHA,* stem height at anthesis; *SSWA,* stem-specific weight at anthesis; *LDR,* leaf senescence rate; *LSD,* leaf senescence duration; *GDD,* cumulative growing degree days; *GWP,* grain weight percentage; *HI,* harvest index; *GW/SW,* grain/stem weight ratio. The correlation (*r*) values highlighted with green, yellow, and pink colors are significant at $p < 0.001$, $p < 0.01$, and $p < 0.05$, respectively.

The study showed that the traits SRM and SRE exhibited wider variations among genotypes under all environments, with higher variations under water-deficit stress conditions. This variability could be attributed to differences in stem-soluble carbohydrate concentrations among the genotypes (Ehdaie et al., 2006; Ruuska et al., 2006; Liu et al., 2020; Vosoghi Rad et al., 2022). Soluble carbohydrates in the stem have been found to mobilize into the grain during the grain-filling period, and stem biomass after anthesis is directly related to the total water-soluble carbohydrate in the stem under water-deficit conditions (Saint Pierre et al., 2010). According to Ruuska et al. (2006), the variation in total water-soluble carbohydrate was attributed mainly to variations in the fructan component, with the other major soluble carbohydrates, sucrose, and hexose varying less. Therefore, the study suggests that the ability to accumulate and remobilize stem reserves is an important trait for improving grain filling in wheat under water-limited environments. The SRM trait could be particularly useful for developing drought-tolerant cultivars; however, its significance under tropical heat stress conditions may be limited.

Previous research showed that genotypes with high potential for utilizing stem reserves for developing grains under stress conditions may be associated with accelerated leaf senescence dynamics (Fokar et al., 1998; Islam et al., 2021; Salem et al., 2021). Therefore, investigating the relationships of SRM with grain-filling duration, leaf senescence dynamics, and source and sink components under different environmental conditions could provide valuable insights. Additionally, molecular studies could better characterize the trait and to understand the mechanism that could help in developing wheat varieties with improved SRM efficiency using appropriate bio-engineering approaches. The heritability estimates of 49% and 56% for the traits SRM and SRE, respectively, suggest that a significant proportion of the observed genotypic variations were due to the genotypic effects (Table 2). Thus, developing and deploying wheat varieties using the SRM trait could be a promising approach to address the challenges of abiotic stresses in tropical agro-regions. The GGEbiplot analysis was found effective in categorizing wheat genotypes based on the SRM trait and representing genotype × environment interactions. According to GGEbiplot analysis, the genotype-by-environment interaction was notably higher (represented by the Axis 2 component) suggesting that the genotypic expression of SRM is highly environment specific. Therefore, unique genotypes should be selected, and distinct selection strategies may be employed for different environments. The genotype(s) with higher and stable SRM capacity across different environments would be ideal to select as valuable genetic resources for the future breeding program. According to the results, genotype 41 (HD 2864) and genotypes 10 (GCP 6) could serve as potential genetic resources for SRM as a trait, as these genotypes exhibited higher SRM in both years.

Overall, stem reserve mobilization is a significant trait for improving the tolerance of plants to water-deficit stress conditions, which is important for the sustainability of agriculture and food security in the face of climate change and other environmental challenges. The impact of weather

variability and climate change on wheat yields, therefore, remains central to food security concerns and directly affects the livelihood of small-scale farmers who control the majority of the landholdings in India and produce 41% of India's food grains (Zaveri, and Lobell, 2019). Furthermore, wheat productivity in the IGP region is largely dependent on irrigation, and the evident depletion of groundwater resources poses a major threat to wheat production in the region. Recent evidence points to irrigation as one of the key factors in explaining wheat yield gaps across the western and eastern parts of the Indo-Gangetic Plain (Jain et al., 2017). In some parts, the IGP regions are already facing moderate-to-severe water scarcity. In that consideration, enhancing plant SRM capacity could be an economic and realistic approach. Therefore, improving stem reserve mobilization in crops would be an important goal for plant breeders and geneticists, as it can lead to more resilient and productive plants under water-limited environments.

## Conclusion

The study concluded that the SRM trait has an increased significance in sustaining wheat yield particularly under stressful tropical environments (water-deficit stress in particular). Water-deficit stress had enhanced SRM over the non-stress environment. Results suggested SRM and stem biomass (or stem reserve capacity) positively influenced the grain development (measured as grain weight spike$^{-1}$) under stressful environments, but not under the non-stress (timely sown well irrigated) condition. However, the SRM-mediated grain yield (yield area$^{-1}$) protection was only evident under the water-deficit stress but not under the heat stress environment, demonstrating the significance of the SRM trait for yield advantage is not certain under heat stress environments, which is possibly attributed to sink inefficiencies and physiological dysfunctions. The defoliation study revealed that in the absence of post-anthesis photosynthesis by leaves, the contribution of SRM toward grain development was substantially increased (36%), being higher under the non-stress environment over the stress environment. The wider genotypic variations in the SRM trait and the yield protection attributed to SRM under water-deficit (drought) stress validated the approach, which needs to be integrated in the wheat-breeding program.

## Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

## Author contributions

SG, AA, and VC conceptualized the overall idea. SG, KH, and AA drafted the manuscript. SG, KH, and HK analyzed the data. SG and HK performed the plant observations. SG, AA, KH, HK, and VC edited and finalized the manuscript. All authors contributed to the article and approved the submitted version.

## Funding

The part of the research was supported under the fund from the DST-SERB project and ICAR-BMGF project (OPP1194767). This research was funded by ICAR–IARI.

## Acknowledgments

The authors thankfully acknowledge the research facilities and manpower provided by ICAR–Indian Institute of Agricultural research (ICAR–IARI), New Delhi. The authors would also like to thank Dr. K. V. Prabhu and Dr. Debashish Chakraborty, ICAR–IARI, New Delhi, for their valuable suggestions and guidelines for successful experimentation.

## 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/fgene.2023.1180941/full#supplementary-material

## References

Alvarado, G., Rodríguez, F. M., Pacheco, A., Burgueño, J., Crossa, J., Vargas, M., et al. (2020). META–R: A software to analyze data from multi–environment plant breeding trials. *Crop J.* 8 (5), 745–756. doi:10.1016/j.cj.2020.03.010

Asseng, S., and Van Herwaarden, A. F. (2003). Analysis of the benefits to wheat yield from assimilates stored prior to grain filling in a range of environments. *Plant Soil* 256 (1), 217–229. doi:10.1023/A:1026231904221

Bheemanahalli, R., Sunoj, V. J., Saripalli, G., Prasad, P. V., Balyan, H. S., Gupta, P. K., et al. (2019). Quantifying the impact of heat stress on pollen germination, seed set, and grain filling in spring wheat. *Crop Sci.* 59 (2), 684–696. doi:10.2135/cropsci2018.05.0292

Chowdhury, A. R., Ghosh, M., Lal, M., Pal, A., Hazra, K. K., Acharya, S., et al. (2020). Foliar spray of synthetic osmolytes alleviates terminal heat stress in late–sown wheat. *Int. J. Plant Prod.* 14, 321–333. doi:10.1007/s42106–020–00087–4

Chowdhury, A. R., Ghosh, M., Lal, M., Pal, A., Hazra, K. K., Acharya, S. S., et al. (2019). Efficacy of calcium chloride and arginine foliar spray in alleviating terminal heat stress in late–sown wheat (*Triticum aestivum* L). J. Agric. Sci. (Camb.) 157 (6), 537–549. doi:10.1017/S002185961900087X

Cruz–Aguado, J. A., Reyes, F., Rodes, R., Perez, I., and Dorado, M. (1999). Effect of source–to–sink ratio on partitioning of dry matter and $^{14}$C–photoassimilates in wheat during grain filling. *Ann. Bot.* 83 (6), 655–665. doi:10.1006/anbo.1999.0869

Daloz, A. S., Rydsaa, J. H., Hodnebrog, Ø., Sillmann, J., van Oort, B., Mohr, C. W., et al. (2021). Direct and indirect impacts of climate change on wheat yield in the Indo-Gangetic plain in India. *J. Agric. Food Res.* 4, 100132. doi:10.1016/j.jafr.2021.100132

Dwivedi, S. K., Basu, S., Kumar, S., Kumar, G., Prakash, V., Kumar, S., et al. (2017). Heat stress induced impairment of starch mobilisation regulates pollen viability and grain yield in wheat: Study in Eastern Indo–Gangetic Plains. *Field Crops Res.* 206, 106–114. doi:10.1016/j.fcr.2017.03.006

Ehdaie, B., Alloush, G. A., Madore, M. A., and Waines, J. G. (2006). Genotypic variation for stem reserves and mobilization in wheat: I. Postanthesis changes in internode dry matter. *Crop Sci.* 46 (5), 735–746. doi:10.2135/cropsci2005.04–0033

Ehdaie, B., Alloush, G. A., and Waines, J. G. (2008). Genotypic variation in linear rate of grain growth and contribution of stem reserves to grain yield in wheat. *Field Crops Res.* 106 (1), 34–43. doi:10.1016/j.fcr.2007.10.012

Farooq, M., Bramley, H., Palta, J. A., and Siddique, K. H. (2011). Heat stress in wheat during reproductive and grain–filling phases. *Crit. Rev. Plant Sci.* 30 (6), 491–507. doi:10.1080/07352689.2011.615687

Farooq, S., Shahid, M., Khan, M. B., Hussain, M., and Farooq, M. (2015). Improving the productivity of bread wheat by good management practices under terminal drought. *J. Agron. Crop Sci.* 201 (3), 173–188. doi:10.1111/jac.12093

Fokar, M., Blum, A., and Nguyen, H. T. (1998). Heat tolerance in spring wheat. II. Grain filling. *Euphytica* 104 (1), 9–15. doi:10.1023/A:1018322502271

Ford, M. A., Blackwell, R. D., Parker, M. L., and Austin, R. B. (1979). Associations between stem solidity, soluble carbohydrate accumulation and other characters in wheat. *Ann. Bot.* 44 (6), 731–738. doi:10.1093/annbot/44.6.731

Islam, M. A., De, R. K., Hossain, M. A., Haque, M. S., Uddin, M. N., Fakir, M. S. A., et al. (2021). Evaluation of the tolerance ability of wheat genotypes to drought stress: Dissection through culm-reserves contribution and grain filling physiology. *Agronomy* 11 (6), 1252. doi:10.3390/agronomy11061252

Jain, M., Singh, B., Srivastava, A. A. K., Malik, R. K., McDonald, A. J., and Lobell, D. B. (2017). Using satellite data to identify the causes of and potential solutions for yield gaps in India's Wheat Belt. *Environ. Res. Lett.* 12 (9), 094011. doi:10.1088/1748-9326/aa8228

Khoshro, H. H., Taleei, A., Bihamta, M. R., Shahbazi, M., Abbasi, A., and Ramezanpour, S. S. (2014). Expression analysis of the genes involved in accumulation and remobilization of assimilates in wheat stem under terminal drought stress. *Plant Growth Reg.* 74 (2), 165–176. doi:10.1007/s10725–014–9908–x

Liu, Y., Yang, X., Tian, D., Cong, R., Zhang, X., Pan, Q., et al. (2018). Resource reallocation of two grass species during regrowth after defoliation. *Front. Plant Sci.* 9, 1767. doi:10.3389/fpls.2018.01767

Liu, Y., Zhang, P., Li, M., Chang, L., Cheng, H., Chai, S., et al. (2020). Dynamic responses of accumulation and remobilization of water soluble carbohydrates in wheat stem to drought stress. *Plant Physiol. biochem.* 155, 262–270. doi:10.1016/j.plaphy.2020.07.024

Lopatecki, L. E., Longair, E. L., and Kasting, R. (1962). Quantitative changes of soluble carbohydrates in stems of solid–and hollow–stemmed wheats during growth. *Can. J. Bot.* 40 (9), 1223–1228. doi:10.1139/b62–113

Ovenden, B., Milgate, A., Lisle, C., Wade, L. J., Rebetzke, G. J., and Holland, J. B. (2017). Selection for water-soluble carbohydrate accumulation and investigation of genetic× environment interactions in an elite wheat breeding population. *Theor. Appl. Genet.* 130, 2445–2461. doi:10.1007/s00122-017-2969-2

Parihar, A. K., Basandrai, A. K., Kushwaha, K. P. S., Chandra, S., Singh, K. D., Bal, R. S., et al. (2018). Targeting test environments and rust–resistant genotypes in lentils (*Lens culinaris*) by using heritability–adjusted biplot analysis. *Crop Pasture Sci.* 69 (11), 1113–1125. doi:10.1071/CP18259

Patterson, H. D., and Williams, E. R. (1976). A new class of resolvable incomplete block designs. *Biometrika* 63, 83–92. doi:10.1093/biomet/63.1.83

Rivera-Amado, C., Molero, G., Trujillo-Negrellos, E., Reynolds, M., and Foulkes, J. (2020). Estimating organ contribution to grain filling and potential for source upregulation in wheat cultivars with a contrasting source–sink balance. *Agronomy* 10 (10), 1527. doi:10.3390/agronomy10101527

Ruuska, S. A., Rebetzke, G. J., van Herwaarden, A. F., Richards, R. A., Fettell, N. A., Tabe, L., et al. (2006). Genotypic variation in water–soluble carbohydrate accumulation in wheat. *Funct. Plant Biol.* 33 (9), 799–809. doi:10.1071/FP06062

Saeidi, M., Moradi, F., and Jalali-Honarmand, S. (2012). The effect of post anthesis source limitation treatments on wheat cultivars under water deficit. *Aust. J. Crop Sci.* 6 (7), 1179–1187.

Saint Pierre, C., Trethowan, R., and Reynolds, M. (2010). Stem solidness and its relationship to water-soluble carbohydrates: Association with wheat yield under water deficit. *Func. Plant Biol.* 37 (2), 166–174. doi:10.1071/FP09174

Salem, K. F., Elabsawy, E. A., Aldahak, L., and Elshamy, H. (2021). Evaluation of stem reserve mobilization in Egyptian bread wheat (*Triticum aestivum* L) genotypes and F1 hybrids under post-anthesis chemical desiccation stress. *J. Genet. Environ. Resour. Conserv.* 9 (1), 176–182.

Santiveri, F., Royo, C., and Romagosa, I. (2002). Patterns of grain filling of spring and winter hexaploidtriticales. *Eur. J. Agron.* 16 (3), 219–230. doi:10.1016/S1161–0301(01)00127–7

Sehgal, A., Sita, K., Siddique, K. H., Kumar, R., Bhogireddy, S., Varshney, R. K., et al. (2018). Drought or/and heat-stress effects on seed filling in food crops: Impacts on functional biochemistry, seed yields, and nutritional quality. *Front. Plant Sci.* 9, 1705. doi:10.3389/fpls.2018.01705

Shearman, V. J., Sylvester–Bradley, R., Scott, R. K., and Foulkes, M. J. (2005). Physiological processes associated with wheat yield progress in the UK. *Crop Sci.* 45 (1), 175–185. doi:10.2135/cropsci2005.0175a

Shirdelmoghanloo, H., Cozzolino, D., Lohraseb, I., and Collins, N. C. (2016). Truncation of grain filling in wheat (*Triticum aestivum*) triggered by brief heat stress during early grain filling: Association with senescence responses and reductions in stem reserves. *Func. Plant Biol.* 43 (10), 919–930. doi:10.1071/FP15384

Talukder, A. S. M. H. M., McDonald, G. K., and Gill, G. S. (2013). Effect of short-term heat stress prior to flowering and early grain set on the grain yield of wheat. *Field Crops Res.* 147, 54–63. doi:10.1016/j.fcr.2014.01.013

Tatar, Ö., Brück, H., and Asch, F. (2016). Photosynthesis and remobilization of dry matter in wheat as affected by progressive drought stress at stem elongation stage. *J. Agron. Crop Sci.* 202 (4), 292–299. doi:10.1111/jac.12160

Thapa, S., Rudd, J. C., Jessup, K. E., Liu, S., Baker, J. A., Devkota, R. N., et al. (2022). Middle portion of the wheat culm remobilizes more carbon reserve to grains under drought. *J. Agron. Crop Sci.* 208 (6), 795–804. doi:10.1111/jac.12508

Vosoghi Rad, M., Jami Moeini, M., Taherian, M., and Armin, M. (2022). Accumulation and remobilization of assimilates in different genotypes of durum wheat under terminal drought stress. *J. Crop Sci. Biotech.* 25 (2), 199–214. doi:10.1007/s12892-021-00123-3

Wang, H., Wang, H., Shao, H., and Tang, X. (2016). Recent advances in utilizing transcription factors to improve plant abiotic stress tolerance by transgenic technology. *Front. Plant Sci.* 7, 67. doi:10.3389/fpls.2016.00067

Yan, W., and Tinker, N. A. (2006). Biplot analysis of multi–environment trial data: Principles and applications. *Can. J. Plant Sci.* 86 (3), 623–645. doi:10.4141/P05–169

Yu, S. M., Lo, S. F., and Ho, T. H. D. (2015). Source–sink communication: Regulated by hormone, nutrient, and stress cross-signaling. *Trends Plant Sci.* 20 (12), 844–857. doi:10.1016/j.tplants.2015.10.009

Zaveri, E., and Lobell, B. D. (2019). The role of irrigation in changing wheat yields and heat sensitivity in India. *Nat. Comm.* 10 (1), 4144. doi:10.1038/s41467-019-12183-9

Zhang, M., Gao, Y., Zhang, Y., Fischer, T., Zhao, Z., Zhou, X., et al. (2020). The contribution of spike photosynthesis to wheat yield needs to be considered in process-based crop models. *Field Crops Res.* 257, 107931. doi:10.1016/j.fcr.2020.107931

Zhu, L., Liang, Z. S., Xu, X., Li, S. H., and Monneveux, P. (2009). Evidences for the association between carbon isotope discrimination and grain yield–Ash content and stem carbohydrate in spring wheat grown in Ningxia (Northwest China). *Plant Sci.* 176 (6), 758–767. doi:10.1016/j.plantsci.2009.02.018

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 21 July 2023
DOI 10.3389/fgene.2023.1198600



OPEN ACCESS

EDITED BY
Sundeep Kumar,
Indian Council of Agricultural Research (ICAR), India

REVIEWED BY
Balpreet Kaur Dhatt,
University of Nebraska-Lincoln, United States
Bahram Heidari,
Shiraz University, Iran

*CORRESPONDENCE
Wang Xing,
xiaoyujiayou_086@163.com

†These authors share first authorship

RECEIVED 01 April 2023
ACCEPTED 30 June 2023
PUBLISHED 21 July 2023

CITATION
Li W, Lin M, Li J, Liu D, Tan W, Yin X, Zhai Y, Zhou Y and Xing W (2023), Genome-wide association study of drought tolerance traits in sugar beet germplasms at the seedling stage.
*Front. Genet.* 14:1198600.
doi: 10.3389/fgene.2023.1198600



# Genome-wide association study of drought tolerance traits in sugar beet germplasms at the seedling stage

Wangsheng Li[1,2†], Ming Lin[3†], Jiajia Li[1,2], Dali Liu[1,2], Wenbo Tan[1,2], Xilong Yin[1,2], Yan Zhai[1,2], Yuanhang Zhou[3] and Wang Xing[1,2*]

[1]National Beet Medium-Term Gene Bank, Heilongjiang University, Harbin, China, [2]Key Laboratory of Sugar Beet Genetics and Breeding, College of Advanced Agriculture and Ecological Environment, Heilongjiang University, Harbin, China, [3]Xinjiang Academy of Agricultural Sciences, Urumqi, China

**Introduction:** Sugar beets are an important crop for global sugar production. Intense drought and the increasing lack of water resources pose a great threat to sugar beet cultivation. It is a priority to investigate favourable germplasms and functional genes to improve the breeding of drought tolerant plants.

**Methods:** Thus, in this study, 328 sugar beet germplasms were used in a genome-wide association study (GWAS) to identify single nucleotide polymorphism (SNP) markers and candidate genes associated with drought tolerance.

**Results:** The results showed that under drought stress (9% PEG-6000), there were 11 significantly associated loci on chromosomes 2, 3, 5, 7, and 9 from the 108946 SNPs filtered using a mixed linear model (MLM). Genome-wide association analysis combined with qRT-PCR identified 13 genes that were significantly differentially expressed in drought-tolerant extreme materials.

**Discussion:** These candidate genes mainly exhibited functions such as regulating sugar metabolism, maintaining internal environmental stability and participating in photosystem repair. This study provides valuable information for exploring the molecular mechanisms of drought tolerance and improvement in sugar beet.

KEYWORDS

**sugar beet germplasms, PEG-6000 drought stress, SNP, GWAS, linkage disequilibrium**

## 1 Introduction

Sugar beet is an important sugar crop. It has strong drought tolerance in the middle and late growth period, but weak drought tolerance at seedling stage. Most regions in the world are facing the problem of drought, and the sugar beet plant area of Heilongjiang Province in China is no exception. In spring, there are many southerly winds, large evaporation and small precipitation. The development of drought-tolerant varieties becomes more and more important (Bloch et al., 2006), and the selection and utilization of excellent drought tolerant germplasm resources and drought tolerant genes play an important role in the utilization efficiency and the improvement of agricultural production in water-shortage areas.

Mapping of quantitative trait loci (QTLs) is an effective tool often used to reveal the genetic basis of complex quantitative traits in crops (Huang et al., 2006). However, using traditional molecular markers such as restriction fragment length polymorphism (RFLP)

and simple sequence repeats (SSR), only a few QTLs associated with drought stress can be identified (Jha et al., 2021; Dwiningsih et al., 2022). With the development of functional genomics and transcriptomics in recent years, a large number of genes have been found to be involved in plant drought tolerance, including protein kinases, transcription factors and some structural genes (Wang et al., 2021b). However, only a few genes have been functional validated. It remains a major challenge to accurately and effectively screen more drought stress-related genes and use them to breed drought-tolerant varieties.

Drought tolerance in plants is a complex quantitative trait that is controlled by multiple genes and involves multiple physiological and biochemical metabolic pathways. Although QTL mapping is a powerful method for detecting genomic regions associated with complex traits, the genetic effects of QTLs may not exist or may simply not be tested in different genetic backgrounds and environments (Shariatipour et al., 2021). Compared with traditional QTL, genome-wide association study (GWAS) can utilize genome-wide single nucleotide polymorphisms (SNPs) as molecular markers to analyze the genetic basis of complex traits, assisted by high-throughput genotyping platforms (Zheng et al., 2021). With the continuous development of high-throughput sequencing and high-resolution metabolic assays, as well as the development of multiple bioinformatics techniques and statistical methods to provide a basis for precise localization of complex trait gene variants, genome-wide association analysis is playing an increasingly important role in mining genetic loci for research (Zheng et al., 2021). Significant markers associated with aboveground biomass, fruit size, lodging score and leaf elongation in beans were identified through a whole genome-wide association analysis of 96 common bean germplasms. Specifically, 7 significant markers were found under irrigated conditions, and 5 significant markers were found under water deficit conditions (Hoyos-Villegas et al., 2017). In 2019, Mathew et al. used genome-wide association analysis to identify 75 significant marker-trait associations for population structure and marker traits of wheat biomass traits under drought and non-stress conditions. A total of 37 presumptive candidate genes were screened through gene annotation on IWGSC RefSeq 1.1 (Mathew et al., 2019). In recent years, the genome-wide association method has been applied in many crop drought tolerance studies and achieved good results. However, the current research on drought tolerance of sugar beet mainly focus on physiological and biochemical responses. Although traditional breeding methods such as analysis of combinin ability of lines have been used for selection of superior sugar beet genotypes for sugar related traits (Hassani et al., 2020), GWAS provide more reliable information about markers linked with sugar trait for use in marker-assisted selection We believe that using genome-wide association analysis in drought tolerance studies of beet germplasm resources can quickly and accurately excavate drought tolerance related genes. Therefore, in this study, 328 beet germplasm resources were simulated by PEG-6000 drought stress. It is hoped to explore significant loci and potential candidate genes related to drought tolerance of beet germplasm resources by genome-wide association analysis, and to provide reference and basis for further research on the isolation of related genes and molecular marker-assisted selection of beet drought tolerance.

# 2 Materials and methods

## 2.1 Test material

The material for this study was provided by the National Beet Medium-Term Gene Bank at Heilongjiang University. The 328 sugar beet germplasm resources used in the trial were from 17 countries (Supplementary Table S1). 200 from China, 9 from Japan, 1 from North Korea, 15 from Russia, 49 from the United States, 1 from Austria, 2 from Belgium, 10 from Poland, 1 from Denmark, 10 from Germany, 3 from France, 12 from the Netherlands, 2 from Romania, 5 from Sweden, 2 from Hungary, 4 from Italy, and 2 from the United Kingdom (Figure 1). This experiment was conducted in a sophisticated climate chamber with a constant environment of 25°C and 60% humidity during the day and 18°C and 60% humidity during the night. There were 4 rows for each genotype, and 6 plants in each row were cultured in Hoagland's solution. After growing one pair of true leaves, two of the four rows of germplasms were selected and treated with PEG-6000 (Livak and Schmittgen) at a concentration of 9% to simulate drought stress (Tan et al., 2023). The remaining two rows of germplasms continued to be cultured in Hoagland's solution as controls (CG). The solution was changed once every 3 days. The materials were sampled after the growth of three pairs of true leaves (Cai et al., 2015).

## 2.2 Phenotype identification and physiological trait assay

The following traits were measured in this test: embryonic axis diameter (stem diameter measured using Vernier callipers) (Yang et al., 2022b), maximum root length (a straightedge was used to measure the maximum length of the root), plant height (measured from the ground to the top of the plant using a straightedge), root fresh weight (the roots were removed, dried on paper towels, and weighed on an electronic balance), leaf fresh weight (all leaves were cut off and weighed on an electronic balance), root dry weight (roots were dried in a ventilated oven at 105°C for 15 min, then transferred to 75°C for drying to a standard weight) (Polash et al., 2018), and leaf dry weight (leaves were dried in a ventilated oven at 105°C for 15 min, then transferred to 75°C for drying to a standard weight) (Yang et al., 2022a). Related tools were Vernier callipers (0.01 mm, Brand: Sheffield, Model: S071012, Origin: Mainland China), electronic balance (0.00001 g, Brand: Mettler-Toledo GmbH, Model: MS105DU/A, Origin: Switzerland), and steel ruler (0.1 cm, Brand: Sheffield, Model: S079020, Origin: China Mainland). The relative leaf water was calculated as follows:

Relative leaf water = (Leaf fresh weight-Leaf dry weight)/(Leaf saturated fresh weight-Leaf dry weight).

Superoxide dismutase activity was evaluated by the nitrogen blue tetrazolium method (Spitz and Oberley, 1989). Soluble sugar levels were evaluated by the Anthrone-colorimetric method (Maness, 2010). Proline levels were evaluated by the acidic ninhydrin method (Carillo and Gibon, 2011). Soluble protein levels were evaluated using a BCA Protein Assay Kit (Brand: biosharp, Model: BL521A-3, Origin: Beijing, China).

The drought tolerance coefficient (DTC) was calculated for each trait:

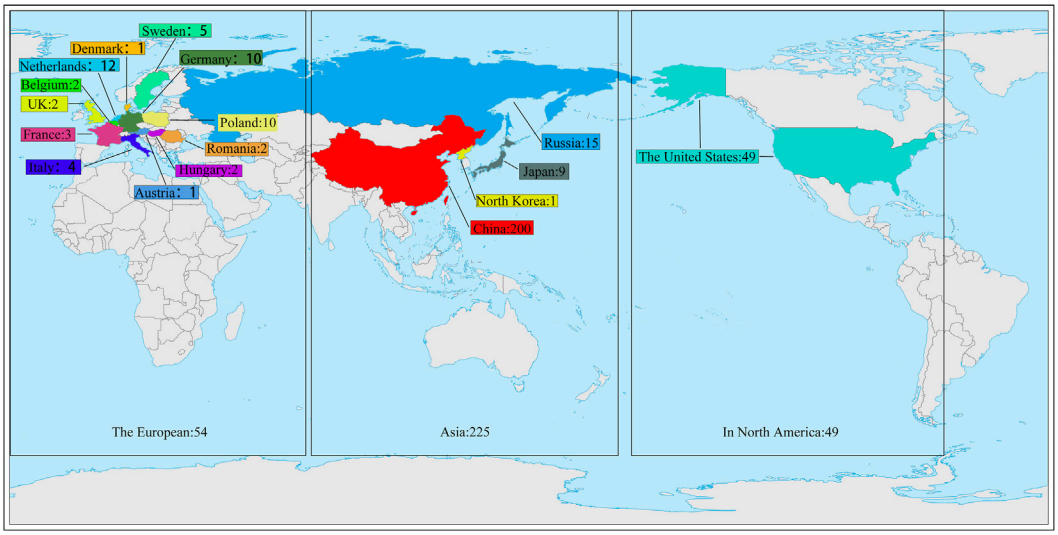



FIGURE 1
328 sugar beet germplasm resources distribution map.

$$DTC_X = DT/NT$$

where DTC_X is the drought tolerance coefficient of each trait; DT is the mean value of the traits under drought treatment; and NT is the mean value of the traits under normal treatment.

## 2.3 Genotype data

Sequencing: SNPs were identified by NOVOGENE (Beijing, China) and sequenced using the Illumina HiSeq sequencing platform with double-end (Paired-End) 150 sequencing. BWA (Li and Durbin, 2009) (0.7.17): Interference information was filtered out of the raw data to obtain high-quality clean data, and effective high-quality sequencing data were aligned to the reference genome by BWA software (parameter: mem-t 4-k 32-M) (reference genome download address: ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/511/025/GCF_000511025.2_RefBeet-1.2.2/GCF_000511025.2_RefBeet-1.2.2_genomic.fna.gz). Samtools (Soskic et al., 2019) (1.9): Samtools was used to transform the format of the sam file and build an index to generate the bai file. GATK (Sun et al., 2019) (4.2.6.1): GATK was used to detect mutations and generate group modification.vcf file. VCFtools (Ratnapriya et al., 2019) (0.1.17): A total of 108946 high-quality loci were screened for subsequent analysis using sequencing depth (dp ≥ 4), minor allele frequency (maf ≥ 0.05) and miss rate (miss ≤ 80%).

## 2.4 Population structure and LD analysis

VCFtools (0.1.17) was first used to convert the.vcf files ped files and then plink (1.9) software was used to convert ped format to bed format (Cadzow et al., 2014). Population structure analysis of bed files was performed using admixture (1.3.0) software, the population size K value was preset to 1–10 for classification, and the minimum value of cross entropy was used to determine the optimal number of classifications (Byrne et al., 2020). Principal component analysis of 108946 high-quality SNP loci was performed using Tassel (5.2.82) software (Breria et al., 2020), and the PCA results were visualized using R (4.1.0) software. The LD of 328 sugar beet germplasms was analysed by PopLDdecay (3.4.2). The LD intensity (r2) between two SNPs within a certain distance was calculated (Xu et al., 2018).

## 2.5 Genome-wide association study

The kinship analysis was performed using Tassel (5.2.82) software to obtain the kinship matrix (Zeng et al., 2017) and the results were visualized using the pheatmap package in R (4.1.0). The mixed linear model (MLM) (Kang et al., 2015) of Tassel software was used to analyse the PCA and kinship analysis results as covariates. This combination (Q + K) is often considered the most powerful, and this model can classify crops into different subgroups based on growth habit or geographical origin. Mixed models corrected with population structure and kinship can effectively reduce false positive results (Wang et al., 2007). Manhattan and Q-Q plots were visualized using the CMplot package of R (4.0.1) (Kumar et al., 2021). The mixed linear model (MLM) equation was as follows:

$$Y = X\alpha + Q\beta + K\mu + e$$

In the equation, Y is the vector of measured phenotypes, X is the fixed effect of SNPs, P is the fixed effect of population structure, K is the random effect of kinship, and e is the random error (Dramadri et al., 2021).

We used the powerful annotation function of the internationally recognized UniProt database (Yang and Wang, 2015) to annotate the SNPs. After screening fake SNPs according to genome

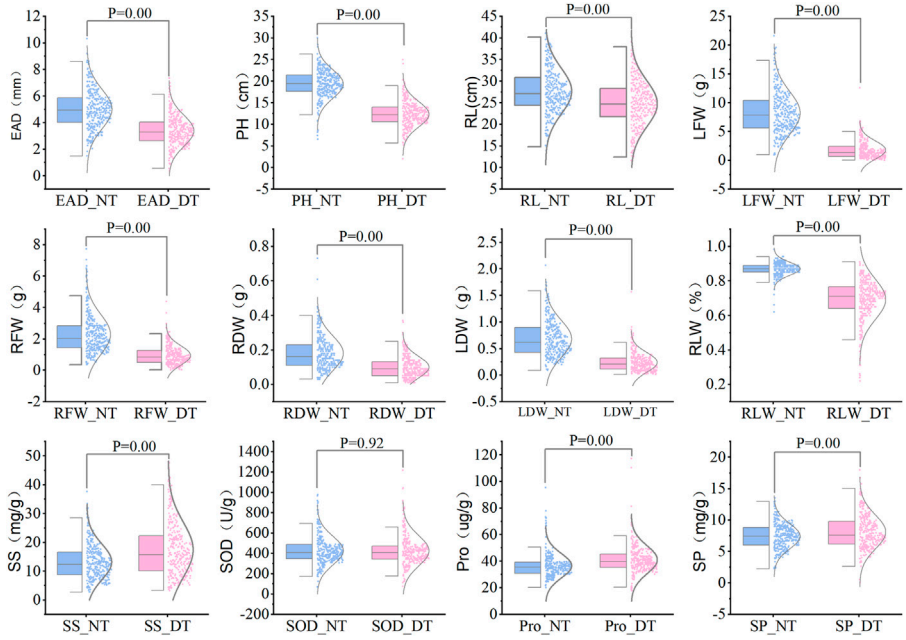



FIGURE 2
Comparison of various traits between drought treatment and normal water supply for 328 sugar beet germplasms. Note: NT, normal treatment; DT, drought treatment, 9% PEG-6000 for 7 days.

**TABLE 1 Analysis of variation in drought tolerance coefficient of 328 sugar beet germplasm for each trait.**

| Traits | Mean | STDEV | Max | Min | Skewness | Kurtosis | CV(%) |
|---|---|---|---|---|---|---|---|
| EAD | 0.708 | 0.243 | 1.676 | 0.171 | 0.718 | 0.848 | 34.38 |
| PH | 0.651 | 0.154 | 1.197 | 0.259 | 0.506 | 0.586 | 23.58 |
| RL | 0.914 | 0.215 | 1.659 | 0.475 | 0.549 | 0.238 | 23.55 |
| LFW | 0.236 | 0.188 | 1.042 | 0.005 | 1.411 | 2.112 | 79.58 |
| RFW | 0.482 | 0.322 | 1.879 | 0.015 | 1.269 | 1.87 | 66.73 |
| RDW | 0.634 | 0.406 | 2.333 | 0.022 | 1.261 | 1.923 | 64.16 |
| LDW | 0.394 | 0.278 | 1.684 | 0.019 | 1.341 | 2.74 | 70.62 |
| RLW | 0.799 | 0.130 | 1.226 | 0.267 | −1.062 | 2.581 | 16.29 |
| SS | 1.611 | 1.225 | 6.943 | 0.152 | 1.725 | 3.249 | 76.03 |
| SOD | 1.023 | 0.263 | 1.994 | 0.204 | 0.532 | 1.409 | 25.72 |
| Pro | 1.156 | 0.319 | 2.185 | 0.362 | 0.539 | 0.446 | 27.57 |
| SP | 1.130 | 0.415 | 2.715 | 0.006 | 0.831 | 0.928 | 36.70 |

Note: EAD, embryonic axis diameter; PH, plant height; RL, maximum root length; LFW, leaf fresh weight; RFW, root fresh weight; RDW, root dry weight; LDW, leaf dry weight; RLW, relative leaf water; SS, soluble sugar levels; SOD, superoxide dismutase activity; Pro, Proline levels; SP, soluble protein levels.

annotation, we obtained high-quality and significant SNPs with a threshold of –log10p> 6.

The SNPs were screened to obtain high-quality and significantly correlated SNPs with a threshold of –log10p> 6. The size of the screened SNP interval was estimated based on LD block analysis, and genes in the region near the locus were obtained and functionally annotated. Based on the functional annotations, we could roughly determine which genes might be associated with the traits of interest and use in subsequent analysis (Bai et al., 2022).

## 2.6 Real-time fluorescence quantitative PCR

Four drought-tolerant (Mono HY53, Neitang578, Gansunongjiazhong, BGRC16137) and four drought-sensitive

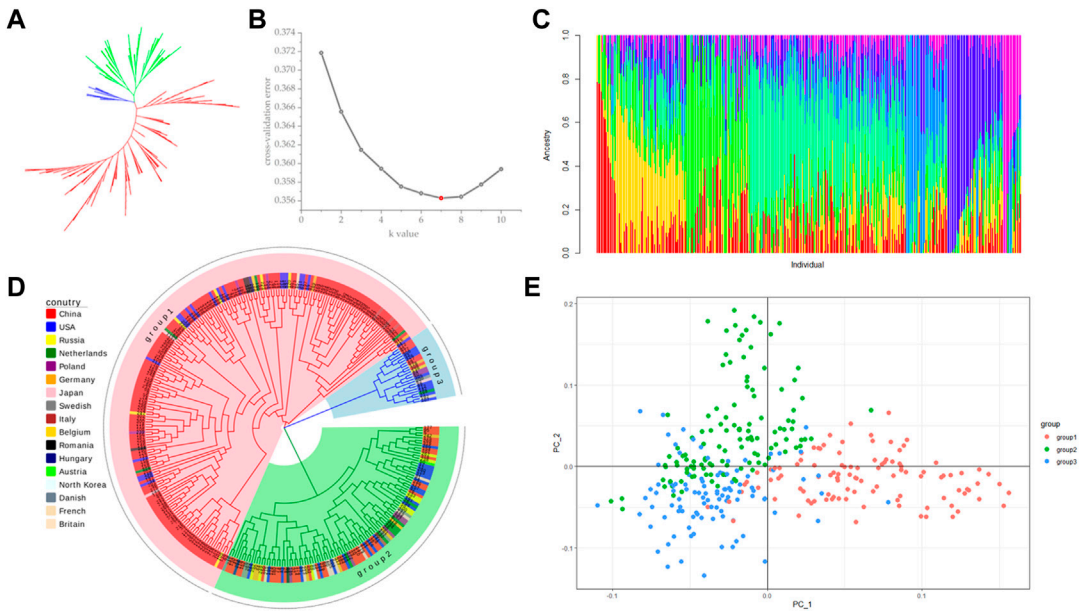



FIGURE 3
Population structure analysis. **(A)** NJ method unrooted evolutionary tree; **(B)** cross-validation (CV) error value; **(C)** estimation of population structure; **(D)** NJ method rooted evolutionary tree; **(E)** principal component analysis (PCA).

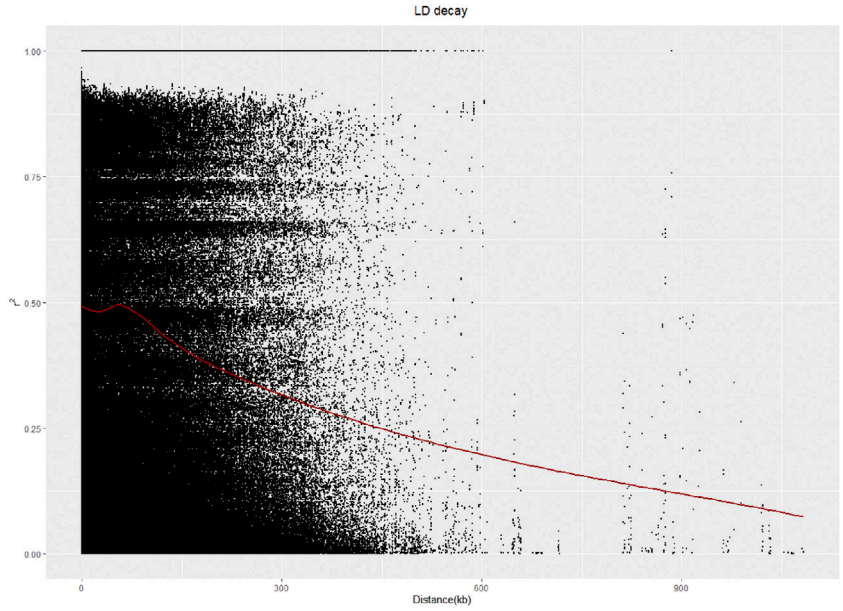



FIGURE 4
LD analysis of 328 sugar beet germplasms.

germplasms [7501A/BCE, 92017/1-8/1, Tianyansanhao (X), 92017/1-4/1] were used as extreme materials to extract total RNA according to the Trizol method. cDNA was synthesized according to the TranScript One-Step gDNA Removal and cDNA Synthesis SuperMix protocol (AT311, Beijing All Style Gold Biotechnology Co., Ltd.). qPCR was performed using the SuperReal PreMix Plus kit (Beijing Tiangen Biochemical Technology Co., Ltd., version FP210831) on a real-time fluorescent quantitative PCR instrument (Thermo Fisher Scientific Instruments, Shanghai. QuantStudio™ 1 Plus) (Liu et al., 2023). Expression level of 14 genes was analyzed with the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001). *BvGAPDH* (NC_024800) was used as an internal control to standardize the expression levels of the different samples. All assays were performed in two independent experiments and replicated three times. Primers were summarized in Supplementary Table S5.

10.3389/fgene.2023.1198600

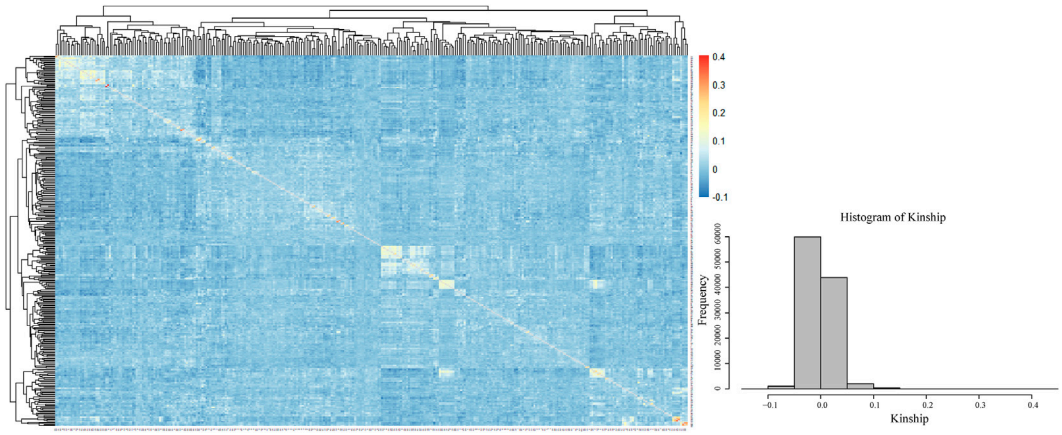



FIGURE 5
Kinship heatmap and cluster analysis.

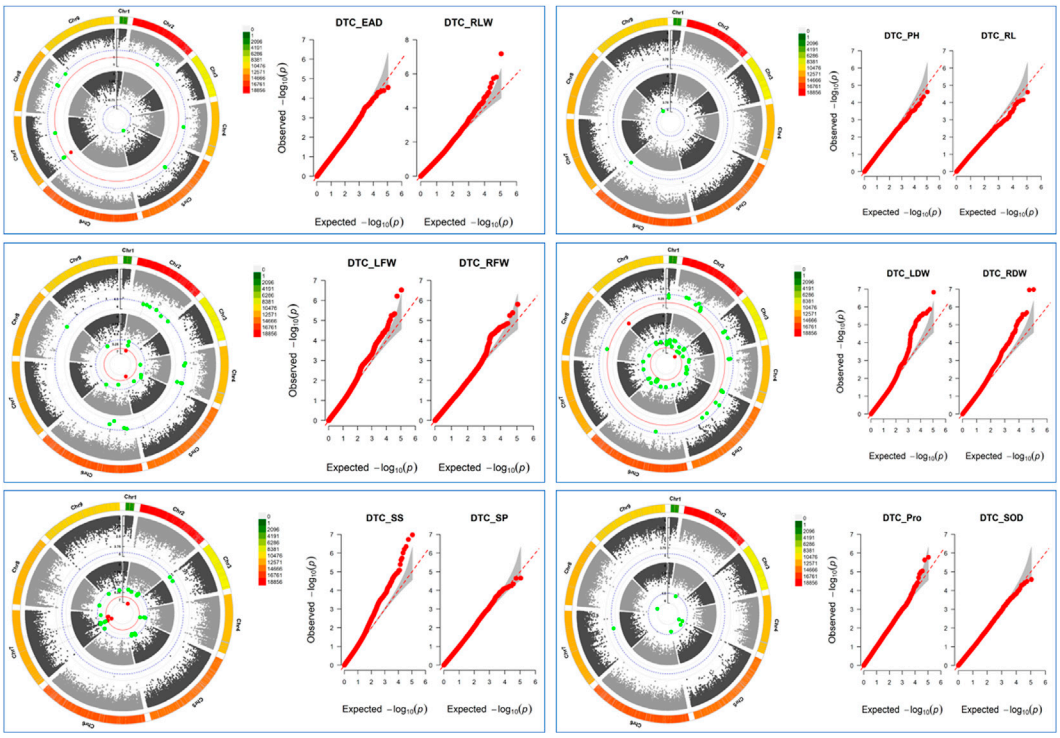



FIGURE 6
Manhattan plot: inner circles indicate single nucleotide polymorphisms (SNPs) associated with the Q-Q plot on the left, outer circles indicate single nucleotide polymorphisms (SNPs) associated with the Q-Q plot on the right, grey dashed lines indicate SNP loci with –log10p = 4.5, red dashed lines indicate SNP loci with –log10p = 6, and the outermost side indicates SNP density on the chromosome. Q-Q plots: horizontal coordinates indicate expected values, vertical coordinates indicate observed values, and grey areas are 95% confidence intervals.

# 3 Results

## 3.1 Phenotypic and physiological analysis

The statistical results showed that the measured data were all close to a normal distribution, indicating that the 12 traits were all typical quantitative traits. Eight traits, including embryonic axis diameter (EAD), plant height (PH), root length (RL), leaf fresh weight (LFW), root fresh weight (RFW), root dry weight (RDW), leaf dry weight (Lander et al., 2001) and leaf relative water content (RLW), were significantly lower in the drought-treated plants than in the control plants. Three indicators, including soluble sugar levels

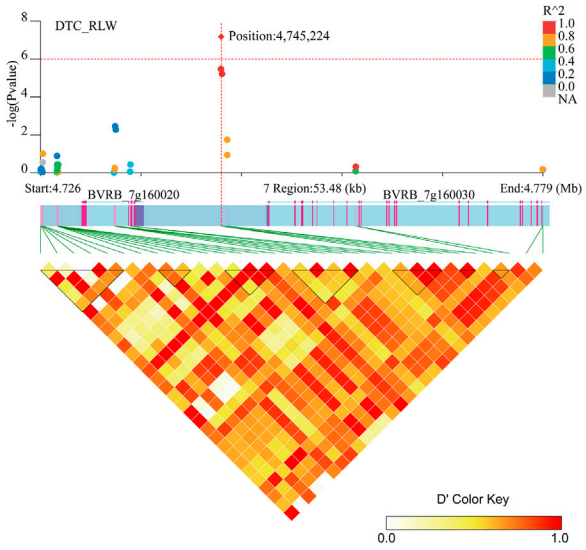



FIGURE 7
Manhattan plot of candidate gene regions of DTC_RLW and heatmap of LD (bottom). The orange vertical line indicates the position of the most significant SNPs, and the orange horizontal line indicates –log10p = 6.

(SS), soluble protein levels (SP) and free proline levels (Pro), were significantly higher in the drought-treated plants than in the control plants (Figure 2). Drought tolerance coefficients for each trait showed that these traits all exhibited suitable variation, indicating that they were controlled by small effect polygenes with coefficients of variation (CV) ranging from 16.29% to 79.58% (Table 1). The CV is an absolute value reflecting the degree of dispersion of each trait in the population. To better estimate the correlations between markers and traits, it is desirable to use highly dispersed trait data for statistical analysis. The traits in this test exhibited a high degree of dispersion and were suitable for genome-wide association analysis.

## 3.2 Population structure and genetic diversity

Population stratification refers to the existence of subgroups within a population, where the diversity among individuals within the subgroups is greater than the average diversity among individuals within the entire population (Wang et al., 2021a). Different allele frequencies at certain loci between different subpopulations can lead to false positive results when using two subpopulations for association analysis (Zheng et al., 2021). Therefore, before performing association analysis, we used a subset of 108946 SNPs without close linkage for population structure prediction. Using Tassel (5.2.82) software to construct phylogenetic trees by the neighbor-joining (NJ) method, 328 sugar beet germplasms were divided into three subgroups, but the differences between the subgroups were small (Figures 3A, D). The results from a principal component analysis (PCA) of the above three subgroups were consistent with the phylogenetic analysis, and no significant grouping was found (Figure 3E). The results indicated that the 328 germplasms without significant grouping were suitable for a GWAS. Using admixture software, the number of subgroups (K) was set to 1–10, and each K value was repeated three times to select the optimal K value according to the maximum likelihood method. The cross-validation error (CV error) value was the smallest when K = 7 (Figure 3B), so the Q-matrix with K = 7 was used as a covariate to improve the accuracy of the GWAS (Figure 3C).

A linkage disequilibrium (LD) analysis was carried out, and the LD distance decreased as the physical location of the SNPs on chromosomes gradually increased. The LD distance of 50 kb was equal to half of the maximum value (Figure 4).

The visualization of the kinship matrix showed that the variation between most of the 328 genotypes was low (light blue), and the genetic similarity between individual genotypes was high. The low genetic variation in this study would reduce the occurrence of false positive results (Figure 5).

## 3.3 GWAS of drought tolerance coefficients for each trait

The GWAS results showed (Figure 6; Supplementary Table S2) that a total of 183 SNPs (–log10p > 4.5) were identified in the 12 drought tolerance-related traits, including 11 SNPs with peaks >6.0, which the study mainly focused on. It was found that 1 SNP on chromosome 7 was associated with DTC_RLW; 2 SNPs associated with DTC_LFW were located on chromosomes 2 and 5; 1 SNP on chromosome 3 was associated with DTC_LDW; 2 SNPs on chromosome 9 were associated with DTC_RDW; and there were 5 SNPs associated with DTC_SS on chromosomes 2 and 7. It is worth noting that the two significant loci associated with DTC_RDW were close to each other with a distance of only 216 bp.

## 3.4 Candidate genes associated with sugar beet drought tolerance

In the five traits (–log10p> 6) of DTC_RLW, DTC_LFW, DTC_LDW, DTC_RDW, and DTC_SS, 50 kb upstream and downstream of the significant loci were analyzed by LD decay. A total of 24 genes in sugar beet associated with drought tolerance were identified and annotated (Supplementary Table S3).

Two genes (*BVRB_7g160020* and *BVRB_7g160030*) were identified on chromosome 7 at position 4,745,224 may be associated with DTC_RLW (Figure 7; Supplementary Table S3). *BVRB_7g160020* encodes a tyrosine-protein phosphatase, RLPH2, which is reported to be a new phospho-tyrosine-specific phosphatase belonging to the phosphoprotein phosphatase (PPP) family (Uhrig et al., 2016). *BVRB_7g160030*, as 1,4-alpha-glucan-branching enzyme 2-2, is mainly involved in amylose synthesis in gluconeogenesis (Han et al., 2022). It was reported that *BVRB_7g160030* catalytically regulates α (1–6) glycosidic bond branching synthesis. We believe that *BVRB_7g160030* is a potential target gene for the drought stress response in sugar beet.

Associated with DTC_LFW, three genes (*BVRB_2g033710*, *BVRB_2g033720* and *BVRB_2g033730*) were found on chromosome 2. Two genes (*BVRB_5g114800* and *BVRB_*

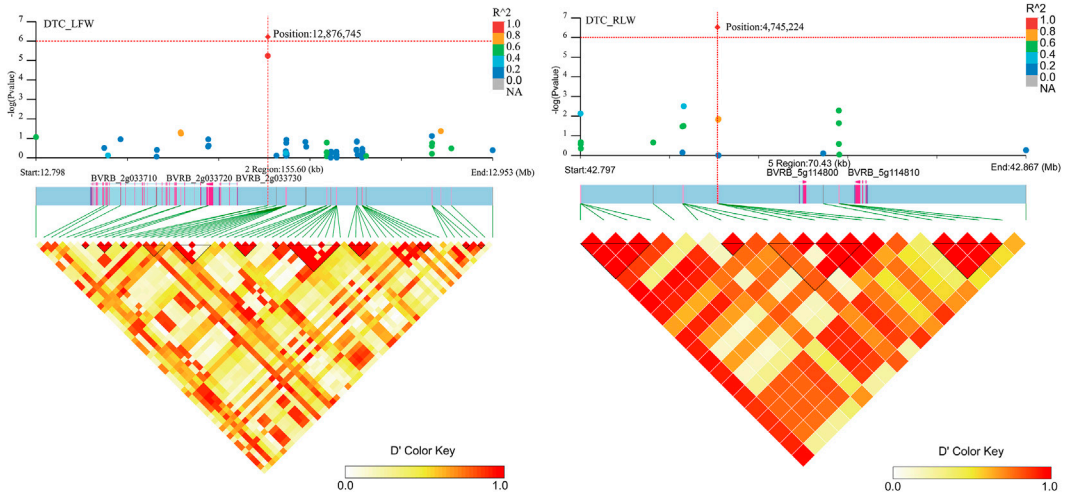



FIGURE 8
Manhattan plot of candidate gene regions associated with DTC_RLW and heatmap of linkage disequilibrium (bottom). The orange vertical line indicates the position of the most significant SNP, and the orange horizontal line indicates −log10p = 6.

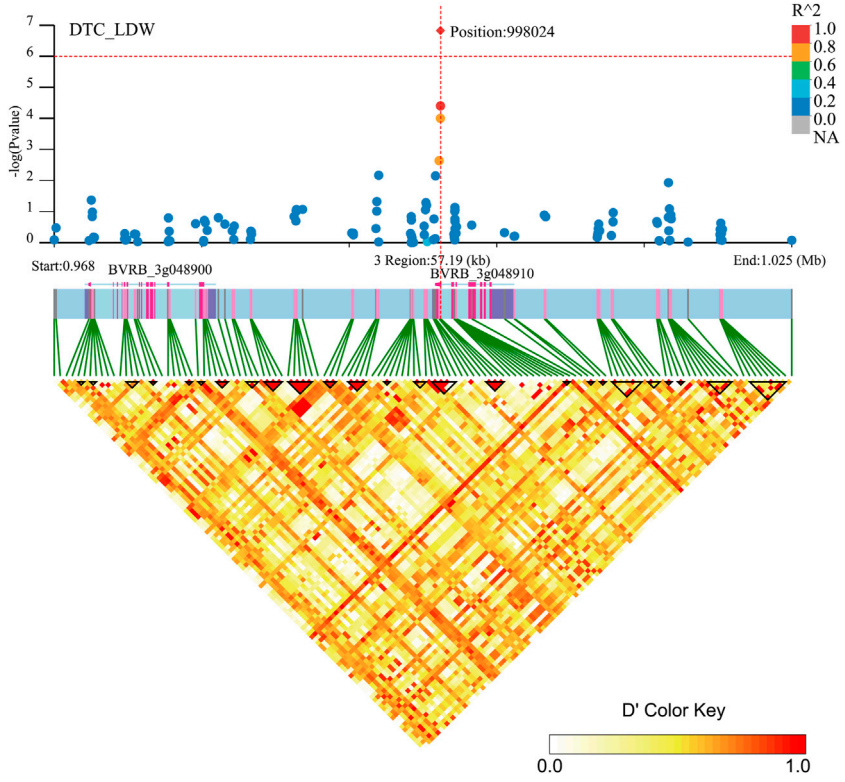



FIGURE 9
Manhattan plot of candidate gene regions associated with DTC_LDW and heatmap of linkage disequilibrium (bottom). The orange vertical line indicates the position of the most significant SNP, and the orange horizontal line indicates −log10p = 6.

*5g114810*) were found on chromosome 5. *BVRB_2g033720*, *BVRB_2g033730* and *BVRB_5g114800* were of unknown function (Figure 8; Supplementary Table S3). Through annotation, we found that *BVRB_2g033710* functions as myosin-15; *BVRB_5g114810* functions as MYB-related protein 306, the MYB structural domain is a peptide segment of approximately 51-52 amino acids containing a series of highly conserved amino acid residues and spacer sequences, and the MYB gene family is

10.3389/fgene.2023.1198600

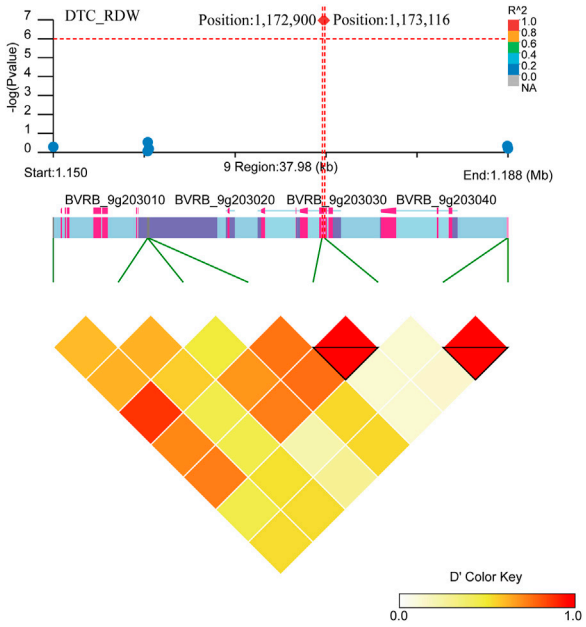



FIGURE 10
Manhattan plot of candidate gene regions associated with DTC_RDW and heatmap of linkage disequilibrium (bottom). The orange vertical line indicates the position of the most significant SNP, and the orange horizontal line indicates −log10p = 6.

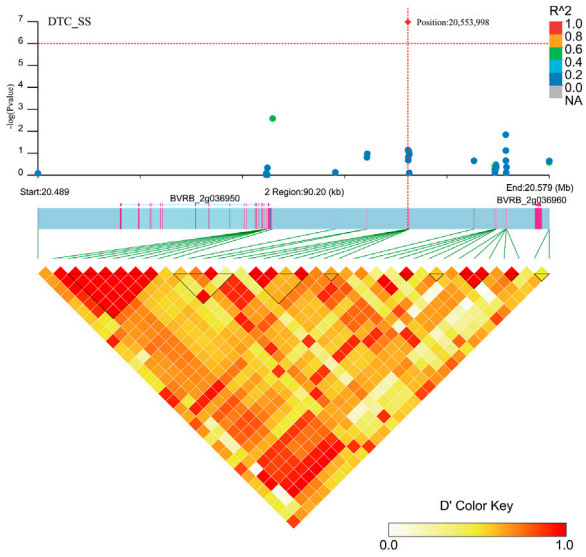



FIGURE 11
Manhattan plot of candidate gene regions associated with DTC_SS and heatmap of linkage disequilibrium (bottom). The orange vertical line indicates the position of the most significant SNP, and the orange horizontal line indicates −log10p = 6.

widely involved in plant metabolic regulation (Jiao et al., 2021; Johnson et al., 2022).

Two genes (*BVRB_3g048910* and *BVRB_3g048900*) were identified on chromosome 3 position: 998,024, may be associated with DTC_LDW (Figure 9; Supplementary Table S3). By annotation, we found that *BVRB_3g048910* functions as the probable methyltransferase PMT3. Methyltransferases use S-adenosylmethionine, betaine and dimethylthetin as methyl donors to generate methionine. *BVRB_3g048900* functions as serine/threonine-protein kinase tricornered (TOR), which functions as a phosphate donor using ATP to phosphorylate serine and threonine residues on proteins to regulate abscisic acid, growth hormone, glucose and sucrose-mediated signalling (Ahn et al., 2011; Xiong et al., 2013; Deng et al., 2016).

Two close significant loci (position: 1,172,900, position: 1,173,116) on chromosome 9 were found to be associated with DTC_RDW (Figure 10; Supplementary Table S3). Six gene were found in the region near these two significant loci (*BVRB_9g203050*, *BVRB_9g203040*, *BVRB_9g203030*, *BVRB_9g203020*, *BVRB_9g203010* and *BVRB_9g20300*). *BVRB_9g203050* functions as photosystem II (PSII) stability/assembly factor HCF136 and is essential for PSII biogenesis. BVRB_9g203030 functions as eukaryotic initiation factor 4A-9, which promotes the hydrolysis of ATP and drives RNA deconvolution against abiotic stress; *BVRB_9g203000* functions as enhancer of AG-4 protein 2, which possesses the dual ability to induce apoptosis and autophagy (Kumar et al., 2015), and is a transcription factor that acts as a flowering repressor by enhancing the expression of genes that delay flowering and suppressing nutritional growth to inhibit reproductive development (Wang et al., 2007).

Two gene regions (*BVRB_2g036960* and *BVRB_2g036950*) were found to be associated with DTC_SS (Figure 11; Supplementary Table S3) in the region near positions 20,553,998 on chromosome 2. *BVRB_2g036950* is an uncharacterized protein, and *BVRB_2g036960* functions as a probable membrane-associated kinase regulator 4. Tyrosine phosphorylation controls the activation of brassinosteroid receptors through membrane-released protein kinase inhibitors.

Four SNP loci significantly associated with DTC_SS (Figure 12; Supplementary Table S3) were found on chromosome 7. No genes were found near position 15,215,427. Two genes (*BVRB_7g163640* and *BVRB_7g163650*) were found near position 15,170,171 and they are hypothetical and uncharacterized proteins. Three genes (*BVRB_7g166470*, *BVRB_7g166480* and *BVRB_7g166490*) were found near position 27,948,364, and *BVRB_7g166470* functions as glycine-rich cell wall structural protein 1, a process that results in the assembly, arrangement of constituent parts, or disassembly of the cell wall. *BVRB_7g166480* functions as PROTEIN LOW PSII ACCUMULATION 1, which is mainly involved in the efficient assembly of photosynthetic system II. BVRB_7g166490 functions as photosystem I reaction centre subunit IV. Two gene (*BVRB_7g166670* and *BVRB_7g166680*) were found near positions 28,950 and *BVRB_7g166670* is an uncharacterized protein, and *BVRB_7g166680* functions as a ras GTPase-activating protein-binding protein.

## 3.5 qRT-PCR validation of candidate gene expression

By examining the overlap of the identified candidate genes with known drought stress response genes in other crop species, we

10.3389/fgene.2023.1198600

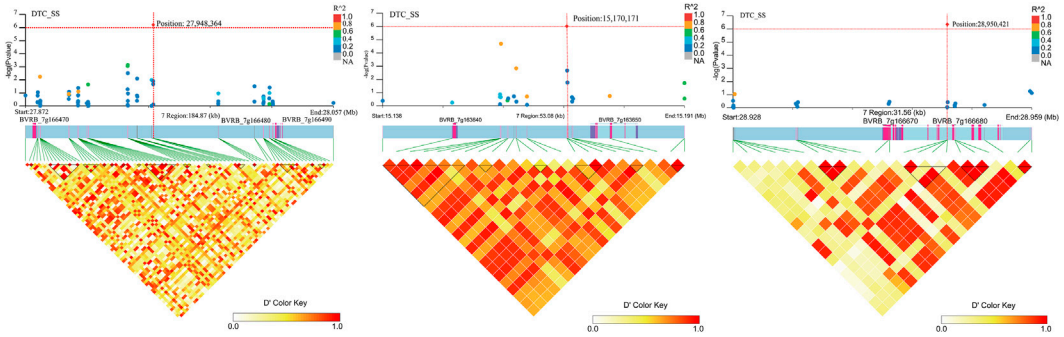



FIGURE 12
Manhattan plot of candidate gene regions associated with DTC_SS and heatmap of linkage disequilibrium (bottom). The orange vertical line indicates the position of the most significant SNP, and the orange horizontal line indicates −log10p = 6.

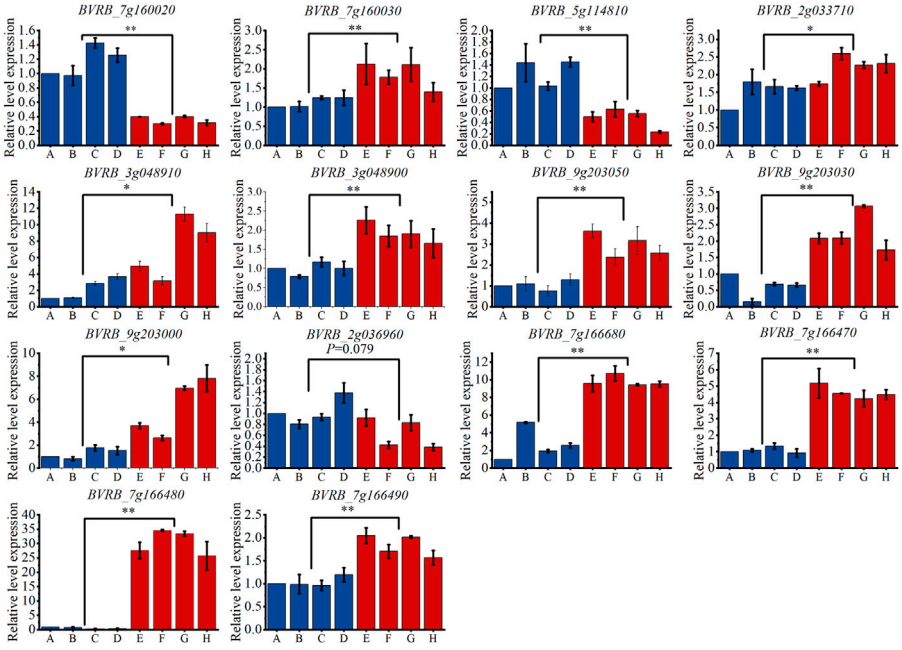



FIGURE 13
Expression levels of candidate genes in drought-tolerant extreme materials. Note: **(A–D)** are drought-sensitive materials. **(E–H)** are drought-tolerant materials. *: Indicates that the expression levels of the two groups of materials are different at the 0.05 level. **: Indicates that the expression levels of the two groups of materials are different at the 0.01 level.

identified 14 genes that may be involved in the drought tolerance process in sugar beet. In order to validate the GWAS results, qRT-PCR was performed on these genes (As shown in Figure 13). The results showed that, except for *BVRB_2g036960*, the other 13 genes were all significantly differently expressed between the drought tolerant and sensitive groups. The expression levels of *BVRB_7g160020* and *BVRB_5g114810* were 3.30-fold and 2.57-fold higher in the drought-sensitive group than them in the drought-tolerant group, respectively. *BVRB_7g160030*, *BVRB_2g033710*, *BVRB_3g048910*, *BVRB_3g048900*, *BVRB_9g203050*, *BVRB_9g203030*, *BVRB_9g203000*, *BVRB_7g166680*, *BVRB_7g166470* and *BVRB_7g166490* in drought-tolerant group were expressed 1.47-4.23-fold of them in the drought-sensitive group. Notably, the expression level of *BVRB_7g166480* in the drought-tolerant group was 46.69-fold higher than that of the drought-sensitive group.

# 4 Discussion

## 4.1 Drought tolerance-related traits in sugar beet

Drought stress is the main abiotic stress factor that causes crop yield reduction worldwide. The sugar beet plant is more drought tolerant in the middle and late growth stages but is drought sensitive

in the seedling stage, which causes great taproot loss (Fugate et al., 2018). Enhancing the drought tolerance of seedlings will greatly help to improve the sugar beet crop yield. Sadeghian and Yavari conducted mannitol-simulated drought stress on 9 sugar beet varieties, assessed the dry weight of cotyledons, fresh weight of cotyledons, RFW, RDW, and RL, and concluded that the drought tolerance traits of sugar beet seedlings can be used for selective breeding of the strain with the strongest tolerance (Sadeghian and Yavari, 2004). Islam et al. (2020) screened 7 beet varieties suitable for drought conditions by carrying out drought stress experiments on 11 varieties and measuring related physiological and biochemical traits. Through stepwise regression analysis, Moosavi et al. (2017) found that root and petiole dry weight were suitable as standard traits for drought tolerance of sugar beet, and 7221 drought tolerance genotypes were screened by these traits.

Although breeding by screening drought-tolerant varieties of sugar beet can improve drought tolerance in new varieties, this method has some limitations. The transfer of agronomic traits is easily limited by intermediate reproductive isolation, which is not conducive to genetic improvement of selected crops using genetic resources of closely or distantly related species. This study aims to improve the efficiency and precision of drought-tolerant breeding of sugar beet through research methods at the molecular level.

## 4.2 Genome-wide association analysis of drought tolerance-related traits in sugar beet

With the development of genotyping technology, SNP markers are characterized by high throughput, low cost and high accuracy. The application of genome-wide association analysis is universal, and it can be used to analyze both qualitative and quantitative traits (Alqudah et al., 2020). In 175 rice samples, 13 significant SNPs were screened out and 50 genes were identified, and 10 of these genes were associated with drought and/or other abiotic stress tolerance (Pantaliao et al., 2016). It was detected 132 single nucleotide polymorphisms in sesame under drought stress, and 13 potential candidate genes were found to be related with drought tolerance (Li et al., 2018). At present, GWAS has been reported in the drought tolerance studies of many crops, but it is rarely reported in sugar beet. In this study, we explored the functional genes related to drought tolerance of sugar beet through genome-wide association analysis. 328 sugar beet germplasms were analyzed, and phylogenetic tree and PCA analysis showed that there was no obvious population stratification, indicating that the heredity of sugar beet used here was less affected by geographical factors. Considering that MLM model can better control false positive results, we used this model to excavate significant sites (Zhang et al., 2010).

## 4.3 Potential candidate genes associated with sugar beet drought tolerance

A total of 11 SNPs loci were identified by genome-wide association analysis and 25 potential candidate genes were screened out. It was found that 14 genes had been reported to be related with drought tolerance. We conducted qRT-PCR to verify gene expression levels in the two groups of drought tolerance extreme materials and found that only *BVRB_2g036960* had no difference in expression between the two groups of materials, and the expression levels of 2 genes in the drought sensitive group were significantly higher than those in the drought tolerant group. The expression levels of 11 genes in drought-tolerant group were significantly higher than those in drought-sensitive group. Therefore, these 13 differentially expressed genes might be key genes which can be used in our next research on drought tolerance in sugar beet.

*BVRB_7g160020* regulated RLPH2 is a tyrosine phosphorylase of the PPP family, and tyrosine phosphorylation is essential for signaling at the cytoplasmic structural domain (Mühlenbeck et al., 2021). We also found that the gene *BVRB_7g160030*, which is adjacent to the gene *BVRB_7g160020*, is involved in straight-chain starch synthesis in sugar metabolism. $\alpha$ (1–6) glycosidic bond branching increases the water solubility of glycogen for storage and increases the number of nonreducing ends, making it easier for biological organisms to aggregate when glycogen is needed for energy supply, and the 1,4-alpha-glucan-branching enzyme regulates glycogen metabolism and participates in the abiotic stress response (Li et al., 2019).

We found that *BVRB_7g166470* is involved in the cell wall composition. *BVRB_3g048910* can regulate methionine, which can mitigate membrane damage by peroxides. It has been shown that the probable methyltransferase PMT3 is involved in the synthesis of phosphatidylcholine (PC), which is an important precursor of lipid signalling (Baldessari and Iglesias, 2012) and a ligand for regulatory proteins (Inatsugi et al., 2009). PC appears to be a major contributor to the adaptive response of plants to environmental stress, and it plays an important role in maintaining cell membrane integrity in complex plant species. The ras GTPase-activating protein-binding protein 2 associated with *BVRB_7g166680* regulates ascorbic acid, which has strong reducing properties that effectively mitigate peroxide damage (Khazaei et al., 2020). *BVRB_9g203030* is eukaryotic initiation factor 4A-9, which promotes ATP hydrolysis to provide energy for material transport and signaling in plants, and myosin, which is a component of the cytoskeleton to provide energy for organelle movement, material transport and apical growth (Han et al., 2021). *BVRB_2g033710* is involved in the components of the cytoskeleton that provide energy for organelle movement, material transport and apex growth (Han et al., 2022). *BVRB_3g048900* on chromosome 3 was found to regulate abscisic acid, and *BVRB_5g114810* on chromosome 5 was related to MYB transcription factors. Abscisic acid can regulate MYB, which can promote lateral root growth, enhance root water absorption, and induce old leaf death to reduce water loss under drought conditions (Moosavi et al., 2017). The ENHANCER OF AG-4 protein regulated by *BVRB_9g203000* can induce apoptosis and prolong nutritional growth (Moosavi et al., 2017; Alqudah et al., 2020).

*BVRB_9g203050* on chromosome 9 regulates HCF136, which is essential for PSII biogenesis and stability (Meurer et al., 1998); the HCF136 protein is essential for the formation of early intermediates in PSII assembly, including D2 (psbD) and cytochrome b559 (Plücken et al., 2002). D2 can repair PSII and D1 expression dysfunction and maintain the stability of photosynthesis (Kanamaru and Tanaka, 2004). Protein LOW PSII ACCUMULATION (BVRB_7g166480), a required chaperone for

efficient PSII assembly, is a component of the chloroplast cyst-like membrane; it binds to psbA during the biogenesis of PSII and plays an important role in repairing PSII damage to maintain photosystem stability (Peng et al., 2006). Photosystem I (PSI) reaction centre subunit IV (BVRB_7g166490) is capable of stabilizing the interaction between PsaC and the PSI core, assisting in the docking of iron oxidation reduction protein to PSI and interacting with iron oxidation reduction protein-NADP oxidoreductase (Dutta et al., 2014).

### 4.4 Potential impact on sugar beet breeding

Although conventional breeding can improve the performance of sugar beet in drought tolerance, artificial populations are limited by population size, insufficient recombination, and often large locus intervals, making selection both time-consuming and inefficient (He et al., 2010). Most of the natural populations used here have been reproduced over long periods of time and multiple generations, with adequate genomic recombination and accurate linkage localization, and these new genomic resources will facilitate genomic evolution in sugar beet and provide a scientific and theoretical basis for relevant research in drought tolerance.

## 5 Conclusion

In this study, 11 loci potentially associated with drought tolerance in sugar beet were screened using a mixed linear model (MLM). Genome-wide association analysis combined with qRT-PCR revealed that 13 genes were significantly differentially expressed between the two extreme groups of materials, with *BVRB_7g166480* being 46.69-fold more expressed in the drought-tolerant group than in the drought-sensitive group. It prompted us to further investigate the role of these interesting genes in sugar beet drought tolerance improvement in the future.

## Data availability statement

The datasets presented in this study can be found in online repositories. The sequencing data has been uploaded. The link is "https://www.ncbi.nlm.nih.gov/sra/PRJNA935487". The accession no. is PRJNA935487.

## Author contributions

JL designed the research. WL analyzed the physiology data. DL and WT drafted the article. XY and YAZ performed the experiment. WX revised the article. ML and YUZ were tested for gene expression. All authors contributed to the article and approved the submitted version.

## Funding

The National Sugar Industry Technology System Project (CARS-170102), the Inner Mongolia Autonomous Region "the open competition mechanism to select the best candidates" project entitled "Creation of Elite Beet Germplasm and Breeding of Varieties Suitable for Mechanized Operation" (2022JBGS0029), the Accurate Identification of Sugar Beet Germplasm Resources (19230709), the Innovative Training Plan for Young Talents in Heilongjiang Ordinary Undergraduate Colleges and Universities (UNPYSCT-2020) and Tianchi Talent Program.

## Conflict of interest

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

## Publisher's note

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

## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1198600/full#supplementary-material

## References

Ahn, C. S., Han, J.-A., Lee, H.-S., Lee, S., and Pai, H.-S. (2011). The PP2A regulatory subunit Tap46, a component of the TOR signaling pathway, modulates growth and metabolism in plants. *Plant Cell.* 23 (1), 185–209. doi:10.1105/tpc.110.074005

Alqudah, A. M., Sallam, A., Baenziger, P. S., and Börner, A. (2020). Gwas: Fast-forwarding gene identification and characterization in temperate cereals: Lessons from barley–a review. *J. Adv. Res.* 22, 119–135. doi:10.1016/j.jare.2019.10.013

Bai, X., Wang, X., Wang, Y., Wei, Y., Fu, Y., Rao, J., et al. (2022). Genome-wide association study of six forage traits in ramie (boehmeria nivea L. Gaud). *Plants* 11 (11), 1443. doi:10.3390/plants11111443

Baldessari, A., and Iglesias, L. (2012). in *Lipases and phospholipases, methods and protocols*. Editor G. Sandoval (New York: Humana Press), 457–470.

Bloch, D., Hoffmann, C. M., and Märländer, B. (2006). Impact of water supply on photosynthesis, water use and carbon isotope discrimination of sugar beet genotypes. *Eur. J. Agron.* 24 (3), 218–225. doi:10.1016/j.eja.2005.08.004

Breria, C. M., Hsieh, C. H., Yen, J.-Y., Nair, R., Lin, C.-Y., Huang, S.-M., et al. (2020). Population structure of the world vegetable center mungbean mini core collection and genome-wide association mapping of loci associated with variation of seed coat luster. *Trop. Plant Biol.* 13 (1), 1–12. doi:10.1007/s12042-019-09236-0

Byrne, R. P., van Rheenen, W., van den Berg, L. H., Veldink, J. H., and McLaughlin, R. L. (2020). Dutch population structure across space, time and GWAS design. *Nat. Commun.* 11 (1), 4556. doi:10.1038/s41467-020-18418-4

Cadzow, M., Boocock, J., Nguyen, H. T., Wilcox, P., Merriman, T. R., and Black, M. A. (2014). A bioinformatics workflow for detecting signatures of selection in genomic data. *Front. Genet.* 5, 293. doi:10.3389/fgene.2014.00293

Cai, S., Jiang, G., Ye, N., Chu, Z., Xu, X., Zhang, J., et al. (2015). A key ABA catabolic gene, OsABA8ox3, is involved in drought stress resistance in rice. *PLoS One* 10 (2), e0116646. doi:10.1371/journal.pone.0116646

Carillo, P., and Gibon, Y. (2011). Protocol: Extraction and determination of proline. *PrometheusWiki* 2011, 1–5.

Deng, K., Yu, L., Zheng, X., Zhang, K., Wang, W., Dong, P., et al. (2016). Target of rapamycin is a key player for auxin signaling transduction in Arabidopsis. *Front. Plant Sci.* 7, 291. doi:10.3389/fpls.2016.00291

Dramadri, I. O., Nkalubo, S. T., Kramer, D. M., and Kelly, J. D. (2021). Genome-wide association analysis of drought adaptive traits in common bean. *Crop Sci.* 61 (5), 3232–3253. doi:10.1002/csc2.20484

Dutta, S., Teresinski, H. J., and Smith, M. D. (2014). A split-ubiquitin yeast two-hybrid screen to examine the substrate specificity of atToc159 and atToc132, two Arabidopsis chloroplast preprotein import receptors. *PloS One* 9 (4), e95026. doi:10.1371/journal.pone.0095026

Dwiningsih, Y., Thomas, J., Kumar, A., Gupta, C., Gill, N., Ruiz, C., et al. (2022). *Identification of QTLs and candidate loci associated with drought-related traits of the K/Z RIL rice population*.

Fugate, K. K., Lafta, A. M., Eide, J. D., Li, G., Lulai, E. C., Olson, L. L., et al. (2018). Methyl jasmonate alleviates drought stress in young sugar beet (Beta vulgaris L) plants. *J. Agron. Crop Sci.* 204 (6), 566–576. doi:10.1111/jac.12286

Han, H., Verstraeten, I., Roosjen, M., Mazur, E., Rýdza, N., Hajný, J., et al. (2021). *Rapid auxin-mediated phosphorylation of Myosin regulates trafficking and polarity in Arabidopsis*. BioRxiv. doi:10.1101/2021.04.13.439603

Han, M., Xu, M., Su, T., Wang, S., Wu, L., Feng, J., et al. (2022). Transcriptome analysis reveals critical genes and pathways in carbon metabolism and ribosome biogenesis in poplar fertilized with glutamine. *Int. J. Mol. Sci.* 23 (17), 9998. doi:10.3390/ijms23179998

Hassani, M., Heidari, B., and Stevanato, P. (2020). Combining abilities of sugar beet genotypes for root-and sugar-related traits under multi-environment trials. *Plant Breed.* 139 (1), 192–206. doi:10.1111/pbr.12755

He, Z., Xia, X., Bonjean, A. P., Li, F., Wang, B., and Wang, J. (2010). The potential of carcinoembryonic antigen, p53, Ki-67 and glutathion Stransferase-π as clinico-histopathological markers for colorectal cancer. *Cereals China* 688 (787), 51–57. doi:10.1016/S1674-8301(10)60008-5

Hoyos-Villegas, V., Song, Q., and Kelly, J. D. (2017). Genome-wide association analysis for drought tolerance and associated traits in common bean. *Plant Genome* 10 (1), 1–17. doi:10.3835/plantgenome2015.12.0122

Huang, X., Cloutier, S., Lycar, L., Radovanovic, N., Humphreys, D., Noll, J., et al. (2006). Molecular detection of QTLs for agronomic and quality traits in a doubled haploid population derived from two Canadian wheats (*Triticum aestivum* L). *Theor. Appl. Genet.* 113, 753–766. doi:10.1007/s00122-006-0346-7

Inatsugi, R., Kawai, H., Yamaoka, Y., Yu, Y., Sekiguchi, A., Nakamura, M., et al. (2009). Isozyme-specific modes of activation of CTP: Phosphorylcholine cytidylyltransferase in *Arabidopsis thaliana* at low temperature. *Plant Cell. Physiology* 50 (10), 1727–1735. doi:10.1093/pcp/pcp115

Islam, M. J., Kim, J. W., Begum, M. K., Sohel, M. A. T., and Lim, Y.-S. (2020). Physiological and biochemical changes in sugar beet seedlings to confer stress adaptability under drought condition. *Plants* 9 (11), 1511. doi:10.3390/plants9111511

Jha, U. C., Nayyar, H., Palakurthi, R., Jha, R., Valluri, V., Bajaj, P., et al. (2021). Major QTLs and potential candidate genes for heat stress tolerance identified in chickpea (Cicer arietinum L). *Front. Plant Sci.* 12, 655103. doi:10.3389/fpls.2021.655103

Jiao, Z., Xu, W., Nong, Q., Zhang, M., Jian, S., Lu, H., et al. (2021). An integrative transcriptomic and metabolomic analysis of red pitaya (Hylocereus polyrhizus) seedlings in response to heat stress. *Genes.* 12 (11), 1714. doi:10.3390/genes12111714

Johnson, A. A., Venkatesh, B., Jayamma, N., Reddy, B. M., Pandurangaiah, M., and Sudhakar, C. (2022). *R2R3 MYB transcription factor, AhMYB94 plays a crucial role in stress tolerance of a salt susceptible groundnut cultivar*.

Kanamaru, K., and Tanaka, K. (2004). Roles of chloroplast RNA polymerase sigma factors in chloroplast development and stress response in higher plants. *Biosci. Biotechnol. Biochem.* 68 (11), 2215–2223. doi:10.1271/bbb.68.2215

Kang, Y., Sakiroglu, M., Krom, N., Stanton-Geddes, J., Wang, M., Lee, Y. C., et al. (2015). Genome-wide association of drought-related and biomass traits with HapMap SNPs in Medicago truncatula. *Plant, Cell. & Environ.* 38 (10), 1997–2011. doi:10.1111/pce.12520

Khazaei, Z., Esmaielpour, B., and Estaji, A. (2020). Ameliorative effects of ascorbic acid on tolerance to drought stress on pepper (Capsicum annuum L) plants. *Physiology Mol. Biol. Plants* 26 (8), 1649–1662. doi:10.1007/s12298-020-00846-7

Kumar, D., Das, B., Sen, R., Kundu, P., Manna, A., Sarkar, A., et al. (2015). Andrographolide analogue induces apoptosis and autophagy mediated cell death in U937 cells by inhibition of PI3K/Akt/mTOR pathway. *PLoS One* 10 (10), e0139657. doi:10.1371/journal.pone.0139657

Kumar, D., Sharma, S., Sharma, R., Pundir, S., Singh, V. K., Chaturvedi, D., et al. (2021). Genome-wide association study in hexaploid wheat identifies novel genomic regions associated with resistance to root lesion nematode (Pratylenchus thornei). *Sci. Rep.* 11 (1), 3572. doi:10.1038/s41598-021-80996-0

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., et al. (2001). Initial sequencing and analysis of the human genome. *Nature* 409 (6822), 860–921. doi:10.1038/35057062

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25 (14), 1754–1760. doi:10.1093/bioinformatics/btp324

Li, D., Dossa, K., Zhang, Y., Wei, X., Wang, L., Zhang, Y., et al. (2018). GWAS uncovers differential genetic bases for drought and salt tolerances in sesame at the germination stage. *Genes.* 9 (2), 87. doi:10.3390/genes9020087

Li, X., Ding, M., Wang, M., Yan, F., Zhang, Z., Liang, W., et al. (2019). Gene cloning of 1,4-α-glucan branching enzyme (GBE) from nostoc flagelliforme and its differential expression in response to drought stress. *Mol. Plant Breed.* 17 (19), 6305–6313.

Liu, D., Gao, Z., Li, J., Yao, Q., Tan, W., Xing, W., et al. (2023). Effects of cadmium stress on the morphology, physiology, cellular ultrastructure, and BvHIPP24 gene expression of sugar beet (Beta vulgaris L). *Int. J. Phytoremediation* 25 (4), 455–465. doi:10.1080/15226514.2022.2090496

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

Maness, N. (2010). "Extraction and analysis of soluble carbohydrates," in *Plant stress tolerance* (Springer), 341–370.

Mathew, I., Shimelis, H., Shayanowako, A. I. T., Laing, M., and Chaplot, V. (2019). Genome-wide association study of drought tolerance and biomass allocation in wheat. *PloS One* 14 (12), e0225383. doi:10.1371/journal.pone.0225383

Meurer, J., Plücken, H., Kowallik, K. V., and Westhoff, P. (1998). A nuclear-encoded protein of prokaryotic origin is essential for the stability of photosystem II in *Arabidopsis thaliana*. *EMBO J.* 17 (18), 5286–5297. doi:10.1093/emboj/17.18.5286

Moosavi, S. G. R., Ramazani, S. H. R., Hemayati, S. S., and Gholizade, H. (2017). Effect of drought stress on root yield and some morpho-physiological traits in different genotypes of sugar beet (Beta Vulgaris L). *J. Crop Sci. Biotechnol.* 20 (3), 167–174. doi:10.1007/s12892-017-0009-0

Mühlenbeck, H., Bender, K. W., and Zipfel, C. (2021). Importance of tyrosine phosphorylation for transmembrane signaling in plants. *Biochem. J.* 478 (14), 2759–2774. doi:10.1042/BCJ20210202

Pantaliao, G. F., Narciso, M., Guimarães, C., Castro, A., Colombari, J. M., Breseghello, F., et al. (2016). Genome wide association study (GWAS) for grain yield in rice cultivated under water deficit. *Genetica* 144 (6), 651–664. doi:10.1007/s10709-016-9932-z

Peng, L., Ma, J., Chi, W., Guo, J., Zhu, S., Lu, Q., et al. (2006). LOW PSII ACCUMULATION1 is involved in efficient assembly of photosystem II in *Arabidopsis thaliana*. *Plant Cell.* 18 (4), 955–969. doi:10.1105/tpc.105.037689

Plücken, H., Müller, B., Grohmann, D., Westhoff, P., and Eichacker, L. A. (2002). The HCF136 protein is essential for assembly of the photosystem II reaction center in *Arabidopsis thaliana*. *FEBS Lett.* 532 (1-2), 85–90. doi:10.1016/s0014-5793(02)03634-7

Polash, M. A. S., Sakil, M. A., Tahjib-Ul-Arif, M., and Hossain, M. A. (2018). Effect of salinity on osmolytes and relative water content of selected rice genotypes. *Trop. Plant Res.* 5 (2), 227–232. doi:10.22271/tpr.2018.v5.i2.029

Ratnapriya, R., Sosina, O. A., Starostik, M. R., Kwicklis, M., Kapphahn, R. J., Fritsche, L. G., et al. (2019). Retinal transcriptome and eQTL analyses identify genes associated with age-related macular degeneration. *Nat. Genet.* 51 (4), 606–610. doi:10.1038/s41588-019-0351-9

Sadeghian, S., and Yavari, N. (2004). Effect of water-deficit stress on germination and early seedling growth in sugar beet. *J. Agron. Crop Sci.* 190 (2), 138–144. doi:10.1111/j.1439-037x.2004.00087.x

Shariatipour, N., Heidari, B., Tahmasebi, A., and Richards, C. (2021). Comparative genomic analysis of quantitative trait loci associated with micronutrient contents, grain quality, and agronomic traits in wheat (*Triticum aestivum* L). *Front. Plant Sci.* 12, 709817. doi:10.3389/fpls.2021.709817

Soskic, B., Cano-Gamez, E., Smyth, D. J., Rowan, W. C., Nakic, N., Esparza-Gordillo, J., et al. (2019). Chromatin activity at GWAS loci identifies T cell states driving complex immune diseases. *Nat. Genet.* 51 (10), 1486–1493. doi:10.1038/s41588-019-0493-9

Spitz, D. R., and Oberley, L. W. (1989). An assay for superoxide dismutase activity in mammalian tissue homogenates. *Anal. Biochem.* 179 (1), 8–18. doi:10.1016/0003-2697(89)90192-9

Sun, D., Cen, H., Weng, H., Wan, L., Abdalla, A., El-Manawy, A. I., et al. (2019). Using hyperspectral analysis as a potential high throughput phenotyping tool in GWAS for protein content of rice quality. *Plant Methods* 15 (1), 54–16. doi:10.1186/s13007-019-0432-x

Tan, W., Li, W., Li, J., Liu, D., and Xing, W. (2023). Drought resistance evaluation of sugar beet germplasms by response of phenotypic indicators. *Plant Signal. Behav.* 18 (1), 2192570. doi:10.1080/15592324.2023.2192570

Uhrig, R. G., Labandera, A.-M., Muhammad, J., Samuel, M., and Moorhead, G. B. (2016). Rhizobiales-like phosphatase 2 from *Arabidopsis thaliana* is a novel phospho-

tyrosine-specific phospho-protein phosphatase (PPP) family protein phosphatase. *J. Biol. Chem.* 291 (11), 5926–5934. doi:10.1074/jbc.M115.683656

Wang, Q., Sajja, U., Rosloski, S., Humphrey, T., Kim, M. C., Bomblies, K., et al. (2007). HUA2 caused natural variation in shoot morphology of A. thaliana. *Curr. Biol.* 17 (17), 1513–1519. doi:10.1016/j.cub.2007.07.059

Wang, M., Fang, Z., Yoo, B., Bejerano, G., and Peltz, G. (2021a). The effect of population structure on murine genome-wide association studies. *Front. Genet.* 12, 745361. doi:10.3389/fgene.2021.745361

Wang, X., Niu, Y., and Zheng, Y. (2021b). Multiple functions of MYB transcription factors in abiotic stress responses. *Int. J. Mol. Sci.* 22 (11), 6125. doi:10.3390/ijms22116125

Xiong, Y., McCormack, M., Li, L., Hall, Q., Xiang, C., and Sheen, J. (2013). Glucose–TOR signalling reprograms the transcriptome and activates meristems. *Nature* 496 (7444), 181–186. doi:10.1038/nature12030

Xu, Y., Yang, T., Zhou, Y., Yin, S., Li, P., Liu, J., et al. (2018). Genome-wide association mapping of starch pasting properties in maize using single-locus and multi-locus models. *Front. Plant Sci.* 9, 1311. doi:10.3389/fpls.2018.01311

Yang, H., and Wang, K. (2015). Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. *Nat. Protoc.* 10 (10), 1556–1566. doi:10.1038/nprot.2015.105

Yang, K., Chen, G., Xian, J., and Chen, W. (2022a). Varying relationship between vascular plant leaf area and leaf biomass along an elevational gradient on the eastern qinghai-tibet plateau. *Front. Plant Sci.* 13, 824461. doi:10.3389/fpls.2022.824461

Yang, Y., Dong, S., Miao, H., Liu, X., Dai, Z., Li, X., et al. (2022b). Genome-wide association studies reveal candidate genes related to stem diameter in cucumber (cucumis sativus L). *Genes.* 13 (6), 1095. doi:10.3390/genes13061095

Zeng, A., Chen, P., Korth, K., Hancock, F., Pereira, A., Brye, K., et al. (2017). Genome-wide association study (GWAS) of salt tolerance in worldwide soybean germplasm lines. *Mol. Breed.* 37 (3), 30–14. doi:10.1007/s11032-017-0634-8

Zhang, Z., Ersoz, E., Lai, C.-Q., Todhunter, R. J., Tiwari, H. K., Gore, M. A., et al. (2010). Mixed linear model approach adapted for genome-wide association studies. *Nat. Genet.* 42 (4), 355–360. doi:10.1038/ng.546

Zheng, J., Zhang, Z., Gong, Z., Liang, Y., Sang, Z., Xu, Y., et al. (2021). Genome-wide association analysis of salt-tolerant traits in terrestrial cotton at seedling stage. *Plants* 11 (1), 97. doi:10.3390/plants11010097

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 21 July 2023
DOI 10.3389/fpls.2023.1147200



OPEN ACCESS

EDITED BY
Dwijesh Chandra Mishra,
Indian Agricultural Statistics Research Institute, Indian Council of Agricultural Research, India

REVIEWED BY
Partha Biswas,
Bangladesh Rice Research Institute, Bangladesh
Bahram Heidari,
Shiraz University, Iran

*CORRESPONDENCE
Hari Krishna
harikrishna.agri@gmail.com
Pradeep Kumar Singh
pksinghiari@gmail.com

RECEIVED 18 January 2023
ACCEPTED 27 June 2023
PUBLISHED 21 July 2023

CITATION
Sunilkumar VP, Krishna H, Devate NB, Manjunath KK, Chauhan D, Singh S, Sinha N, Singh JB, T. L. P, Pal D, Sivasamy M, Jain N, Singh GP and Singh PK (2023) Marker-assisted selection for transfer of QTLs to a promising line for drought tolerance in wheat (*Triticum aestivum* L.). *Front. Plant Sci.* 14:1147200. doi: 10.3389/fpls.2023.1147200



# Marker-assisted selection for transfer of QTLs to a promising line for drought tolerance in wheat (*Triticum aestivum* L.)

V. P. Sunilkumar[1], Hari Krishna[1*], Narayana Bhat Devate[1], Karthik Kumar Manjunath[1], Divya Chauhan[1], Shweta Singh[1], Nivedita Sinha[1], Jang Bahadur Singh[1], Prakasha T. L.[1], Dharam Pal[1], M. Sivasamy[1], Neelu Jain[1], Gyanendra Pratap Singh[2] and Pradeep Kumar Singh[1*]

[1]Indian Agricultural Research Institute, ICAR, New Delhi, India, [2]National Bureau of Plant Genetic Resources, ICAR, New Delhi, India

Wheat crop is subjected to various biotic and abiotic stresses, which affect crop productivity and yield. Among various abiotic stresses, drought stress is a major problem considering the current global climate change scenario. A high-yielding wheat variety, HD3086, has been released for commercial cultivation under timely sown irrigated conditions for the North Western Plain Zone (NWPZ) and North Eastern Plain Zone NEPZ of India. Presently, HD3086 is one of the highest breeder seed indented wheat varieties and has a stable yield over the years. However, under moisture deficit conditions, its potential yield cannot be achieved. The present study was undertaken to transfer drought-tolerant QTLs in the background of the variety HD3086 using marker-assisted backcross breeding. QTLs governing Biomass (BIO), Canopy Temperature (CT), Thousand Kernel Weight (TKW), Normalized Difference Vegetation Index (NDVI), and Yield (YLD) were transferred to improve performance under moisture deficit conditions. In $BC_1F_1$, $BC_2F_1$, and $BC_2F_2$ generations, the foreground selection was carried out to identify the plants with positive QTLs conferring drought tolerance and linked to traits NDVI, CT, TKW, and yield. The positive homozygous lines for targeted QTLs were advanced from $BC_2F_2$ to $BC_2F_4$ *via* the pedigree-based phenotypic selection method. Background analysis was carried out in $BC_2F_5$ and obtained 78-91% recovery of the recurrent parent genome in the improved lines. Furthermore, the advanced lines were evaluated for 2 years under drought stress to assess improvement in MABB-derived lines. Increased GWPS, TKW, and NDVI and reduced CT was observed in improved lines. Seven improved lines were identified with significantly higher yields in comparison to HD3086 under stress conditions.

KEYWORDS

**HD3086, drought tolerance, foreground selection, QTLs, MABB**

## Introduction

Wheat is the most important food grain grown worldwide as well as in India. On average, 35% of the world's population depends on wheat as their staple food (IDRC, 2010). More than two-thirds of the world's wheat is used for human consumption, while one-fifth is used for animal feed (Grote et al., 2021). It is estimated that a 60% increase can be expected in the consumption of wheat-based products by 2050 due to an expanding global population and food demand. In order to meet global demand, wheat yields need to be increased by 1.6% per year (Tilman et al., 2011). However, the average yield of wheat worldwide is much lower than its potential (Asseng et al., 2020; Rashid et al., 2022). Due to change in climatic conditions, wheat crop faces challenges like biotic and abiotic stress, which significantly reduce crop yield. New diseases and pests have been posing a significant threat to wheat production; coupled with increasing drought and heat stress due to changing environmental conditions (Gupta et al., 2010). Among abiotic stresses, recurrent drought has a major impact on agriculture through alteration in the phenology of crops and changes in diseases and insect dynamics, which ultimately reduces the potential yield (Daryanto et al., 2016; Nalley et al., 2018). Currently, 70% of the wheat-growing area worldwide experiences moisture deficiency stress (Portmann et al., 2010); nearly 50% of the wheat grown in developing nations is rain-fed, receiving an average of 600 mm of precipitation per year, with occasional lows of 350 mm. Additionally, only 1-2 irrigations are given to 66 - 80% of the wheat grown in irrigated conditions, which results in a reduction in yield (Joshi et al., 2007). It is anticipated that annual precipitation will fall by 4–27% in various parts of the world and the temperature will hike by 1.5°C as a result of global warming over the next 10 years (IPCC, 2021). These two abiotic stresses hamper production because they frequently coexist throughout the grain-filling stage in dry or semi-arid environments (Wardlaw and Willenbrink, 2000; Sallam et al., 2015). Wheat crop yield loss due to heat and drought can reach up to 86% and 69%, respectively (Prasad et al., 2011; Yang et al., 2021). Drought is an inadequacy of water, including precipitation and stored soil moisture, required for crop growth, both in terms of distribution and quantity, which results in a restricted expression of the genetic yield potential (Sinha et al., 1986).

The North Western Plains Zone and North Eastern Plains Zone of India together contribute 78 MT of wheat production in an area of 21 mha. In recent years, the ICAR-Indian Agriculture Research Institute (IARI) has contributed several high-yielding wheat varieties, including HD2967 and HD3086, which together account for 40% of the nation's total wheat-grown area. In Indo-Gangetic plains, the high-yielding wheat variety HD3086 has been made available for commercial production under timely sown, irrigated conditions. It alone accounts for 11.6% of the wheat varieties indent and 34% of breeder seed indent in India (https://seednet.gov.in/). Due to its adaptability and higher yield, the breeder seed requirement for this variety is growing every year. However, its potential yield is reduced under moisture deficit stress conditions. According to previous reports, 68-70% of arable Indian land is under drought stress, especially in the wheat growing belt, including NWPZ and NEPZ, due to erratic rainfall and depleting stored water content (Mondal et al., 2017) leading to reduced wheat yield.

Drought tolerance is a complex trait; expression is controlled by polygenes and is influenced by various environmental factors (Gupta et al., 2012; Farooq et al., 2014). The complex inheritance mechanism of stress tolerance traits, limited genetic diversity of yield components under stress conditions, and the dearth of efficient selection techniques limit traditional breeding strategies for developing drought-tolerant varieties (Gupta et al., 2012; Gupta et al., 2017; Salarpour et al.),. An understanding of the genetic architecture of drought-related traits and information on relevant candidate genes/QTLs to develop drought-tolerant cultivars are necessary. In the past, many QTLs/meta-QTLs for yield and associated traits have been identified in wheat under drought stress which accounts 19-59% of the phenotypic variance (Quarrie et al., 2006; (Kirigwi et al., 2007; Shukla et al., 2015; Tahmasebi et al., 2016; Salarpour et al., 2021; Gupta et al., 2017). Apart from this, drought-related QTL mapping studies were also conducted under drought stress (Salarpour et al., 2021 and Salarpour et al., 2020) and heat stress conditions (Tahmasebi et al., 2016) in wheat. Early heading and anthesis, canopy temperature (CT), normalized difference vegetative index (NDVI), water-soluble carbohydrates (WSC), and chlorophyll content are key target agronomic and physiological traits for enhancing wheat crop ability to withstand drought (Bayoumi et al., 2008; Dolferus, 2014; Naeem et al., 2015; Afzal et al., 2017; Shamuyarira et al., 2019; Sobhaninan et al., 2019). Due to the complex inheritance pattern of drought tolerance, conventional breeding for these traits is difficult and time-consuming, hence schemes including quantitative trait loci (QTL) mapping and marker-assisted breeding should be used (Bahari et al., 2014). Additionally, it has been demonstrated that traditional breeding, in conjunction with marker-assisted selection (MAS), is effective for breeding complex traits, such as resistance to biotic and abiotic stresses, in a variety of crops, including wheat (Bustos et al., 2001; Gupta et al., 2010; Kumar et al., 2010; Kumar et al., 2011; Tyagi et al., 2014). The wheat variety DBW43 is known to perform better under moisture deficit conditions and also carries genes for leaf rust and yellow rust resistance. Hence, in the present study, DBW43 was used as a donor parent to transfer QTLs linked to component traits of drought tolerance in the background of HD3086 using marker-assisted backcross breeding (MABB).

## Materials and methods

### Plant materials and generation of improved lines

In the present MABB scheme, drought-tolerant germplasm line DBW43 was used as a donor parent, whereas HD3086 was used as the recurrent parent. Recurrent parent HD3086 is a popular high-yielding variety with a pedigree "DBW14/HD2733//HUW468". It has semi-erect growth (99-101cm), has a 143-day period to maturity, and provides an average yield of 5.46 t/ha, but its

potential yield is reduced under moisture deficit stress conditions. However, the donor parent DBW43 is the introduced line from CIMMYT, with pedigree "BABAX/*Lr42*//BABAX*2/3/VIVITSI", and is identified at IIWBR, Karnal. The variety has an erect growth habit (110 cm) and is known to perform well under moisture deficit conditions; it also shows resistance to some races of yellow and leaf rust. Furthermore, DBW43 is a moisture deficit stress and heat-tolerant germplasm line; it is highly resistant to yellow and leaf rust and provides an average yield of 49 q/ha in RI conditions (Harikrishna, 2017; Singh et al., 2016).

The $F_1$ plant's hybridity was tested to check true F1s based on the presence of the hybrid band, using the linked SSR marker X*gwm484*. X*gwm484* was a polymorphic marker between parents and also a linked marker with the QTL related to drought-tolerant traits such as BIO, WSC, and Yield. True individual $F_1$ plants were backcrossed twice with the recurrent parent HD3086 to generate $BC_1F_1$ and $BC_2F_1$. The following generations were handled by the procedure shown in Figure 1. The scheme comprises selection in each backcross generation, the foreground selection was performed with SSR markers linked to targeted drought-tolerant QTL regions (Table 1), followed by phenotypic selection (PS) for the plants similar to the HD3086 recurrent parent phenome (RPP). Foreground selection was undertaken in $BC_1F_1$, $BC_2F_1$, and $BC_2F_2$ (Table 2). In each step, foreground screening for QTLs was carried out to select the plants before making backcross in $F_1$ and $BC_1F_1$. The $BC_2F_1$ plants were also screened for drought-tolerant QTLs, and the selected plants were subjected to selfing to generate $BC_2F_2$ progenies. Furthermore, the selected homozygous $BC_2F_2$ progenies for targeted QTLs were advanced *via* pedigree-based phenotypic selection up to $BC_2F_4$ generation. The advancement of improved lines was done in New Delhi during the main season and in off-season nurseries such as Wellington and Lahul-Spiti. In each advanced generation, the phenotypic selection was done to achieve maximum recovery of the recurrent parent phenome. The selected $BC_2F_4$ and $BC_2F_5$ lines were tested for two years for tolerance to moisture stress in field conditions. The $BC_2F_5$ individuals that performed better than the recurrent parent for targeted traits were advanced to seed multiplication for nomination under the MABB trials of the All India Coordinated Research Projects (AICRP).

## Foreground selection

The techniques outlined by Hospital and Charcosset (1997) were used for foreground selection. The linked QTLs to drought-tolerant traits such as NDVI, CT, CC, DH, and Yld were used in MABB. SSR markers linked to these QTLs were validated using the RIL population generated by a cross between HI1500 x DBW43 (Harikrishna, 2017). The QTL Qndvi6.iari-4A with the phenotypic variance of 14% ($R^2$ = 0.14) linked to SSR marker *Xwmc617*, a CT and TKW related meta QTL 'MQTL24' located on 3A linked to SSR marker *Xwmc640* with PV of 14.01%, and a biomass, WSC, and yield-related QTL located on 2D linked to SSR marker *Xgwm484* with a PV of 9.45% (Yang et al., 2007; Pinto et al., 2010; Kadam et al., 2012; Acuñ-Galindo et al., 2015; Gu et al., 2015; Harikrishna, 2017) were used in this study to transfer respective traits.

Genomic DNA was isolated from fresh leaf tissue using the CTAB procedure, DNA was quantified using a Nanodrop, and purity was tested using 0.8% agarose gel electrophoresis with λ DNA as a standard. PCR reaction was performed with a volume of 15 µl comprising 30-40 ng of template DNA, 5 pmol of each primer, 0.05 mM dNTPs, and 10X PCR buffer (10 mM Tris, pH 8.4, 50 mM KCl, 1.8 mM MgCl2, and 0.5 U of *Taq DNA polymerase*) (Bangalore

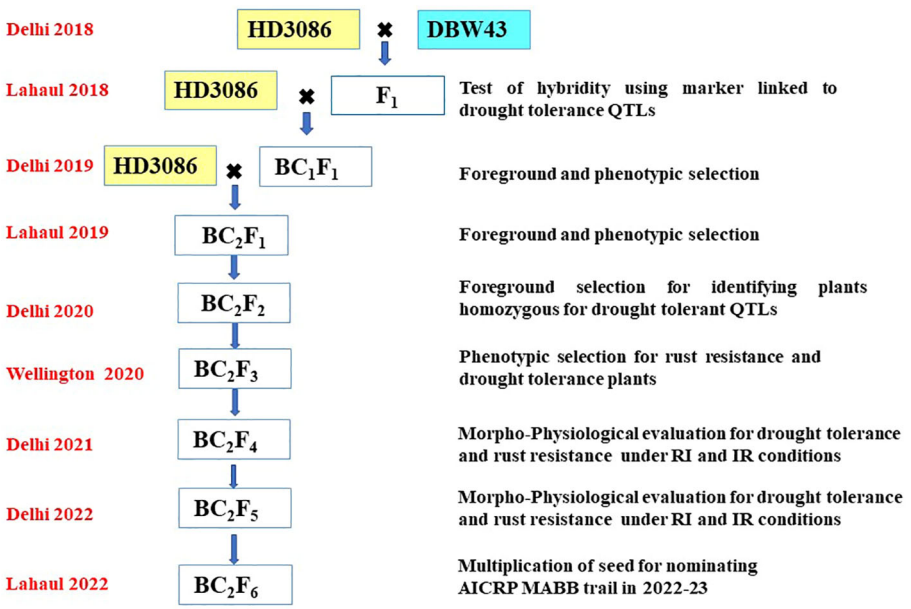



FIGURE 1
Schematic workflow of marker-assisted backcross breeding of HD3086*2/DBW43; the crossing pattern, season, and location of advancement and steps followed in each season are given in the figure.

10.3389/fpls.2023.1147200

TABLE 1 Sequence of SSR primers and their amplification details for linked SSR markers used to transfer drought-tolerant QTLs in HD3086*2/DBW43 population.

| SL.NO | Gene/ Marker | Chromosome | Primer sequence (5′ – 3′) | Position (cM) | Annealing temperature (°C) | PVE (%) | Product size (bp) | | Associated/Reported Traits | Reference |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | HD3086 | DBW43 | | |
| 1 | *Xgwm484* | 2D | F - ACATCGCTCTTCACAAACCC<br>R - AGTTCCGGTCATGGCTAGG | 45 | 55 | 9.45 Kadam et al. (2012) | 160 | 180 | BIO, WSC, YLD | Yang et al. (2007) |
| 2 | *Xwmc617* | 4A | F - CCACTAGGAAGAAGGGGAAACT<br>R - ATCTGGATTACTGGCCAACTGT | 13 | 61 | 14 (Harikrishna, 2017) | 220-230 | 230-260 | YLD | Pinto et al., 2010 |
| 3 | *Xwmc640* | 3A | F - AATTTATCTCGATCATGTGAGC<br>R - TGAGTAGTTCCCTTAGGACCTT | 43-53 | 61 | 14.01 Gu et al. (2015) | 130<br>180 | 140 | TKW, CT, DM (MQTL24) | Acuña-Galindo et al. (2015) |

Genei Pvt. Ltd., India). SSR markers from the *Xgwm* and *Xwmc* series were used to perform PCR amplification of the template DNA (Table 1). The amplified product was resolved in gel electrophoresis using 3% Agarose SFR gel and observed on Gel documentation systems (Supplementary Figure 1).

## Background analysis

To identify plants with the maximum recovery of the recurrent parent genome (RPG), selected $BC_2F_5$ lines along with parents were genotyped utilizing hybridization-based Wheat Breeder's 35K Axiom Array SNP chips of the Affymetrix GeneTitan$^R$ system. This array contained 35,143 SNPs that were evenly dispersed throughout the wheat genome. However, after filtering for monomorphic alleles, minor allele frequency (MAF), and missing data, a total of 3706 polymorphic SNPs were chosen between parents DBW43 and HD3086 and were used for background analysis. Using the GGT 2.0 (Graphical Geno Typing 2.0) software, the background recovery of the recurrent parent was graphically visualized (Van Berloo, 2008). The following formula was used to determine the recurrent parent's contribution to the background of MABB-generated lines:

$$G = [(B + 1/2A) \times 100]/N$$

were,

N = total number of parental polymorphic markers screened

B = number of markers showing homozygosity for recurrent parent allele

A = number of markers showing heterozygosity for parental alleles.

## Screening for leaf rust resistance at the seedling stage

In each backcross population, HD3086*2/DBW43 and parents were tested at the seedling stage for resistance to *P. triticina* races 77-5 and 77-9 using Single Race Testing (SRT) at IARI New Delhi and IARI Regional Station, Indore. Seedlings, at approximately 8–10 days old were inoculated with spores/suspension early in the evening. After inoculation, seedlings were raised under a humid glass chamber for 36 hours and maintained an ambient temperature of 23 ± 2°C and 85% relative humidity. Leaf rust resistance was scored for each race using the 0–4 modified Stakman Scale (Roelfs, 1992; Table 3). The disease reaction was recorded 12–14 days after the inoculation.

## Screening of genotypes at the adult stage for leaf and yellow rust resistance

Field screening for leaf and stripe rust was done in each backcross generation at IARI, New Delhi during the main season, and in off seasons at Wellington and Shimla. "Agra local", a rust-susceptible landrace, served as the infector and was planted around the experimental plots after every 20 rows of backcross population were planted for screening. The lines under testing were sprayed

TABLE 2 Number of plants selected in each generation in HD3086*2/DBW43 population.

| Generation | No. of plants | Foreground Selection | |
|---|---|---|---|
| | | No. of plants selected | |
| $BC_1F_1$ | 98 | 44 | FS+PS |
| $BC_2F_1$ | 160 | 25 | FS+PS |
| $BC_2F_2$ | 830 | 140 | FS+PS |
| $BC_2F_3$ | 140 | 120 | PS |
| $BC_2F_4$ | 120 | 42 | BS+PS |
| $BC_2F_5$ | 10 | 7 | PS |
| $BC_2F_6$ | 7 | 1 | PS |

FS, Foreground selection; PS, Phenotypic selection.

with a mixture of urediospores from the prevalent stripe and leaf rust pathotypes to create artificial rust epiphytotic. To guarantee uniform disease propagation, rust-infected pots were positioned in fields between the experimental materials. The infected leaves of the susceptible host were suspended in water containing urediniospores (5.6 g/ha) and 0.75 l/ml Tween20 (surfactant) following the method described by Imtiaz et al. (2003). According to the modified Cobb scale provided by Peterson et al. (1948), the disease severity (DS) and infection response (IR) were recorded at the reproductive stage. The DS was expressed in percentage (0-100%) and the IR was recorded as S (susceptible), MS (moderately susceptible), MR (moderately resistant), and TR (trace).

## Evaluation for drought tolerance traits

The selected backcross populations were evaluated for agronomic and physiological parameters for two years at the IARI experimental farm in New Delhi (280 40'N, 770 13'E, MSL228m). In the year 2020-2021, parents (viz., HD3086, DBW43), three check varieties (viz., HI1500, GW322, and BABAX), and the selected positive $BC_2F_4$ lines were evaluated in an augmented design under irrigated conditions (IR) and restricted irrigation (RI) conditions. Irrigated trials received a total of six irrigations, whereas restricted irrigated trials received only one irrigation (21 days after sowing, in addition to pre-sowing irrigation) to induce terminal drought stress. The experiment was conducted in the same field with a divider to restrict irrigation water flow to the RI treatment. Weather data during the growing season of wheat (November to March) during 2020-2021 and 2021-2022 are included in Supplementary Table 1. A plot size of 0.63 $m^2$ was maintained, in which each plot consisted of three rows with a spacing of 23 cm between rows. The $BC_2F_5$ lines, which had performed better in the previous generation, were evaluated in RCBD (randomized complete block design) using large plots of size 7.2 $m^2$ (6 m x 1.2 m) under IR, RI, and late sown (LS) conditions in the year 2021-22. The wheat crop was raised using standard agronomic management techniques.

The "Wheat Physiological Breeding II: A Field Guide to Wheat Phenotyping" (Pask et al., 2012) manual was used as the guide for

TABLE 3 Leaf rust scoring method according to modified Stakman scale (Roelfs, 1992).

| Infection types | Leaf rust response/reaction | Disease response |
|---|---|---|
| 0 | No flecks or uredinia | Immune |
| 0; | faint Hypersensitive flecks | Highly resistant |
| ; | Hypersensitive flecks | Highly resistant |
| 1 | Small uredinia with necrosis | Resistant |
| 2 | Small to medium uredinia with necrosis | Moderately susceptible |
| 3 | Moderate to large size uredinia with/without chlorosis, | Susceptible |
| 4 | Very large uredinia without chlorosis | Highly susceptible |
| X | Mesothetic, a mixture of resistant pustule types | Mesothetic |
| "+" | Indicates slightly larger uredinia | – |
| "-" | Indicates slightly smaller uredinia | – |

If more than one kind of reaction is observed in a plant, they are written with consecutive scores, For example, if a plant has both 1 and 2 reactions, it is written as "1 2". Similarly, if a line has; and 1 reactions, it is written as "; 1"

the standard method of data collection in improved lines (Supplementary Table 2). Agronomic traits like days to heading (DH), days to maturity (DM), plant height (PH), plot yield (PY), thousand kernel weight (TKW), and grain weight per spike (GWPS) were recorded from each entry. Physiological parameters including NDVI, CT, and soil plant analysis development (SPAD) chlorophyll content were measured at three distinct stages, namely, the vegetative stage (late boot stage, Z49), grain filling stage (early milk stage, Z73), and grain maturity stage (late milk stage, Z85) (Zadoks et al., 1974).

## Statistical analysis

Descriptive statistics and analysis of variance were calculated for $BC_2F_4$ individuals using an R package 'augmented RCBD' (Aravind et al., 2021). Comparisons of individuals were conducted based on adjusted mean calculated with the formula (Federer, 1961)

$Vi = ui - bj$ Where

Vi is the Adjusted mean of $i^{th}$ variety

$ui$ is the Unadjusted mean of $i^{th}$ variety

$bj$ is $j^{th}$ block effect

Whereas the $BC_2F_5$ lines were planted in RCBD with three replications, and the analysis of variance, Least Significant Difference (LSD), and Coefficient of Variation (CV) were calculated from MS Excel following standard procedure. The lines that significantly performed better under moisture deficit stress were identified.

# Results

## Development of backcross population following foreground selection and background analysis

### Foreground selection for drought-tolerant QTLs

To transfer the drought tolerance QTLs, the recurrent parent HD3086 was crossed with the donor parent DBW43, and $BC_1F_1$ progenies were produced. The true hybrid plants from the cross were backcrossed with recurrent parent HD3086 to enhance the recurrent genome portion in the progenies. A total of 98 $BC_1F_1$ plants were screened for the presence of different foreground markers linked with QTLs of interest. Among them, 44 plants were found to contain positive alleles for required QTLs and were selected for the second round of backcrossing. To generate $BC_2F_1$, the selected plants were again backcrossed with the recurrent parent HD3086. Among 160 $BC_2F_1$ plants, 25 plants were positive for targeted drought tolerant QTLs and were selfed to develop $BC_2F_2$ progenies. All the 830 $BC_2F_2$ plants obtained were grown and homozygosity for the markers linked with donor QTLs was identified from foreground selection. 140 $BC_2F_2$ homozygous plants were selected *via* foreground selection with a different combination of QTLs related to drought tolerance and were advanced up to $BC_2F_4$ through pedigree-based phenotypic selection.

## Background recovery of the recurrent parent genome

Improved $BC_2F_5$ homozygous lines were subjected to background analysis for recovery of the recurrent parent genome, and eight lines were identified with maximum recovery of the recurrent parent genome along with drought-tolerant QTLs from the donor (Figure 2). The recurrent genome recovery of improved lines ranged from 78%-91%, with an average recovery of 85.5%. Improved lines had maximum visual similarity at the phenotypic level with recurrent parent HD3086. Among these, two lines, viz., HD3086-7-1-210-26 and HD3086-3-15-174-22, had a maximum recurrent parent genome recovery of 92% and 90%, respectively. The targeted QTLs from the donor parent, which are spread over 3A, 4A, and 2D chromosomes, were transferred to the progenies developed in this study (Figure 2).

## Phenotypic screening for rust resistance

Single race testing (STR) for leaf rust resistance at the seedling stage confirmed a similar susceptibility reaction of recurrent parent HD3086 to that of check Agra local ($3\ 3^+$) for both pathotypes, 77-5 and 77-9. However, donor parent DBW43 had a Highly Resistant (HR) response (0); for pathotypes 77-5 and 77-9 (Table 4). Furthermore, 15 $BC_2F_5$ drought-tolerant lines were subjected to SRT for *P. triticina* pathotypes 77-5 and 77-9, of which 11 lines showed resistance reactions (such as 3HR, 3R, and 5MR), and the rest of the 4 lines displayed susceptible reactions at the seedling stage. Among these improved lines, HD3086-5-1-189-24, HD3086-6-6-209-25, and HD3086-13-10-279-34 showed a highly resistance reaction, which is similar to the resistance of parent DBW43.

Backcross-derived lines ($BC_1F_1$, $BC_2F_1$, $BC_2F_2$) with positive drought-tolerant QTLs and resistance to leaf and stripe rust were advanced to the next generation. In $BC_2F_3$, the rust severity of progenies was assessed along with the parents. The recurrent parent HD3086 showed a severity of 60S DS for leaf rust, whereas improved lines showed 0 to 20S DS. Improved lines for drought tolerance and complete resistance to leaf and stripe rust were selected in advanced generations.

## Morpho-physiological performance of the selected lines for drought tolerance

The improved lines from the cross HD3086*2/DBW43 carrying drought tolerance QTLs were evaluated for 2 years under IR and RI conditions. In the year 2020-2021, a total of 120 $BC_2F_4$ lines were evaluated in an augmented design, among which, 42 lines were found significantly superior to the recurrent parent and checks for introgressed traits. Furthermore, these 42 superior $BC_2F_4$ lines were evaluated for component traits of drought tolerance, and seven improved lines superior to recurrent parent HD3086 and the maximum recurrent parent phenome were identified. In the current study, leaf chlorophyll index (LCI) in $BC_2F_5$ ranged from

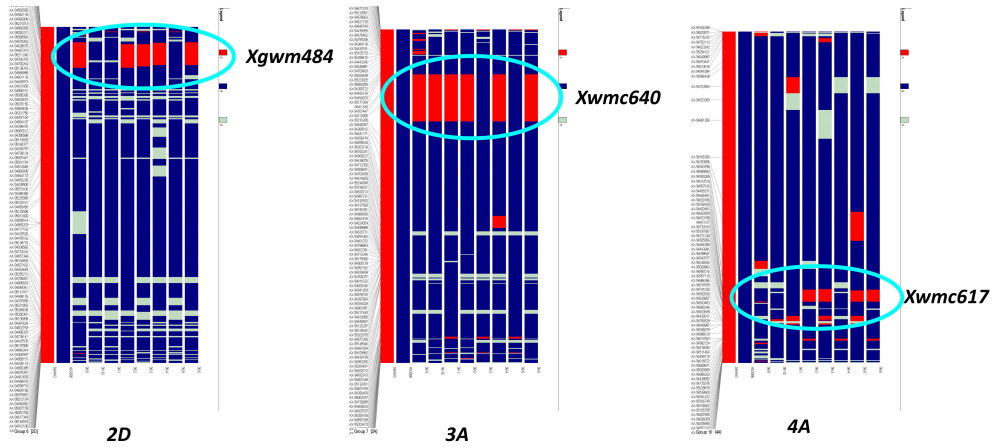



FIGURE 2
Recurrent parent recovery of chromosomes 2D, 3A, and 4A; first two chromosomes indicate donor (red) and recipient (blue) chromosomes followed by progeny. The red portion indicates donor genome segment; the blue portion indicates recipient genome segments. Chromosomal location and donor segment linked to SSR markers *Xgwm484*, *Xwmc640*, and *Xwmc617* are shown with circles.

44.8 to 52.4, and approximately seven progenies outperformed the recurrent parent HD3086 in terms of LCI. The NDVI value of these selected lines ranged from 0.34 to 0.55, with a mean value of 0.45, which is more than the mean value (0.32) of recurrent parent HD3086 during the grain maturity stage. Eight progenies show lower canopy temperature than the recurrent parent; the mean CT value was 32.8 °C and ranged from 32.1°C to 34.4°C. Similarly, improved lines showed superior agronomic performance for the traits like TKW, GWPS, and yield under RI conditions (Table 5).

The distinctiveness, uniformity, and stability (DUS) of selected $BC_2F_5$ lines were assessed. Finally, 3 out of 10 progenies (viz., HD3086-6-6-209-25, HD3086-7-11-220-30, and HD3086-13-10-

TABLE 4 Screening of marker-assisted derived wheat genotypes for leaf rust races at IARI New Delhi and IARI regional station, Indore ($BC_2F_5$).

| Sl.no | Progenies | Rust race (77-5) | | Rust race (77-9) | | Reaction of genotypes |
|---|---|---|---|---|---|---|
| | | IARI RS, Indore | IARI, New Delhi | IARI RS, Indore | IARI, New Delhi | |
| 1 | HD3086 (Recipient parent) | $33^{+}$ | 3 | 2 3 | 3 | Susceptible (S) |
| 2 | DBW43 (Donor parent) | 0; | 0 | 0; | 0; | Highly Resistant (HR) |
| 3 | HD3086-1-3-126-21 | 3 | 0 | 3 | 0 | S |
| 4 | HD3086-3-15-174-22 | ; 1 | 1 | ; 2 | ; | R |
| 5 | HD3086-4-4-184-23 | ; | ; 1 | ; 2 | 2 | MR |
| 6 | HD3086-5-1-189-24 | ; | ; | ; | 1 | HR |
| 7 | HD3086-6-6-209-25 | ; 1 | 1 | ; 1 | ;1 | HR |
| 8 | HD3086-7-1-210-26 | 0 | 1 2 | ; 1 | 2 | MR |
| 9 | HD3086-7-3-212-27 | ; 1 | 1 | ; $1^{-}$ | 1 2 | R |
| 10 | HD3086-7-6-215-28 | ; 1 | 1 2 | ; 2 | 1 2 | MR |
| 11 | HD3086-7-7-216-29 | ; 1 | ; | ; 2 | ; 1 | R |
| 12 | HD3086-7-11-220-30 | 0; | 2 | 3 | 1, 2 | MS |
| 13 | HD3086-10-1-244-31 | 0 | 0, 1 | $3^{+}$ | 0,; | MS |
| 14 | HD3086-11-6-257-32 | ; 2 | 1; | ; 2 | 1 2 | MR |
| 15 | HD3086-13-6-275-33 | 0; | 1 | 2 3 | ; 1 | MR |
| 16 | HD3086-13-10-279-34 | 0; | 0; | ; | 1 | HR |
| 17 | HD3086-15-10-293-35 | $3^{+}$ | 1 | ; 2 | 0 | S |
| 18 | Agra Local | 3+ | 3+ | 3+ | 3+ | HS |

HR, Highly resistant; R, Resistant; MR, Moderately resistant; MS, Moderately Susceptible; S, Susceptible; HS, Highly susceptible.

TABLE 5 Morpho-physiological characteristics of HD3086*2/HI1500 selected $BC_2F_5$ lines for component traits of drought tolerance under moisture deficit stress condition.

| S.no | Selected progeny | QTLs linked trait | Chlorophyll content | | Canopy temperature (°C) | | | NDVI | | | GWPS (g) | TKW (g) | YLD (q/ha) | Total RPG % |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | VS | GFS | VS | GFS | GMS | VS | GFS | GMS | | | | |
| 1 | HD3086-1-3-126-21 | CT | 49 | 50 | 25.6 | 30.3 | 32.9 | 0.78 | 0.65 | 0.34 | 1.24 | 33.5 | 44.18 | 78.31 |
| 2 | HD3086-3-15-174-22 | NDVI, TKW, YLD | 50.5 | 51 | 24.9 | 28.8 | 32.5 | 0.79 | 0.73 | 0.48 | 1.63 | 38.1 | 49.58 | 89.81 |
| 3 | HD3086-4-4-184-23 | NDVI, BIO, YLD | 44.8 | 48.9 | 25.3 | 30.6 | 34.4 | 0.77 | 0.7 | 0.45 | 1.61 | 34.8 | 45.57 | 86.47 |
| 4 | HD3086-5-1-189-24 | BIO, GWPS | 47.9 | 51.1 | 24.7 | 30 | 33.5 | 0.75 | 0.67 | 0.51 | 1.53 | 33.2 | 49.17 | 88.46 |
| 5 | HD3086-6-6-209-25 | TKW, YLD | 48.8 | 51.2 | 25.3 | 29.5 | 32.3 | 0.81 | 0.72 | 0.55 | 1.69 | 27.6 | 36.98 | 80.28 |
| 6 | HD3086-7-1-210-26 | NDVI, TKW, YLD | 48.5 | 52.4 | 25.8 | 29.3 | 32.1 | 0.77 | 0.62 | 0.53 | 2.09 | 42.3 | 50.28 | 91.31 |
| 7 | HD3086-7-11-220-30 | BIO | 46.7 | 49.5 | 27.6 | 32.5 | 32.3 | 0.75 | 0.64 | 0.4 | 1.47 | 36 | 40.30 | 83.47 |
| 8 | HD3086-10-1-244-31 | GWPS, CT | 48.1 | 49.8 | 27.3 | 32.1 | 32.1 | 0.75 | 0.56 | 0.44 | 1.38 | 36.5 | 43.35 | 87.94 |
| 9 | HD3086-11-6-257-32 | NDVI, TKW | 48.6 | 49.6 | 27.1 | 30.3 | 33.5 | 0.78 | 0.69 | 0.45 | 1.83 | 32.7 | 46.54 | 83.75 |
| 10 | HD3086-13-10-279-34 | CT, TKW | 46.3 | 49.6 | 27.7 | 30.5 | 32.9 | 0.76 | 0.66 | 0.4 | 1.17 | 34.4 | 41.00 | 85.43 |
| | HD3086(RI) | | 48.5 | 49.8 | 29.1 | 31.3 | 33.2 | 0.78 | 0.61 | 0.32 | 1.01 | 30.4 | 42.66 | |
| | HD3086(IR) | | 49.4 | 52.6 | 22 | 31.3 | 32 | 0.79 | 0.72 | 0.54 | 1.65 | 44.6 | 62.60 | |
| | DBW43(RI) | | 50.1 | 52.3 | 26 | 30.5 | 31.9 | 0.75 | 0.66 | 0.5 | 1.58 | 36.5 | 49.86 | |
| | DBW43(IR) | | 51.2 | 52.5 | 22.7 | 28.6 | 31.6 | 0.76 | 0.7 | 0.58 | 2.01 | 39.7 | 53.88 | |
| | Mean | | 47.98 | 50.43 | 26.37 | 30.48 | 32.80 | 0.77 | 0.66 | 0.45 | 1.52 | 34.67 | 44.96 | |
| | LSD at 5% | | 2.93 | 2.84 | 1.39 | 1.33 | 2.18 | 0.03 | 0.03 | 0.02 | 0.09 | 1.95 | 2.32 | |

LSD, Least Significant Difference. NDVI, Normalized Difference Vegetation Index; GL, Grain Length; GWPS, Grain Weight Per Spike; TKW, Thousand Kernel Weight; YLD, Yield; RPG, Recurrent Parent Genome; RI, Rainfed; IR, Irrigated; VS, Vegetative stage; GFS, Grain Filling Stage; GMS, Grain Maturity Stage.

279-34) showed lower performance than the recurrent parent, and the rest of the seven progeny lines (*viz*., HD3086-1-3-126-21, HD3086-3-15-174-22, HD3086-4-4-184-23, HD3086-5-1-189-24, HD3086-7-1-210-26, HD3086-10-1-244-31, and HD3086-11-6-257-32) performed better in both the years than recurrent parent HD 3086. Furthermore, they were found to have a maximum recovery of both RPG and RPP and were further nominated for All India Coordinated Wheat Improvement Project (AICWIP) trials (Figure 3). Recently, in 2022-23, the HD3086-7-1-210-26 derived line was proposed for national trials for testing and further release for commercial cultivation under the designation HD3470.

## Discussion

Abiotic stress, particularly heat and drought, significantly declines wheat productivity. The ability to absorb nutrients is significantly affected by drought stress, which results in stunted growth and low yields in arid zones (Yasmin et al., 2019). Drought stress causes DNA, lipid, and membrane damage as well as the generation of reactive oxygen species in mitochondria, peroxisomes, and chloroplasts, which ultimately leads to the destruction of plant metabolism (Rashid et al., 2022). The majority of traits related to better performance under drought stress, such as RWC, days to heading, awn length, and spikelets per spike, had a positive relationship with yield (Sobhaninan et al., 2019) and are quantitatively inherited and governed by many genes/QTLs. Thousands of QTLs for agronomic and yield-contributing traits have been identified under different stress conditions in several previous studies (reviewed in Pinto et al., 2010; Gupta et al., 2012; Gupta et al., 2017; Shashikumara et al., 2020; Devate et al., 2022). However, their particle utilization in plant breeding needs to be emphasized with marker-assisted breeding programs. MABB is one of the practical and affordable methods of marker-assisted selection (MAS) that involves the transfer of desired traits from a donor

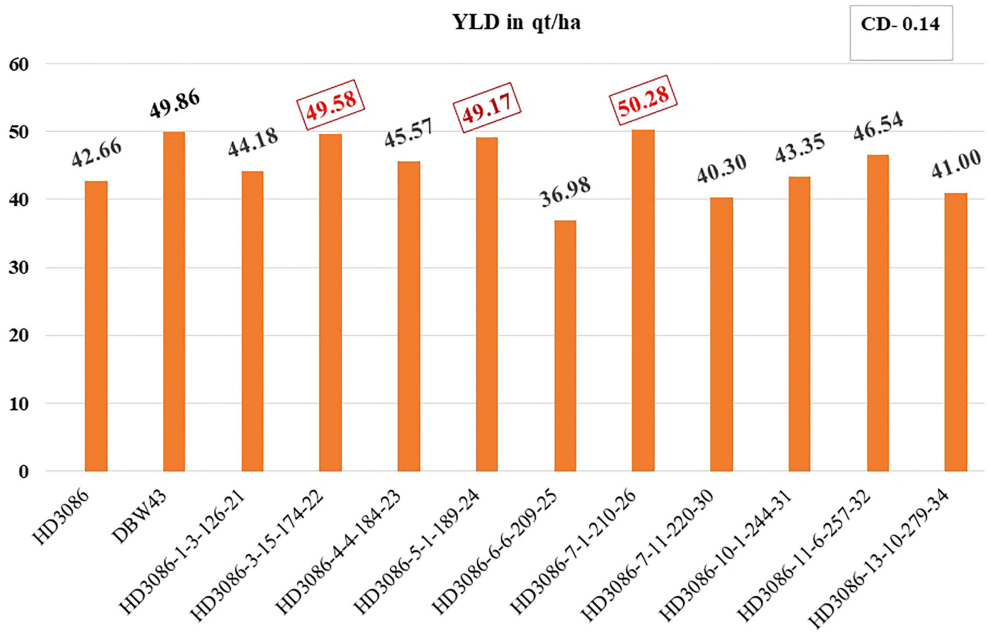



FIGURE 3
Improvement in grain yield of selected lines of HD3086*2/DBW43 derived BC2F5 population under moisture stress. Grain yield performance of parental lines (HD3086 and DBW43) along with the improved lines are given in quintals per hectare.

parent to an elite recurrent parent, with minimum changes in the recurrent parent's genetic background (Collard and Mackill, 2008; Bapela et al., 2022). MAS has been utilized frequently in bread wheat for the introgression of traits, including disease resistance and quality enhancement (Kumar et al., 2010; Malik et al., 2015; Shah et al., 2017; Sharma et al., 2021). However, a small number of studies have previously transferred QTLs for promising wheat varieties to increase drought tolerance and yield under moisture deficiency conditions (Merchuk-Ovnat et al., 2016; Rai et al., 2018; Gautam et al., 2020; Todkar et al., 2020; Sunilkumar et al., 2022). Wheat variety HD3086 is the choice of millions of farmers in Indo-Gangetic plains; its potential yield is reduced under restricted irrigation (RI) conditions. To achieve this, the current study aimed to transfer QTLs linked to component traits of drought tolerance in the genetic background of HD3086 through MABB.

Foreground selection was carried out in $BC_1F_1$, $BC_2F_1$, and $BC_2F_2$ using the markers linked to QTLs for component traits of drought tolerance, such as NDVI, CT, TKW, and yield. NDVI sensor enables quick ground-level measurements of crops with the resolution required to characterize the canopy for its biomass, nutrient content, leaf area, and green area indices (Prasad et al., 2011). The SPAD meter measures leaf chlorophyll index *via* light transmittance that is differentially observed by chlorophyll and estimates leaf chlorophyll content and nitrogen content (Weier and Herring, 2000; Araus et al., 2008). DH, CT at the grain filling stage, and TKW are the most commonly used traits as indirect selection to improve yield under drought and heat stress (Tahmasebi et al., 2014). Similarly, Kernel number, grain yield, and chlorophyll content have been found to co-localize with regions controlling other drought adaptation traits, such as canopy temperature, and lead to deeper root system development and help in greater water absorption under stress (Pinto et al., 2010; Pinto and Reynolds, 2015; Diab et al., 2008; Olivares-Villegas et al., 2008). Therefore, in our study, three QTLs (BIO, WSC, and YLD: *Xgwm484*; CT and YLD: *Xwmc617*; TKW, CT, DM (MQTL24): *Xwmc640*) associated with component traits of drought tolerance were successfully transferred in the genetic background of HD3086 using MABB and phenotypic selection. The presence of introgressed QTLs located on wheat chromosomes 2D, 3A, and 4A explained a phenotypic variance from 9.45% to 14.01% previously validated in our lab using the RIL population HI1500 x DBW43, Harikrishna (2017) and other independent studies (Kadam et al., 2012; Acuña-Galindo et al., 2015; Gu et al., 2015). In the present study, a positive correlation between NDVI, LCI, and TKW with yield and a negative correlation between CT and DH with yield were observed. Similar correlation results were also reported by Lopes and Reynolds (2012), Harikrishna et al. (2016), and Ramya et al. (2016), indicating that productivity under moisture deficit stress was improved by the increased performance of component traits of drought tolerance, such as NDVI, chlorophyll content, and canopy temperature depression. In the current study, seven promising lines that perform better than their recurrent parent under RI conditions were developed. The two lines that possess QTL combinations for Biomass, TKW, and yield showed a 15-17% improvement in yield over recurrent parent HD3086. In a related earlier study by Rai et al. (2018), five prospective varieties were developed in the background of HD2733 by transferring NDVI, CT, and chlorophyll content

linked QTLs by MABB and were found to perform well under rainfed conditions. Similar to this, wheat variety GW322 was improved for traits such as NDVI, stay green, chlorophyll content, and yield, and 18 superior $BC_2F_3$ drought-tolerant progenies were identified in advanced generations of MABB-derived lines (Todkar et al., 2020). In this study, along with drought tolerance, lines were also screened for rust resistance in each generation, and the final improved lines showing resistance to both leaf and stripe rust were selected. The donor germplasm line DBW43 has the *Lr42* gene in their pedigree, and recurrent parent HD3086 showed resistance to 78S84 and 46S119 pathotypes of stripe rust (Singh et al., 2014); this might be the reason the drought-tolerant improved lines in cross HD3086*2/DBW43 also showed resistance to both leaf and stripe rust.

The potential applications of MAS in wheat have been further expanded by the advent of high-throughput sequencing, precision phenotyping, crop molecular physiology, and computational tools. Wheat has the biggest genome size (~16GB; Somers et al., 2004), and it is difficult to cover entire genomic areas with uniformly spaced polymorphic SSR markers. Background selection was performed using 35k SNPs in $BC_2F_5$ along with parents and identified plants with a maximum RPG ranging from 78%-91%. SNPs are almost 300 times more cost-effective than SSR markers for background analysis; therefore, SNPs could be a better choice for background selection in MABB (Khanna et al., 2015). Through the combination of phenotypic selection and marker-assisted background selection, the genome of the recurrent parent HD3086 recovered more quickly and to a greater extent compared to the estimated average recovery percentage. It has been observed in previous research that two backcrosses with selection for RPP were sufficient for better genome recovery in important cereal crops like wheat (Xu et al., 2017; Sharma et al., 2021; Mallick et al., 2022; Sunilkumar et al., 2022), rice (Ellur et al., 2016; Grover et al., 2020), and maize (Hossain et al., 2018; Zunjare et al., 2018). In our study, two drought-tolerant improved lines, HD3086-7-1-210-26 and HD3086-3-15-174-22, had maximum recovery rates of 92% and 90%, respectively. Similar genome recovery rates, i.e., 89.2% to 95.4%, were observed by Rai et al. (2018) in MABB for drought tolerance QTLs in the background of HD2733. In a study containing gene pyramiding of leaf rust resistance genes into an elite cultivar, HD2687, 94.55% genome recovery rates in $BC_2$ generations were observed by Bhawar et al. (2011). Similarly, 98.25% RPG recovery was observed in the transfer of gene *LrTrk* from *Triticum turgidum* cv. *Trinakria* to hexaploidy wheat variety HD2932 by Mallick et al. (2022) through MABB.

## Conclusion

The current work demonstrates the back-crossing breeding technique that combines phenotypic selection and marker-assisted selection to transfer QTLs related to drought tolerance in the background of the well-known wheat variety HD3086. We have successfully transferred three QTLs governing component traits under drought, viz, NDVI, CT, BIO, WSC, YLD, TKW, and DM, to develop improved lines that are performing well under drought stress using the MABB scheme. The transfer of QTLs led to the development of superior varieties in the genetic background of the existing variety HD3086. We have identified seven superior lines over the parent HD3086 under drought stress. The improved lines with more than 90% genomic similarity with the recurrent parent can be released as a superior variety over the existing variety HD3086 under restricted irrigation conditions. Millions of farmers in major wheat-growing regions choose wheat variety HD3086 because of its excellent yielding ability and quality. Furthermore, expanding the area of cultivation in NWPZ and NEPZ with limited irrigations is made possible by improving this variety for moisture deficit stress tolerance.

## Data availability statement

Data is uploaded to Dryad data base, available at following link. https://datadryad.org/stash/share/Bf6VjldDzhcIS0YNqyTRd_qVOllEckhCx1_-TkJK8YE.

## Author contributions

Conduct of experiment and writing of manuscript draft: VPS and HK; Supervision: PS, GS, and NJ; Phenotypic data collection: VPS, SS, DC, ND, and JS; Evaluation of rust: VPS, HK, PL, DP, and SM; Genotypic data: VPS, PS, HK, SS, DC, NS, and KM; Statistical analysis: VPS, ND, and HK. All authors contributed to the article and approved the submitted version.

## Funding

Part of this research is supported by a grant from the Bill and Melinda Gates Foundation (OPP1194767) for SNP genotyping and funding from the Department of Biotechnology (DBT) Govt. of India.

## Acknowledgments

VS acknowledges the ICAR-Indian Agriculture Research Institute (IARI), New Delhi for providing scholarships to complete this work as part of the Ph.D. thesis.

## Conflict of interest

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

## Publisher's note

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

## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1147200/full#supplementary-material

## References


Acuña-Galindo, M. A., Mason, R. E., Subramanian, N. K., and Hays, D. B. (2015). Meta-analysis of wheat QTL regions associated with adaptation to drought and heat stress. *Crop Sci.* 55 (2), 477–492. doi: 10.2135/cropsci2013.11.0793

Afzal, F., Reddy, B., Gul, A., Khalid, M., Subhani, A., Shazadi, K., et al. (2017). Physiological, biochemical and agronomic traits associated with drought tolerance in a synthetic-derived wheat diversity panel. *Crop Pasture Sci.* 68 (3), 213–224. doi: 10.1071/CP16367

Araus, J. L., Slafer, G. A., Royo, C., and Serret, M. D. (2008). Breeding for yield potential and stress adaptation in cereals. *Crit. Rev. Plant Sci.* 27, 377–412. doi: 10.1080/07352680802467736

Aravind, J., Mukesh Sankar, S., Wankhede, D. P., and Kaur, V. (2021) *augmentedRCBD: analysis of augmented randomised complete block designs*. Available at: https://aravindj.github.io/augmentedRCBD/https://cran.rproject.org/package=augmentedRCBD.

Asseng, S., Guarin, J. R., Raman, M., Monje, O., Kiss, G., Despommier, D. D., et al. (2020). Wheat yield potential in controlled-environment vertical farms. *Proc. Natl. Acad. Sci.* 117 (32), 19131–19135. doi: 10.1073/pnas.2002655117

Bahari, N., Karpisheh, L., and Bighdilu, B. B. (2014). The effect of drought stress on the traits related to the remobilization of wheat genotype in the cold region of ardabil. *Int. J. Biosci. (IJB)* 4 (6), 180–184.

Bapela, T., Shimelis, H., Tsilo, T. J., and Mathew, I. (2022). Genetic improvement of wheat for drought tolerance: progress, challenges and opportunities. *Plants* 11 (10), 1331. doi: 10.3390/plants11101331

Bayoumi, T. Y., Eid, M. H., and Metwali, E. M. (2008). Application of physiological and biochemical indices as a screening technique for drought tolerance in wheat genotypes. *Afr. J. Biotechnol.* 7 (14), 2141–2352.

Bhawar, K. B., Vinod, , Sharma, J. B., Singh, A. K., Sivasamy, M., Singh, M., et al. (2011). Molecular marker assisted pyramiding of leaf rust resistance genes Lr19 and Lr28 in bread wheat (Triticum aestivum l.) variety HD2687. *Indian J. Genet.* 71 (4), 304–311.

Bustos, A. D., Rubio, P., Soler, C., Garcia, P., and Jouve, N. (2001). "Marker assisted selection to improve HMW-glutenins in wheat," in *Wheat in a global environment* (Dordrecht: Springer), 171–176.

Collard, B. C., and Mackill, D. J. (2008). Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Philos. Trans. R. Soc Lond. B. Biol. Sci.* 363, 557–572. doi: 10.1098/rstb.2007.2170

Daryanto, S., Wang, L., and Jacinthe, P. A. (2016). Global synthesis of drought effects on maize and wheat production. *PloS One* 11 (5), e0156362. doi: 10.1371/journal.pone.0156362

Devate, N. B., Krishna, H., Parmeshwarappa, S. K. V., Manjunath, K. K., Chauhan, D., Singh, S., et al. (2022). Genome wide association mapping for component traits of drought and heat tolerance in wheat. *Front. Plant Sci.* 13, 943033. doi: 10.3389/fpls.2022.943033

Diab, A. A., Kantety, R. V., Ozturk, N. Z., Benscher, D., Nachit, M. M., and Sorrells, M. E. (2008). Drought-inducible genes and differentially expressed sequence tags associated with components of drought tolerance in durum wheat. *Sci. Res. Essay* 3, 009–026.

Dolferus, R. (2014). To grow or not to grow: a stressful decision for plants. *Plant Sci.* 229, 247–261. doi: 10.1016/j.plantsci.2014.10.002

Ellur, R. K., Khanna, A., Bhowmick, P. K., Vinod, K. K., Nagarajan, M., Mondal, K. K., et al. (2016). Marker-aided incorporation of Xa38, a novel bacterial blight resistance gene, in PB1121 and comparison of its resistance spectrum with xa13+ Xa21. *Sci. Rep.* 6 (1), 1–8. doi: 10.1038/srep29188

Farooq, M., Hussain, M., and Siddique, K. H. (2014). Drought stress in wheat during flowering and grain-filling periods. *Crit. Rev. Plant Sci.* 33 (4), 331–349. doi: 10.1080/07352689.2014.875291

Federer, W. T. (1961). Augmented designs with one-way elimination of heterogeneity. *Biometrics* 17 (3), 447–473. doi: 10.2307/2527837

Gautam, T., Saripalli, G., Kumar, A., Gahlaut, V., Gadekar, D. A., Oak, M., et al. (2020). Introgression of a drought insensitive grain yield QTL for improvement of four Indian bread wheat cultivars using marker assisted breeding without background selection. *J. Plant Biochem. Biotechnol.* 30 (1), 172–183. doi: 10.1007/s13562-020-00553-0

Grote, U., Fasse, A., Nguyen, T. T., and Erenstein, O. (2021). Food security and the dynamics of wheat and maize value chains in Africa and Asia. *Front. Sustain. Food Syst.* 4, 617009. doi: 10.3389/fsufs.2020.617009

Grover, N., Kumar, A., Yadav, A. K., Gopala Krishnan, S., Ellur, R. K., Bhowmick, P. K., et al. (2020). Marker assisted development and characterization of herbicide tolerant near isogenic lines of a mega basmati rice variety,"Pusa basmati 1121". *Rice* 13 (1), 1–13. doi: 10.1186/s12284-020-00423-2

Gu, A., Hao, P., Lv, D., Zhen, S., Bian, Y., Ma, C., et al. (2015). Integrated proteome analysis of the wheat embryo and endosperm reveals central metabolic changes involved in the water deficit response during grain development. *J. Agric. Food Chem.* 63 (38), 8478–8487. doi: 10.1021/acs.jafc.5b00575

Gu, L., Wei, B., Fan, R., Jia, X., Wang, X., and Zhang, X. (2015). Development, identification and utilization of introgression lines using chinese endemic and synthetic wheat as donors. *J. Integr. Plant Biol.* 57 (8), 688–697. doi: 10.1111/jipb.12324

Gupta, P. K., Balyan, H. S., and Gahlaut, V. (2017). QTL analysis for drought tolerance in wheat: present status and future possibilities. *Agronomy* 7 (1), 5. doi: 10.3390/agronomy7010005

Gupta, P. K., Balyan, H. S., Gahlaut, V., and Kulwal, P. L. (2012). Phenotyping, genetic dissection, and breeding for drought and heat tolerance in common wheat: status and prospects. *Plant Breed. Rev.* 36, 85–168. doi: 10.1002/9781118358566.ch2

Gupta, P. K., Kumar, J., Mir, R. R., and Kumar, A. (2010). 4 marker-assisted selection as a component of conventional plant breeding. *Plant Breed. Rev.* 33, 145. doi: 10.1002/9780470535486.ch4

Harikrishna, Singh, G. P., Jain, N., Singh, P. K., Prasad, S., Divya, A., et al. (2016). Physiological characterization and grain yield stability analysis of RILs under different moisture stress conditions in wheat (Triticum aestivum l.). *Ind. J. Plant Physiol.* 21, 576–582. doi: 10.1007/s40502-016-0257-9

Harikrishna (2017). *'Marker assisted recurrent selection (MARS) for drought tolerance in wheat'* (New Delhi: Indian Agricultural Research Institute).

Hossain, F., Muthusamy, V., Pandey, N., Vishwakarma, A. K., Baveja, A., Zunjare, R. U., et al. (2018). Marker-assisted introgression of opaque2 allele for rapid conversion of elite hybrids into quality protein maize. *J. Genet.* 97, 287–298. doi: 10.1007/s12041-018-0914-z

IDRC (2010). *Facts and figures on food and biodiversity* (Canada: IDRC Communications, International Development Research Centre). Available at: https://www.idrc.ca/en/research-in-action/facts-figures-food-and-biodiversity.

Imtiaz, M., Cromey, M. G., Hampton, J. G., and Ahmad, M. (2003). Inheritance of durable adult plant resistance to stripe rust (Puccinia striiformisf. sp. tritici) in 'Otane'wheat (Triticum aestivum). *New Z. J. Crop Hortic. Sci.* 31 (1), 23–31. doi: 10.1080/01140671.2003.9514232

IPCC (2021). *Climate change 2021: The physical science basis. contribution of working group i to the sixth assessment report of the intergovernmental panel on climate change* (Cambridge: Cambridge University), 1–394.

Joshi, A. K., Mishra, B., Chatrath, R., Ortiz Ferrara, G., and Singh, R. P. (2007). Wheat improvement in India: present status, emerging challenges and future prospects. *Euphytica* 157 (3), 431–446. doi: 10.1007/s10681-007-9385-7

Kadam, S., Singh, K., Shukla, S., Goel, S., Vikram, P., Pawar, V., et al. (2012). Genomic associations for drought tolerance on the short arm of wheat chromosome 4B. *Funct. Integr. Genomics* 12, 447–464. doi: 10.1007/s10142-012-0276-1

Khanna, A., Sharma, V., Ellur, R. K., Shikari, A. B., Gopala Krishnan, S., Singh, U. D., et al. (2015). Development and evaluation of near-isogenic lines for major blast resistance gene (s) in basmati rice. *Theor. Appl. Genet.* 128 (7), 1243–1259. doi: 10.1007/s00122-015-2502-4

Kirigwi, F. M., Van Ginkel, M. A. R. T. I. N., Brown-Guedira, G., Gill, B. S., Paulsen, G. M., and Fritz, A. K. (2007). Markers associated with a QTL for grain yield in wheat under drought. *Mol. Breed.* 20 (4), 401–413. doi: 10.1007/s11032-007-9100-3

Kumar, J., Jaiswal, V., Kumar, A., Kumar, N., Mir, R. R., Kumar, S., et al. (2011). Introgression of a major gene for high grain protein content in some Indian bread wheat cultivars. *Field Crops Res.* 123 (3), 226–233. doi: 10.1016/j.fcr.2011.05.013

Kumar, J., Mir, R. R., Kumar, N., Kumar, A., Mohan, A., Prabhu, K. V., et al. (2010). Marker-assisted selection for pre-harvest sprouting tolerance and leaf rust resistance in bread wheat. *Plant Breed.* 129 (6), 617–621. doi: 10.1111/j.1439-0523.2009.01758.x

Lopes, M. S., and Reynolds, M. P. (2012). Stay-green in spring wheat can be determined by spectral reflectance measurements (normalized difference vegetation index) independently from phenology. *J. Exp. Bot.* 63, 3789–3798. doi: 10.1093/jxb/ers071

Malik, N., Vinod, S. J. B., Tomar, R. S., Sivasamy, M., and Prabhu, K. V. (2015). Marker-assisted backcross breeding to combine multiple rust resistance in wheat. *Plant Breed* 134 (2), 172–177. doi: 10.1111/pbr.12242

Mallick, N., Jha, S. K., Agarwal, P., Kumar, S., Mall, A., Niranjana, M., et al. (2022). Marker-assisted transfer of leaf and stripe rust resistance from Triticum turgidum var. durum cv. trinakria to wheat variety HD2932. *Front. Genet.* 13. doi: 10.3389/fgene.2022.941287

Merchuk-Ovnat, L., Barak, V., Fahima, T., Ordon, F., Lidzbarsky, G. A., Krugman, T., et al. (2016). Ancestral QTL alleles from wild emmer wheat improve drought resistance and productivity in modern wheat cultivars. *Front. Plant Sci.* 7, 452. doi: 10.3389/fpls.2016.00452

Mondal, B., Singh, A., Yadav, A., Tomar, R. S. S., Singh, G. P., and Prabhu, K. V. (2017). QTL mapping for early ground cover in wheat (Triticum aestivum l.) under drought stress. *Curr. Sci.* 112 (6), 1266–1271.

Naeem, M. K., Ahmad, M., Kamran, M., Shah, M. K. N., and Iqbal, M. S. (2015). Physiological responses of wheat (Triticum aestivum l.) to drought stress. *Int. J. Plant Soil Sci.* 6 (1), 1–9. doi: 10.9734/IJPSS/2015/9587

Nalley, L., Dixon, B., Chaminuka, P., Naledzani, Z., and Coale, M. J. (2018). The role of public wheat breeding in reducing food insecurity in South Africa. *PloS One* 13 (12), e0209598. doi: 10.1371/journal.pone.0209598

Pandurangan Villegas, J. J., Reynolds, M. P., William, H. M., McDonald, G. K., and Ribaut, J. M. (2008). "Drought adaptation attributes and associated molecular markers *via* BSA in the Seri/Babax hexaploid wheat (*Triticum aestivum* l.) population," in *Proceedings of the 11th international wheat genetics symposium*, Brisbane, Australia, 24–29 Aug 2008 (Sydney: University Press).

Pask, A. J. D., Pietragalla, J., Mullan, D. M., and Reynolds, ,. M. P. (2012). *Physiological breeding II: A field guide to wheat phenotyping* (Mexico, D.F: CIMMYT). Available at: http://hdl.handle.net/10883/1288

Peterson, R. F., Campbell, A. B., and Hannah, A. E. (1948). A diagrammatic scale for estimating rust intensity on leaves and stems of cereals. *Can. J. Res.* 26 (5), 496–500. doi: 10.1139/cjr48c-033

Pinto, R. S., and Reynolds, M. P. (2015). Common genetic basis for canopy temperature depression under heat and drought stress associated with optimized root distribution in bread wheat. *Theor. Appl. Genet.* 128 (4), 575–585. doi: 10.1007/s00122-015-2453-9

Pinto, R. S., Reynolds, M. P., Mathews, K. L., McIntyre, C. L., Olivares-Villegas, J. J., and Chapman, S. C. (2010). Heat and drought adaptive QTL in a wheat population designed to minimize confounding agronomic effects. *Theor. Appl. Genet.* 121, 1001–1021. doi: 10.1007/s00122-010-1351-4

Portmann, F. T., Siebert, S., and Döll, P. (2010). MIRCA2000–global monthly irrigated and rainfed crop areas around the year 2000: a new high-resolution data set for agricultural and hydrological modeling. *Global biogeochemical cycles* 24 (1), GB1011. doi: 10.1029/2008GB003435

Prasad, P. V. V., Pisipati, S. R., Momčilović, I., and Ristic, Z. (2011). Independent and combined effects of high temperature and drought stress during grain filling on plant yield and chloroplast EF-tu expression in spring wheat. *J. Agron. Crop Sci.* 197 (6), 430–441. doi: 10.1111/j.1439-037X.2011.00477.x

Quarrie, S. A., Pekic Quarrie, S., Radosevic, R., Rancic, D., Kaminska, A., Barnes, J. D., et al. (2006). Dissecting a wheat QTL for yield present in a range of environments: from the QTL to candidate genes. *J. Exp. Bot.* 57 (11), 2627–2637. doi: 10.1093/jxb/erl026

Rai, N., Bellundagi, A., Kumar, P. K., Kalasapura Thimmappa, R., Rani, S., Sinha, N., et al. (2018). Marker-assisted backcross breeding for improvement of drought tolerance in bread wheat (Triticum aestivum l. *em Thell). Plant Breed.* 137 (4), 514–526. doi: 10.1111/pbr.12605

Ramya, P., Singh, G. P., Neelu, J., Singh, P. K., Manoj, K. P., Kavita, S., et al. (2016). Effect of recurrent selection on drought tolerance and related morpho-physiological traits in bread wheat. *PloS One* 11 (6), e0156869. doi: 10.1371/journal.pone.0156869

Rashid, U., Yasmin, H., Hassan, M. N., Naz, R., Nosheen, A., Sajjad, M., et al. (2022). Drought-tolerant bacillus megaterium isolated from semi-arid conditions induces systemic tolerance of wheat under drought conditions. *Plant Cell Rep.* 41 (3), 549–569. doi: 10.1007/s00299-020-02640-x

Roelfs, A. P. (1992). *Rust diseases of wheat: concepts and methods of disease management* (Mexico: CIMMYT).

Salarpour, M., Abdolshahi, R., Pakniyat, H., Heidari, B., and Aminizadeh, S. (2021). Mapping quantitative trait loci for drought tolerance/susceptibility indices and estimation of breeding values of doubled haploid lines in wheat (Triticum aestivum). *Crop Pasture Sci.* 72 (7), 500–513. doi: 10.1071/CP20526

Salarpour, M., Pakniyat, H., Abdolshahi, R., Heidari, B., Razi, H., and Afzali, R. (2020). Mapping QTL for agronomic and root traits in the Kukri/RAC875 wheat (Triticum aestivum l.) population under drought stress conditions. *Euphytica* 216, 1–19. doi: 10.1007/s10681-020-02627-5

Sallam, A., Hashad, M., Hamed, E. S., and Omara, M. (2015). Genetic variation of stem characters in wheat and their relation to kernel weight under drought and heat stresses. *J. Crop Sci. Biotechnol.* 18 (3), 137–146. doi: 10.1007/s12892-015-0014-z

Shah, L., Ali, A., Zhu, Y., Wang, S., Si, H., and Ma, C. (2017). Wheat resistance to fusarium head blight and possibilities of its improvement using molecular marker-assisted selection. *Czech J. Genet. Plant Breed* 53, 47–54. doi: 10.17221/139/2016-CJGPB

Shamuyarira, K. W., Shimelis, H., Tapera, T., and Tsilo, T. J. (2019). Genetic advancement of newly developed wheat populations under drought-stressed and non-stressed conditions. *J. Crop Sci. Biotechnol.* 22 (2), 169–176. doi: 10.1007/s12892-018-0262-0

Sharma, A., Srivastava, P., Mavi, G. S., Kaur, S., Kaur, J., Bala, R., et al. (2021). Resurrection of wheat cultivar PBW343 using marker-assisted gene pyramiding for rust resistance. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.570408

Shashikumara, P., Harikrishna, , Manu, B., Sunil, B., Sunilkumar, V. P., Nivedita, S., et al. (2020). Mapping genomic regions of moisture deficit stress tolerance using backcross inbred lines in wheat (Triticum aestivum l.). *Sci. Rep.* 10, 21646. doi: 10.1038/s41598-020-78671-x

Shukla, S., Singh, K., Patil, R. V., Kadam, S., Bharti, S., Prasad, P., et al. (2015). Genomic regions associated with grain yield under drought stress in wheat (Triticum aestivum l.). *Euphytica* 203 (2), 449–467. doi: 10.1007/s10681-014-1314-y

Singh, G. P., Prabhu, K. V., Singh, P. K., Singh, A. M., Jain, N., Ramya, P., et al. (2014). HD 3086: a new wheat variety for irrigated, timely sown condition of North Western plains zone of India. *J. Wheat Res.* 6 (2), 179–180.

Sinha, S. K., Aggarwal, P. K., Chaturvedi, G. S., Singh, A. K., and Kailasnathan, K. (1986). Performance of wheat and triticale cultivars in a variable soil–water environment i. grain yield stability. *Field Crops Res.* 13, 289–299. doi: 10.1016/0378-4290(86)90031-6

Sobhaninan, N., Heidari, B., Tahmasebi, S., Dadkhodaie, A., and McIntyre, C. L. (2019). Response of quantitative and physiological traits to drought stress in the SeriM82/Babax wheat population. *Euphytica* 215, 1–15. doi: 10.1007/s10681-019-2357-x

Somers, D. J., Isaac, P., and Edwards, K. (2004). A high-density microsatellite consensus map for bread wheat (Triticum aestivum l.). *Theor. Appl. Genet.* 109, 1105–1114. doi: 10.1007/s00122-004-1740-7

Sunilkumar, V. P., Krishna, H., Devate, N. B., Manjunath, K. K., Chauhan, D., Singh, S., et al. (2022). Marker assisted improvement for leaf rust and moisture deficit stress tolerance in wheat variety HD3086. *Front. Plant Sci.* 13, 4167. doi: 10.3389/fpls.2022.1035016

Tahmasebi, S., Heidari, B., Pakniyat, H., and McIntyre, C. L. (2016). Mapping QTLs associated with agronomic and physiological traits under terminal drought and heat stress conditions in wheat (Triticum aestivum l.). *Genome* 60 (1), 26–45. doi: 10.1139/gen-2016-0017

Tahmasebi, S., Heidari, B., Pakniyat, H., and Jalal Kamali, M. R. (2014). Independent and combined effects of heat and drought stress in the seri M82× babax bread wheat population. *Plant Breed.* 133 (6), 702–711.

Tilman, D., Balzer, C., Hill, J., and Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. *Proc. Natl. Acad. Sci.* 108 (50), 20260–20264. doi: 10.1073/pnas.1116437108

Todkar, L., Singh, G. P., Jain, N., Singh, P. K., and Prabhu, K. V. (2020). Introgression of drought tolerance QTLs through marker assisted backcross breeding in wheat (Triticum aestivum l.). *Indian J. Genet. Plant Breed.* 80 (02), 209–212. doi: 10.31742/IJGPB.80.2.12

Tyagi, S., Mir, R. R., Kaur, H., Chhuneja, P., Ramesh, B., Balyan, H. S., et al. (2014). Marker-assisted pyramiding of eight QTLs/genes for seven different traits in common wheat (Triticum aestivum l.). *Mol. Breed.* 34 (1), 167–175. doi: 10.1007/s11032-014-0027-1

Van Berloo, R. (2008). GGT 2.0: versatile software for visualization and analysis of genetic data. *J. Heredity* 99 (2), 232–236. doi: 10.1093/jhered/esm109

Wardlaw, I. F., and Willenbrink, J. (2000). Mobilization of fructan reserves and changes in enzyme activities in wheat stems correlate with water stress during kernel filling. *New Phytol.* 148 (3), 413–422. doi: 10.1046/j.1469-8137.2000.00777.x

Weier, J., and Herring, D. (2000). Measuring vegetation (NDVI and EVI). *NASA Earth Observatory* 20 (2).

Xu, H., Cao, Y., Xu, Y., Ma, P., Ma, F., Song, L., et al. (2017). Marker-assisted development and evaluation of near-isogenic lines for broad-spectrum powdery mildew resistance gene Pm2b introgressed into different genetic backgrounds of wheat. *Front. Plant Sci.* 8, 1322. doi: 10.3389/fpls.2017.01322

Yang, H., Hu, W., Zhao, J., Huang, X., Zheng, T., and Fan, G. (2021). Genetic improvement combined with seed ethephon priming improved grain yield and drought resistance of wheat exposed to soil water deficit at tillering stage. *Plant Growth Regul.* 95 (3), 399–419. doi: 10.1007/s10725-021-00749-x

Yang, D. L., Jing, R. L., Chang, X. P., and Li, W. (2007). Identification of quantitative trait loci and environmental interactions for accumulation and remobilization of water-soluble carbohydrates in wheat (Triticum aestivum l.) stems. *Genetics* 176, 571–584. doi: 10.1534/genetics.106.068361

Yasmin, H., Nosheen, A., Naz, R., Keyani, R., and Anjum, S. (2019). "Regulatory role of rhizobacteria to induce drought and salt stress tolerance in plants," in *Field crops: sustainable management by PGPR* (Islamabad, Pakistan: Department of Biosciences, COMSATS University Islamabad (CUI)), 279–335.

Zadoks, J. C., Chang, T. T., and Cal Konzak, C. F. (1974). A decimal code for the growth stages of cereals. *Weed Res.* 14 (6), 415–421. doi: 10.1111/j.1365-3180.1974.tb01084

Zunjare, R. U., Hossain, F., Muthusamy, V., Baveja, A., Chauhan, H. S., Bhat, J. S., et al. (2018). Development of biofortified maize hybrids through marker-assisted stacking of β-carotene hydroxylase, lycopene-ϵ-cyclase and opaque2 genes. *Front. Plant Sci.* 9, 178. doi: 10.3389/fpls.2018.00178

# Frontiers in Genetics

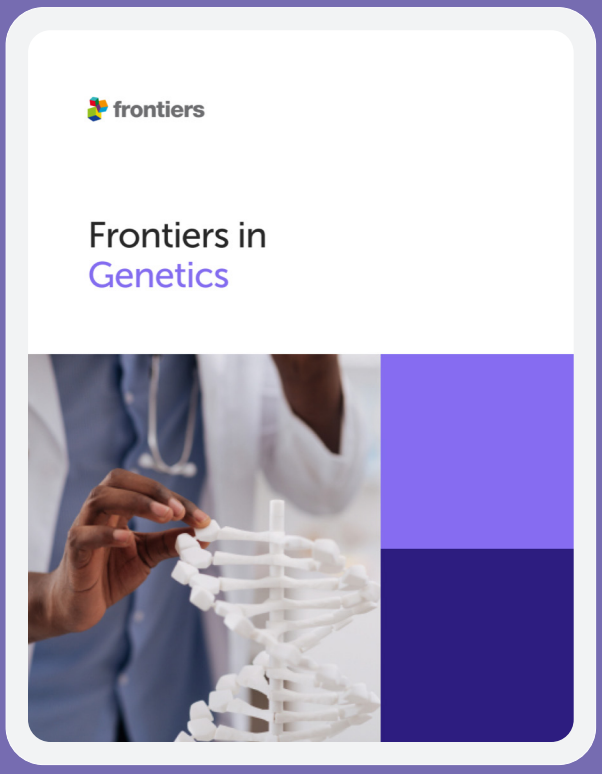
**Highlights genetic and genomic inquiry relating to all domains of life**

The most cited genetics and heredity journal, which advances our understanding of genes from humans to plants and other model organisms. It highlights developments in the function and variability of the genome, and the use of genomic tools.

---

## Discover the latest Research Topics

See more →




**Frontiers**

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

**Contact us**

+41 (0)21 510 17 00
frontiersin.org/about/contact